nature cities

Article

https://doi.org/10.1038/s44284-024-00060-6

War city profiles drawn from satellite images

Received: 27 October 2023

Accepted: 14 March 2024

Published online: 09 April 2024

Check for updates

Zhengyang Hou[1,7], Ying Qu[1,7], Liqiang Zhang[1] ✉, Jun Liu[2], Faqiang Wang[2], Qiwei Yu[1], An Zeng[3], Ziyue Chen[1], Yuanyuan Zhao[4], Hong Tang[1], Yuebin Wang[1], Xingang Li[1], Yang Li[1], Shuwen Peng[1], Jing Ran[5], Xin Yao[1], Xichen Meng[1], Suhong Liu[1] & Chenghu Zhou[6] ✉

The extent of war-induced destruction in urban areas is critical information for international relief efforts, impact assessments and restoration decisions. However, precise geotargeting of zones with severe destruction is still a great challenge. Here we present a novel temporal-knowledge-guided detection scheme (TKDS) with a pixel-based transformer network (PtNet) for monitoring urban destruction using satellite imagery, applied to conflict zones in the Syrian civil war and the Russia–Ukraine conflict. Compared with state-of-the-art methods, the TKDS-PtNet model enhances war damage identification by 44.0 (72.5 versus 28.5) in the F1 score for six Syrian cities and 34.2 (83.5 versus 49.3) for four Ukrainian cities. The identified damaged buildings are further utilized to estimate the affected population and damage to critical infrastructures such as hospitals and schools in these areas. Our results demonstrate the high potential of a repeatable and relatively low-cost scheme for the near real-time monitoring of damage in urban areas resulting from wars, earthquakes or extreme weather events. Our findings underscore the crucial importance of taking action to stop the conflict and developing mechanisms to prevent present and future urban-related damage from military actions.

War in urban areas is one of the main threats to development and well-being, yet over 70 ongoing conflicts, including the Russian–Ukraine conflict, have resulted in devastating destruction, pain and disbelief in cities[1,2]. The widespread devastation of urban buildings resulting from heavy weapon attacks manifests in injuries, fatalities and the collapse of essential services such as water and power supplies as well as hospitals. This not only results in severe destruction and fragmentation within the cities where it occurs but also hinders the sustained maintenance of international peace and governance. In this context, prompt and accurate information regarding the destruction within urban landscapes is vital to understanding the conflict's scale and adherence to international humanitarian law, and becomes instrumental in formulating more effective strategies for urban conflict resolution and reconstruction. Such knowledge also serves as crucial evidence supporting decision-making in humanitarian aid and guides mitigation responses to alleviate the impact on affected populations.

Monitoring building destruction from conflict cities with timeliness and accuracy remains a formidable challenge[1,3–6]. The violence and unsafe conditions in conflict regions inhibit traditional damage assessment methods, such as manual detection and on-site survey during wartime[5,7–9]. Nowadays, satellite imagery analysis emerges as a promising tool for remote assessment of infrastructure destruction. Deep learning-based methods have demonstrated unique strengths in destruction identification automatically[9–13]. However, it is still difficult to precisely geotarget zones with severe destruction due to intricate urban environments. Moreover, previous methods primarily focus on evaluating damage caused by natural disasters[8,14–19], where destruction usually shows strong geographical patterns and correlation features.

[1]State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing, China. [2]School of Mathematical Sciences, Beijing Normal University, Beijing, China. [3]School of Systems Science, Beijing Normal University, Beijing, China. [4]College of Land Science and Technology, China Agricultural University, Beijing, China. [5]School of Architecture and Planning, Hunan University, Changsha, China. [6]State Key Laboratory of Resources and Environment Information System, Institute of Geographical Science and Natural Resources, Chinese Academy of Sciences, Beijing, China. [7]These authors contributed equally: Zhengyang Hou, Ying Qu. ✉e-mail: zhanglq@bnu.edu.cn; zhouch@lreis.ac.cn

https://doi.org/10.1038/s44284-024-00060-6

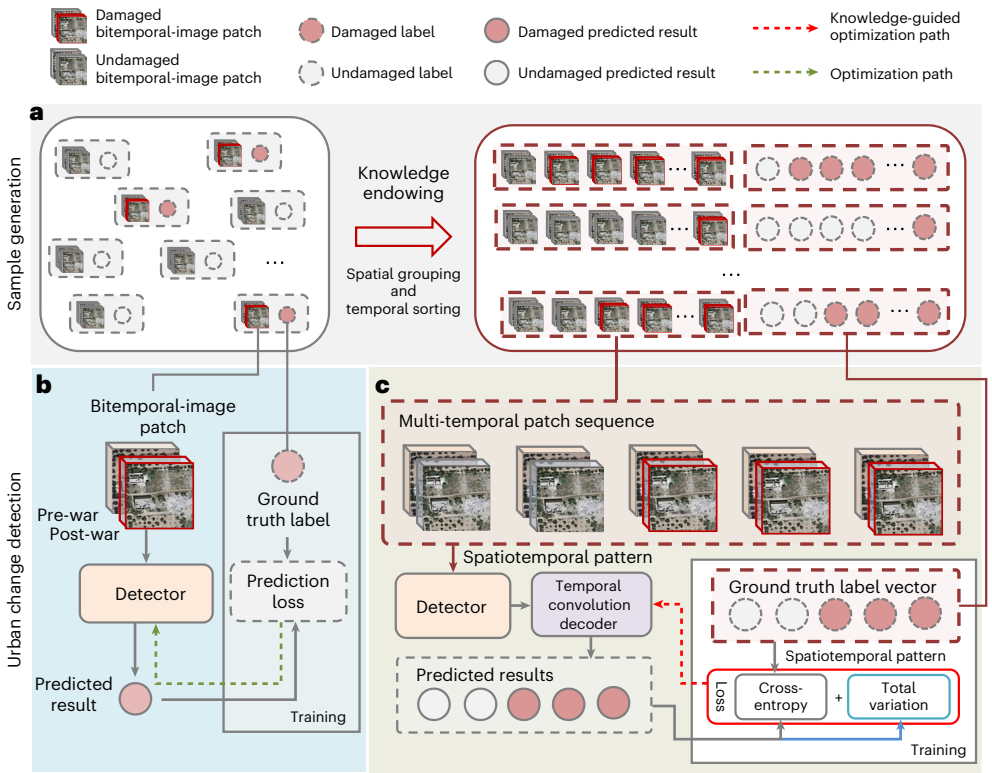

Fig. 1 | The comparison between the commonly used urban change detection scheme and TKDS. The detector can be PtNet, ResNet or any other detection methods. **a**, Generation of the image patch samples. **b**, The commonly used destruction detection methods. **c**, The proposed TKDS scheme.

In contrast, buildings destroyed in wars are often sparsely distributed in urban environments, occupying only a small proportion of urban buildings, with the majority of the landscape remaining undamaged. This leads to an extreme class imbalance, where undestroyed buildings far outnumber the destroyed ones[5]. Such a class imbalance combined with the heterogeneity of urban environments poses a major obstacle, explaining why the assessment of destruction across entire cities has not been thoroughly explored[5,20].

Another challenge lies in the resolution of satellite imagery. So far, very few methods demonstrate the potential to precisely detect war damage automatically[5,7,21]. In addition, their deployment relies on high-resolution satellite imagery (less than 1 m). Due to real-time capability limitations and confidentiality agreements, such high-resolution satellite data during wartime are not publicly available. Notably, medium-resolution satellite images (such as 10 m) available to the public have the advantages of worldwide coverage and timely revisit frequencies. However, it remains a challenge to assess war damage with medium-resolution images, since each building may only occupy less than 5 pixels (10 m × 10 m), meaning that the detailed information about building shape and textures is missing (Supplementary Fig. 1). Furthermore, influenced by illumination variations, a building captured at different times may have distinct colors on the images. Such a temporal domain shift tends to substantially increase the false positive rate (Supplementary Technical Terms).

This, naturally, raises a question: is it possible to accurately monitor war destruction from medium-resolution (such as 10 m) satellite images, especially when there exists an extreme level of class imbalance? Here, we propose a new temporal-knowledge-guided detection scheme (TKDS), which can monitor urban destructions with either high- or medium-resolution satellite imagery. TKDS applies the knowledge that a war in urban may last several weeks or even months[5,22–24], indicating that if an urban block is destroyed at a certain time, it is unlikely to be reconstructed during a war. Integrating such knowledge into deep learning model has the potential to greatly reduce the false positive rate caused by low resolution and class imbalance (Methods). Compared with the commonly used change detection schemes[5,12,14–16], our TKDS is a flexible scheme (Fig. 1), which can regularize the detector with the spatial and temporal patterns of the building change. A pre-war image and each of the images available during the war are stacked to form a group of bitemporal image patches. Then, the bitemporal-image patches at a given location are sorted by date to construct a multi-temporal patch sequence as input to TKDS (Methods). To extract the semantic vector representation of each bitemporal-image patch, a pixel-based transformer network (PtNet) is developed (Methods and Fig. 2). A temporal total variation (TTV) regularization function incorporating building-damaged temporal patterns is utilized as a temporal global constraint to guide PtNet to learn the damage patterns (Methods). TKDS further uses a temporal convolution decoder (TCD) to obtain the contextual relationship of the bitemporal-image patches in the temporal dimension based on the semantic vectors extracted by PtNet, and generates the detection results.

We compare our model TKDS-PtNet with a previous method CNN-STS[5](convolutional neural network integrated spatial–temporal smoothing; Supplementary Note and and Supplementary Fig. 2), revealing that TKDS-PtNet far exceeds the monitoring performance of CNN-STS. To the best of our knowledge, this is the first approach with high generalization abilities to precisely monitor urban destruction using different-resolution satellite imagery. As illustrated in Fig. 3, the TKDS-PtNet model is then employed to monitor damage in six Syrian cities and four Ukrainian cities. It demonstrates outstanding damage assessment capability in these cities from high-resolution or medium-resolution remote sensing images. It allows us to map the spatial distribution of destroyed buildings, identify where buildings are most severely damaged, and further reveal the affected populations and their potential impacts. The contributions of this study are threefold. (1) We propose a new temporal knowledge-guided detection scheme, TKDS,

https://doi.org/10.1038/s44284-024-00060-6

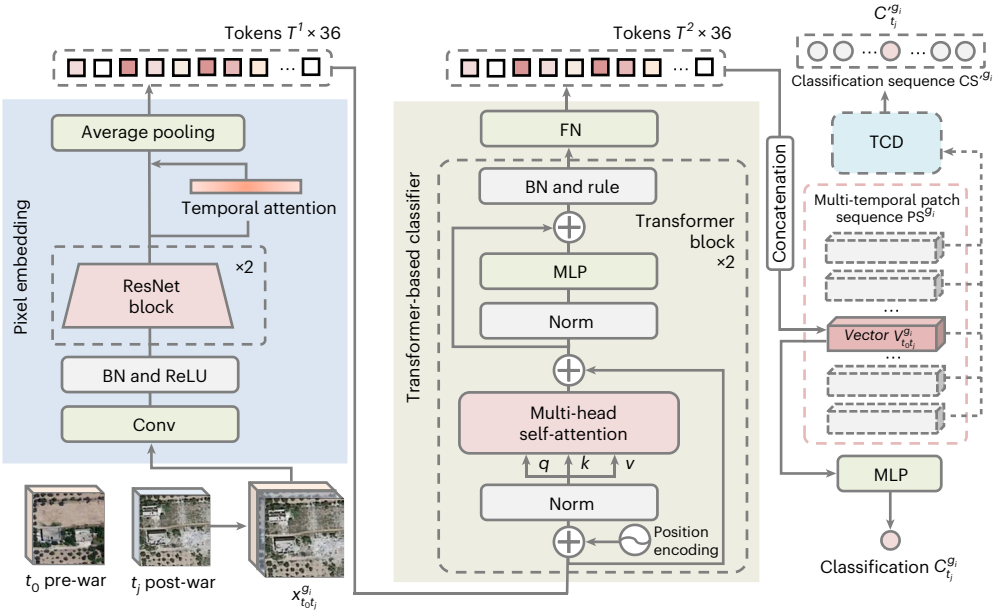

Fig. 2 | The TKDS-PtNet architecture. BN, FN and MLP stand for batch normalization, feature normalization, and multilayer perceptron, respectively.

that can be flexibly integrated with any traditional detectors to improve their detection performance. (2) We develop a semi-supervised domain adaptation (SSDA) strategy that combines maximum mean discrepancy (MMD)[25] loss and supervised contrastive learning (SCL)[26] (Supplementary Technical Terms) to improve the generalization ability of the proposed TKDS-PtNet model. (3) The TKDS-PtNet model offers a rapid and practical way to detect urban destruction in data-sparse and inaccessible environments. It allows us to track the destruction of urban critical infrastructure with a high degree of accuracy and has the potential to be applied to urban contexts to investigate war damage, earthquakes or some major weather events, and estimate the scope and distribution of damage to guide relief efforts.

Results

As shown in Fig. 3, we comprehensively validated the proposed approach and analyzed the impacts of war building damage.

Urban damage monitoring

Damage detection in the Syrian civil war. From 2011 to 2018, the Syrian civil war resulted in substantial casualties, refugees and severe urban damage. The TKDS-PtNet model was employed to detect damaged buildings in six Syrian cities (Aleppo, Hama, Homs, Raqqa, Deir-Ez-Zor and Idlib) using remote sensing images with 0.5 m and 10 m resolutions.

Using the completely damaged building samples labeled by the United Nations Operational Satellite Applications Program (UNOSAT) of the United Nations Institute for Training and Research (UNITAR)[27], we constructed two groups of datasets covering the six Syrian cities. The first group of samples includes 0.5-m-resolution bitemporal-image patches collected from the Worldview satellite imagery. To evaluate the performance of the TKDS-PtNet model on medium-resolution satellite images, we downsampled each bitemporal-image patch with a factor of 20 to generate the second dataset with 10-m-resolution remote sensing imagery (Methods). Based on the two datasets, we built two multi-temporal patch sequences with spatial resolutions of 0.5 m and 10 m. Then, the training and validation samples were generated to tune the TKDS-PtNet model (Methods).

The number of damaged buildings (positive samples) in each city was very small, and their spatiotemporal distribution was uneven. As listed in Supplementary Table 1, the proportion of the positive samples in the bitemporal image patch dataset was 3.9%. Accordingly, the temporal positive samples accounted for 7.3% in the multi-temporal patch sequence dataset. Five matrices, that is, precision, recall, F1 score, the area under the curve (AUC) and average precision (AP) (Supplementary Technical Terms), were used to evaluate the performance of damage detection. We observe that the commonly used deep learning model ResNet-18 can predict 35.2% of the damaged buildings, and 51.8% of the buildings were misidentified as damaged buildings. Thus, the F1 score for ResNet-18 was only 40.7 (Table 1 and Supplementary Fig. 3).

To evaluate the advantage of TKDS, we incorporated PtNet as well as ResNet-18 and ResNet-50 into the scheme, respectively. As listed in Table 1, with the 0.5-m-resolution images, TKDS not only improved the accuracy of PtNet by 31.5 (TKDS-PtNet) in terms of F1 score, but also substantially enhanced the performance of ResNet-50 and ResNet-18 by 43.2 (TKDS-ResNet-50) and 44.5 (TKDS-ResNet-18), respectively. The TKDS-PtNet model further improved the F1 score of PtNet from 54.7 to 86.1. The precision–recall curve and receiver operating characteristic curve (Supplementary Fig. 3 and Supplementary Technical Terms) confirmed the high performance of TKDS-PtNet. As displayed in Table 1 and Supplementary Fig. 3, all the models incorporating with TKDS showed notable performance improvements.

We also compared TKDS-PtNet with CNN-STS[5]. As listed in Table 1, on the 0.5 m satellite image dataset, although CNN-STS achieved higher precision than CNN by 29.7, it had the lowest recall (19.7). This reveals that a large number of destroyed buildings were identified as nondestroyed buildings by CNN-STS. Notably, TKDS-PtNet far exceeded the monitoring performance of CNN-STS (86.14 versus 28.26 in terms of F1 score). This is because CNN-STS is a two-stage machine learning model. In the first stage, a CNN classifier predicted damaged buildings without incorporating any prior knowledge, leading to low precision. In the second stage, the real damage patterns are difficult to learn from the low-precision results of the first stage, which made it easy to overlook some isolated damaged areas. As a comparison, TKDS-PtNet is an end-to-end model that employed TCD and TTV constraints to incorporate the temporal knowledge at both local and global levels, which outperformed CNN-STS substantially in building damage detection performance. Please see Supplementary Note and Supplementary Fig. 2 for more details.

The six cities are different in scale, building type and landscape, and their building sample and image availability differ dramatically (Supplementary Table 1). In such cases, TKDS-PtNet achieved F1 scores

https://doi.org/10.1038/s44284-024-00060-6

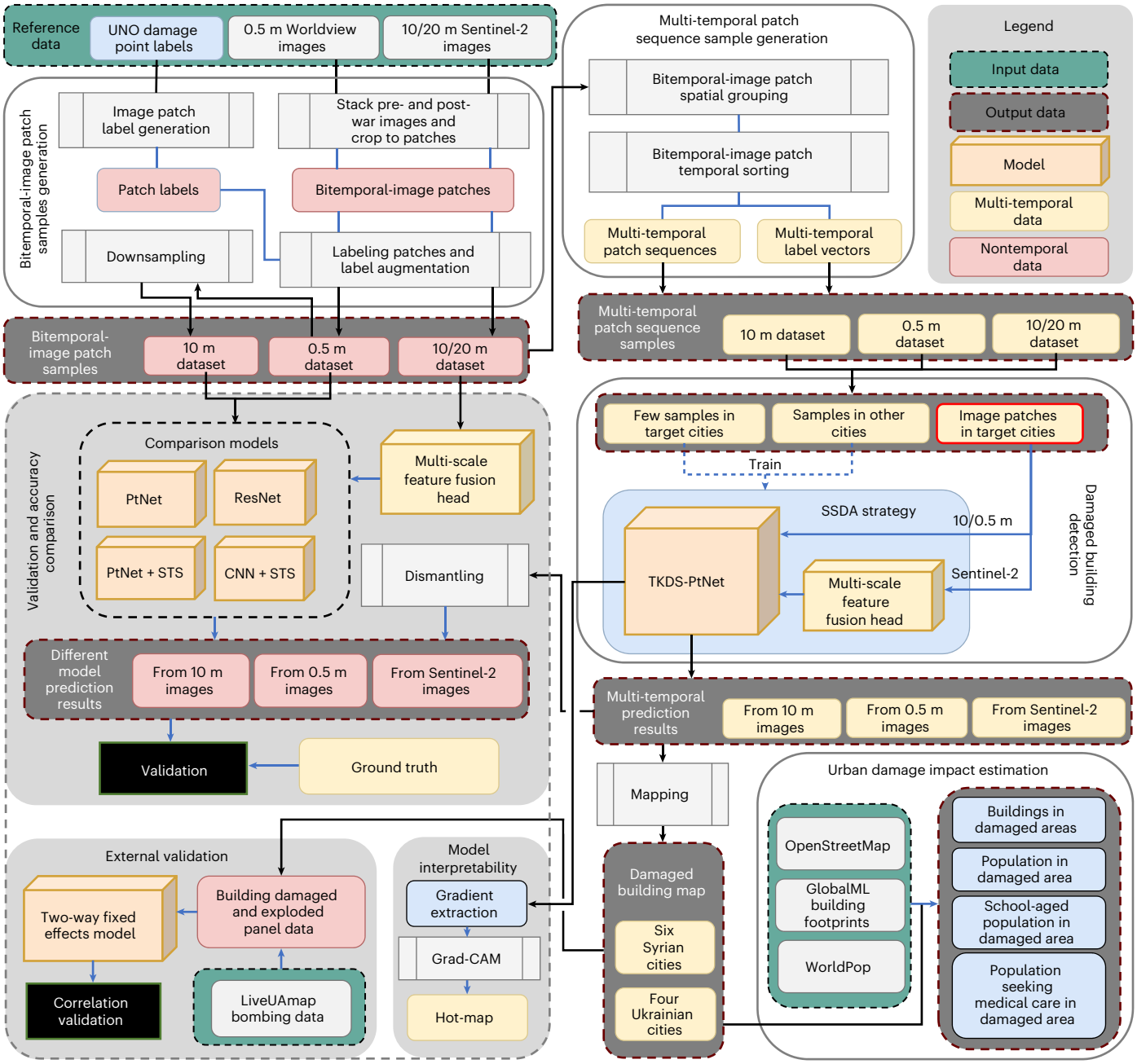

Fig. 3 | The urban destruction monitoring workflow.

of 72.5 and 86.1 with 10-m-resolution and 0.5-m-resolution satellite images, and had the greatest accuracy improvement compared with PtNet in Idlib and Homs cities (Supplementary Table 2). The accuracy was six times better than ResNet-50 (5.69). In other cities like Homs, Hama and Raqqa with relatively high positive sample proportions (10.2%, 10.0% and 7.4%, respectively, as listed in Supplementary Table 1), TKDS-PtNet achieved reliable accuracies (F1 >70) even with the 10-m-resolution images, which were comparable to the results detected with the 0.5-m-resolution images.

Urban damage monitoring in the Russia–Ukraine conflict. Compared with the over-10-year-long Syrian civil war, the intensity and density of building damage triggered by the Russia–Ukraine conflict in Ukrainian cities were relatively low. Using publicly available 10-m-resolution Sentinel-2 satellite images, we monitored the building damage in four cities in Ukraine, that is, Rubizhne, Sievierodonetsk, Volnovakha and Mariupol (Supplementary Fig. 4). The damage assessment in the cities has three difficulties. First, the damage monitoring faced an extreme level of sample class imbalance. Second, high-resolution satellite imagery of the post-war Ukrainian cities was not available. Third, urban building styles were more diverse and complex.

To improve the detection accuracy from Sentinel-2 satellite images, a multi-scale feature fusion head was incorporated into PtNet to better utilize the four 10-m-resolution bands and six 20-m-resolution bands of the images (Methods). As listed in Table 1, compared with the detection results from the three 10-m-resolution image bands, the F1 score of PtNet using the fused images was enhanced by 24.0 (64.1 versus 40.1), much higher than ResNet (F1 of 49.3 in ResNet-18, and F1 of 48.5 in ResNet-50), revealing that PtNet could automatically extract the spectral information of satellite images.

https://doi.org/10.1038/s44284-024-00060-6

Table 1 | The performance of the models in Syrian cities using 0.5-m-resolution and 1-0m-resolution satellite images, and using multi-spectral Sentinel-2 satellite images in Ukrainian cities

Country	Image	Model	Precision	Recall	F1	AUC	AP
		ResNet-50	48.62	34.70	40.50	84.22	39.73
		ResNet-18	48.23	35.16	40.67	86.21	39.83
		CNN	20.60	33.75	25.58	81.27	23.29
		CNN-STS	50.29	19.65	28.26	86.77	33.17
	0.5 m	PtNet	55.55	53.83	54.68	88.15	56.36
		PtNet-STS	88.45	41.06	56.09	92.54	68.33
		TKDS-ResNet-50	90.49	77.84	83.69	97.84	88.39
Syria		TKDS-ResNet-18	89.84	80.91	85.14	98.40	90.24
		TKDS-PtNet	**92.05**	**80.95**	**86.14**	**98.52**	**90.77**
	10 m	ResNet-50	24.78	33.38	28.52	80.74	24.31
		ResNet-18	24.95	31.92	28.00	80.14	23.69
		PtNet	33.90	44.70	38.56	85.57	38.64
		TKDS-ResNet-50	59.64	75.08	66.48	96.23	77.92
		TKDS-ResNet-18	62.66	73.70	67.73	95.72	77.31
		TKDS-PtNet	**72.19**	**72.17**	**72.48**	**96.25**	**79.29**
		PtNet (band 2-4)	34.87	47.24	40.13	90.13	37.99
		ResNet-50 (all bands)	47.55	49.39	48.45	92.12	49.85
Ukraine	Sentinel-2	ResNet-18 (all bands)	63.15	40.37	49.25	90.85	50.23
		PtNet (all bands)	60.68	67.98	64.12	90.48	69.06
		TKDS-PtNet (all bands)	**78.81**	**88.77**	**83.50**	**98.08**	**91.46**

Bold format indicates the best accuracy.

The TKDS-PtNet model further increased the F1 score (83.5) by 19.4 compared with PtNet (64.1). As shown in Supplementary Fig. 5, TKDS-PtNet had much higher detection accuracy in terms of AP and AUC (AP of 91.5 and AUC of 98.1) than ResNet-18 (AP of 50.2 and AUC of 90.9) and ResNet-50 (AP of 49.9 and AUC of 92.1). To demonstrate the robustness of TKDS-PtNet, we conducted a tenfold cross-validation on the Sentinel-2 satellite imagery. As shown in Supplementary Fig. 6, it achieved high performance as average F1 of 86.2, AUC of 98.6 and AP of 92.2.

Domain transfer capability, interpretability and validation

Model domain transfer capability. Under the SSDA strategy, we used zero or 5% of the samples in the target city to validate the domain transfer of TKDS-PtNet trained with the SSDA strategy (Methods). As listed in Table 2, in the four Ukrainian cities, the F1 score of TKDS-PtNet decreased by 27.7 on average compared with the training test using 70% of the samples in all cities (Table 2, column 3) but it was still higher than ResNet-50 in Rubizhne, Sievierodonetsk and Volnovakha cities. Then we randomly added 5% of the samples from the target city. The test results across the four cities exhibited only an average accuracy loss of 5.9 (ranging from 1.2 to 12.2) compared with 70% of the samples in all cities used to train the model (Table 2, column 3) For Sievierodonetsk with an extreme level of class imbalance (positive sample ratios of 0.4% in the bitemporal-image patches; Supplementary Table 2), the F1 score decreased by 5.3. In Volnovakha where the positive sample ratio was relatively high (3.3%), the accuracy loss was as low as 1.2. Thus, our model had a good capability that was robust against severe class imbalance. Using zero or a few labels in the target city, it achieved high performance in building damage detection.

External validation. To ensure the certainty of our detection results, we performed manual verification. Specifically, we validated our detection results through high-spatial-resolution images and media reporting. In the six Syrian cities, we first performed the manual verification through high spatial resolution images, and then searched the media reporting including the war photos and texts about the four cities to validate our detected damage. In the four Ukrainian cities, we conducted a similar verification except without using high-spatial-resolution images due to these data shortage. Next, we used a two-way fixed effects model to validate the reliability of temporal sequence detections for building damage leveraging external bombing events from the Live Universal Awareness Map project (LiveUAmap)[28]. Over 1.3 million predictions from Syria and over 2 million from Ukraine were used in the regression model. To test whether TKDS-PtNet could obtain more positive pre-war predictions after bombing events, the times of 940 bombing events in the six Syrian cities and 102 bombing events in the four Ukrainian cities were geographically marked in the patch image temporal sequences (Methods). As listed in Supplementary Table 3, the results showed a strong positive association between our detections and bomb events. As shown in Supplementary Fig. 7, the predicted values of TKDS-PtNet for Syria and Ukraine after the bomb events increased by 112% and 440%, compared with the baseline (the average of the detection values), indicating that our detections are reliable.

Model interpretability. We used the gradient-weighted class activation mapping (Grad-CAM; Supplementary Technical Terms)[29] as a visualization technique to understand which parts of an image were important for each detection prediction. The Grad-CAM functions identify the areas of an image that had the most impact on the output of the model by creating a heatmap.

Supplementary Fig. 8 shows the activation values of damaged and undamaged buildings using ResNet-50, PtNet and TKDS-PtNet, respectively, on 0.5-m-resolution images in Aleppo. Both ResNet-50 and PtNet exhibited misclassifications. The attention of TKDS-PtNet

https://doi.org/10.1038/s44284-024-00060-6

Table 2 | Performance and generalization of the models in Ukrainian cities

City	(1)		(2)		(3)		(4)		(5)	
	ResNet-50		PtNet		TKDS-PtNet		Generalization (0%)		Few-shot (5%)	
	F1	AUC	F1	AUC	F1	AUC	F1	AUC	F1	AUC
Mariupol	65.53	93.62	74.86	92.87	75.31	97.13	42.63	85.25	70.27	90.62
Rubizhne	43.60	93.12	61.04	91.75	90.40	99.71	56.11	96.49	78.24	98.92
Sievierodonetsk	17.09	76.42	15.51	69.33	42.46	88.30	22.74	91.32	37.12	93.71
Volnovakha	58.26	94.35	81.56	90.26	94.38	99.05	70.21	96.90	93.10	99.59
All	48.45	92.12	64.12	90.48	83.50	98.08				

The samples in each city were split into training and testing sets with a proportion of 70% and 30%, respectively. To test the performance of the models, the TKDS-PtNet model was trained on the training sets of the four cities (columns 1–3). To test the generalization ability in each city, TKDS-PtNet was trained on all the samples from the other three cities. In the few-shot building damage detection, TKDS-PtNet was trained using the SSDA strategy. All samples from the three cities were fed into the network, but only 5% of the annotated samples from the target city were retained. The remaining 95% of the annotated samples were not used in the training procedure. TKDS-PtNet was used to detect building damage for the given cities after 20 training epochs (columns 4 and 5). The performance and generalization ability in each city were tested on their respective testing sets.

on damaged areas was consistent and correct, meaning that TKDS empowered PtNet to incorporate multi-temporal spatial features. Consequently, the activation maps accurately highlighted the range of the damaged areas, leading to marked performance improvement in urban damage monitoring tasks.

Spatial distribution of the destructions and their impacts

We constructed war urban destruction maps from the detection results for facilitating evaluation of the war's impact on population and city life. Before this, we manually performed the verification of the TKDS-PtNet model output to ensure that the information about the spatial distribution and scale of the most severely damaged areas in the Syrian civil war and the Russian–Ukrainian conflict was accurate. On the maps, we overlapped detection maps with population and critical infrastructure data from international data sources, including WorldPop[30], the OpenStreetMap[31] and Microsoft's GlobalMLBuildingFootprints[32] datasets. The population in the damaged zones was estimated separately using the maximum and minimum values of 2018–2020 and 2011–2018 WorldPop population grids in Ukraine and Syria. The assessment of the war's impact on city life, including healthcare and education, was gauged by considering the population and school-aged children (aged 6–12 in Syria and 6–18 in Ukraine) within the 1 km radius around the damaged hospitals and schools.

Impacts from the Syrian civil war. The Syrian civil war since 2011 has triggered substantial urban damage. The TKDS-PtNet model identified a large number of completely damaged buildings and the presence of bomb craters. As shown in Fig. 4, approximately 53,700 buildings in the six Syrian cities were located in the identified damaged areas, where about 116,800–158,600 people lived. Aleppo, Hama, Homs and Deir-Ez-Zor cities had continuous damage areas in a strip-like pattern, indicating that these cities have experienced intense bombing. Aleppo, as the largest city in Syria and one of the oldest human settlements in the world, had a considerable number of landmark cultural and heritage buildings. Unfortunately, it experienced severe destruction, with 38,514 buildings, including 23 schools and 6 hospitals as well as very large portions of the historic buildings, and about 78,300–103,400 people in damaged areas. In Homs and Hama cities, destruction was clustered heavily in some neighborhoods. These areas were densely populated, which might result in massive casualties. There were 6,740 buildings, including 1 school and 1 hospital, and 26,100 people in the damaged areas in Homs City, ranking as Syria's third-largest city. Such damage to infrastructure might have additional repercussions on healthcare and education. In the six Syrian cities, we identified 9 hospitals and 39 schools in damaged areas, potentially leading to a shortage of healthcare services for about 151,700–203,500 people and a disruption in education of about 19,400–26,400 school-aged children.

Impacts from the Russia–Ukraine conflict. Compared with the damage inflicted by the Syrian civil war, urban destruction resulting from the Russia–Ukraine conflict shows more clustered spatial pattern and more severe damage to industrial buildings. In the four Ukrainian cities, as is shown in Fig. 4, about 3,850 buildings were located in the damaged zones, and about 3,971–4,192 people may be directly affected during the war. There were three schools (no hospitals) in identified damaged zones, potentially leading to a disruption in education of about 912–960 school-aged children. Among the four Ukrainian cities, Mariupol City experienced the most severe building damage, as shown in the building damage map of Mariupol City on 28 May 2022 (Supplementary Fig. 9). The majority (73.41%) of the damaged areas were concentrated in the central and southern areas of Mariupol, which include Livoberezhnyi, Zhovtnevyi and Azovstal areas. Notably, in addition to substantial damage to urban houses and apartments, Mariupol experienced severe damage to 66 industrial buildings and 11 garages. Specifically, the Azovstal zone, a major industrial area in Mariupol, witnessed damage to 31 industrial buildings, while Livoberezhnyi also identified damage to 18 industrial buildings. As an important industrial city in Ukraine, Rubizhne had more than 1,200 buildings located in the damaged zones, including 1 school and 22 industrial buildings, while Sievierodonetsk had about 200 buildings in the damaged zones, 19 of which were industrial buildings.

Discussion

To uncover war city profiles and further avoid catastrophe, remote sensing imagery is an alternative cost-effective and noninvasive option to gather a complete view of urban destruction[7]. TKDS offers a reliable and rapid way to monitor urban destruction using high- or medium-resolution satellite imagery and limited samples. The scheme is highly flexible and independent, such that its backbone model can be replaced by other deep learning models according to different visual recognition tasks or application scenarios. Even using 10-m-resolution satellite imagery, TKDS-PtNet was able to obtain superior precision when monitoring building damage, consistently achieving overall F1-score values above 72.5 in the six Syrian cities and 83.5 in the four Ukrainian cities. Performance on this monitoring task is far higher than the CNN-STS[5] and ResNet models. Our study provides reliable evidence of urban destruction during the Syrian civil war and the Russia–Ukraine conflict. Substantial completely damaged buildings have been confirmed in the entire Mariupol City by the Russia–Ukraine conflict, in particular, Livoberezhnyi, Zhovtnevyi and Azovstal areas. Also, our results show that Aleppo was the most seriously damaged city where a large number of landmark cultural and heritage buildings were completely destroyed. The populations living in these damaged areas amounted to approximately 120,800–162,800 (116,800–158,600 in Syria and 4,000–4,200 in Ukraine).

https://doi.org/10.1038/s44284-024-00060-6

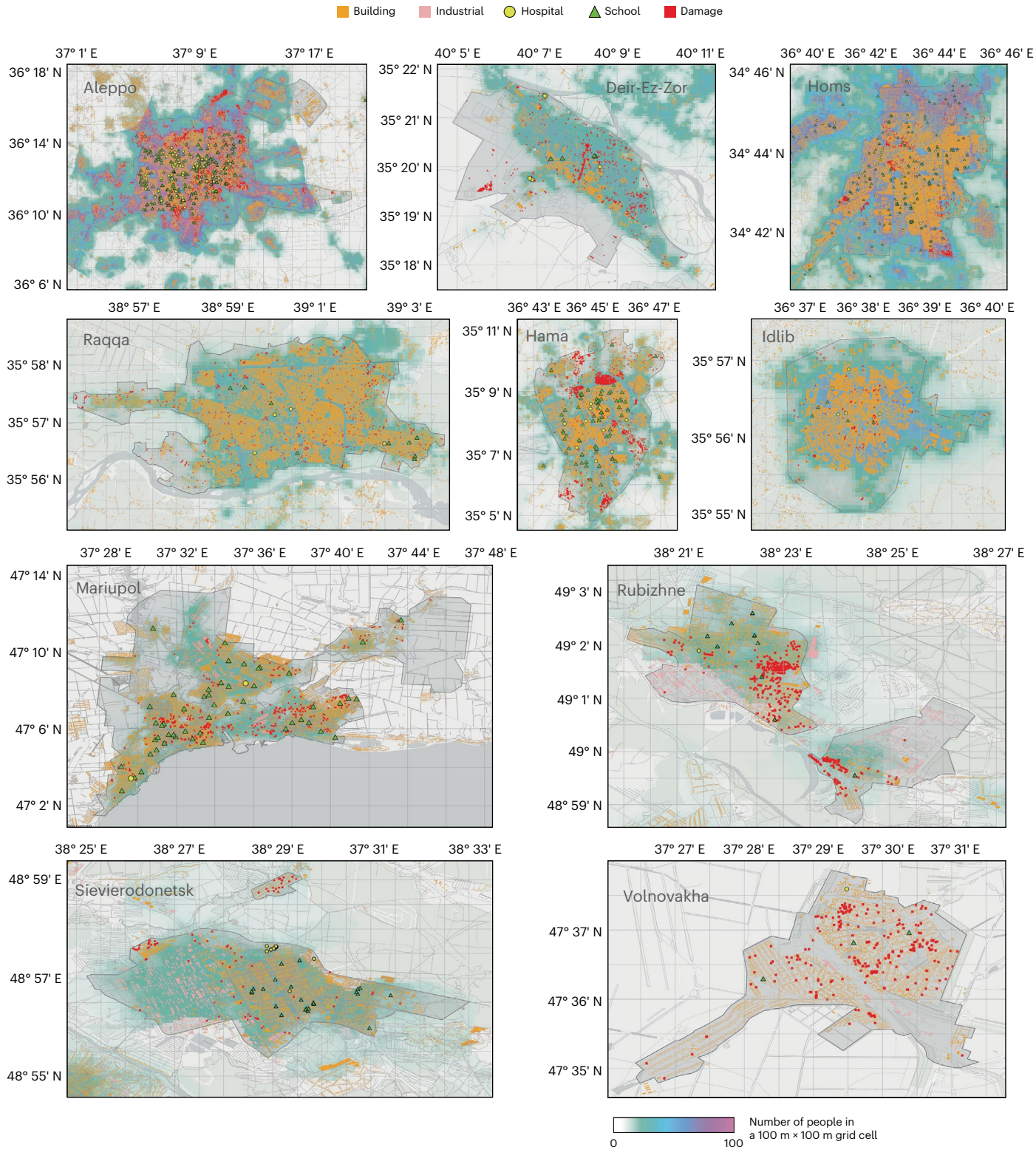

Fig. 4 | Building damage and population distribution in Syrian Aleppo, Homs, Raqqa, Hama, Deir-Ez-Zor and Idlib cities and Ukrainian Mariupol, Rubizhne, Sievierodonetsk and Volnovakha cities. The red rectangles represent the completely damaged building footprints (the sizes of the squares have been adjusted for visual comfort). The locations of hospitals (yellow circles) and schools (green triangles) are obtained from OSM data. The orange and pink polygons represent all the building footprints of the city. In Homs, Raqqa, Hama and Idlib cities where the OSM data are large missing, GlobalMLBuildingFootprints data are used as alternative data for building footprints. The background map color indicates the average population density in each 100 m × 100 m grid cell during the war.

This work illustrates that our model allows policymakers to gain a good overview of the level of destruction in a city, both temporally and spatially. It also provides estimates of the affected infrastructure and population. However, to get precise numbers on critical infrastructure that is damaged, policymakers need to manually verify the model output since a high level of certainty is of significance[33]. In these

cases, one major contribution of the model is that verification of critical infrastructure damage may be done more efficiently through the use of TKDS. The generated urban infrastructure destruction data also opens up possibilities for a set of new research agendas in urban sciences[5,34–36]. For instance, our findings may help understand the microlevel forcing drivers of urban conflict. What measures need to be taken to reduce civilian fatalities during urban conflict? What are the effects of building destruction on displacement, health burden and educational attainment?

This study has several limitations. We focus on monitoring completely damaged buildings and exclude moderate and severe damage. This is because the destruction appearance of the latter might not be visible in satellite images[5]. Identification of completely damaged buildings can be solved by using multi-angle satellite imagery. Although the TKDS-PtNet model has achieved high detection accuracy with imbalanced samples, the precision performance of the model is still affected by heavily imbalanced samples. The model may suffer from model biases caused by biased training data[33]. Thus, it is essential to acknowledge that the precise war impact assessments still require manual verifications of the model output. None of the limitations undermine the validity of our analysis, but they do highlight important avenues for future work. The model accuracy can be improved with extra ground-truth surveys and/or higher-resolution satellite imagery. As higher-resolution satellite imagery becomes more available and artificial-intelligence-powered tools for object detection advance, the captured building destruction samples might be more diverse and prevalent. These will further enhance the applicability of this work. In future work, we will collect the damaged critical infrastructure samples such as roads and railways, and identify the destruction using our method. Based on these data, we will develop a fit-for-purpose resilience framework for post-conflict peace building of critical infrastructure.

Methods

Datasets

Satellite imagery. To monitor completely damaged buildings caused by the Syrian civil war, the training and validation data sets in the six cities are generated using 81 satellite images (approximately 0.5 m resolution) with red–green–blue and three-band channels obtained from Google Earth[37] during 2009–2018. The image collection in each city contains at least one pre-war image patch.

Sentinel-2 Multi-Spectral Instrument Bottom of Atmosphere reflectance images (Level-2A) are selected for monitoring Mariupol, Sievierodonetsk, Volnovakha and Rubizhne cities of Ukraine. We obtain all Sentinel-2 images with cloud coverage less than 90% in Mariupol and Volnovakha cities between 24 February 2022 and 1 July 2022, in Rubizhne and Sievierodonetsk cities between 24 February 2022 and 1 October 2022, and those in the corresponding areas from 10 September 2021 are taken as pre-war images. Since Sentinel-2 has a high revisit frequency of 5 days, all the collected Sentinel-2 images with cloud removal after the war still can completely cover the studied areas.

Labeled samples. The UNITAR/UNOSAT[27] use 0.5-m-resolution satellite images from Google Earth and other data sources such as OpenStreetMap (OSM) data to visually identify and label multiple instances of building damage in the six Syrian cities and four Ukrainian cities. The completely damaged building samples in Syria mainly consist of residential areas in conflict zones. For Mariupol City, building damage samples are only available in three areas, that is, Livoberezhnyi, Zhovtnevyi and Azovstal. We focus on detecting the completely damaged buildings, since the damaged patterns of other levels are not always clearly visible in remote sensing images[5]. The positive or negative samples are defined according to the UNITAR/UNOSAT annotation. We divide the study area into a series of nonoverlapping grids $g_1, g_2, \cdots g_i \cdots, g_k$ with a size of 60 m × 60 m, where k denotes the number of the grids. For any grid g_i, it includes a pre-war image patch $p_{t_0}^{g_i}$ and n ongoing war image patches $p_{t_1}^{g_i}, p_{t_2}^{g_i}, \cdots, p_{t_n}^{g_i} (n > 1$, and n may vary across grids). Each of the ongoing war images $p_{t_j}^{g_i}$ in g_i is stacked with $p_{t_0}^{g_i}$ to construct n bitemporal-image patches $x_{t_0t_1}^{g_i}, x_{t_0t_2}^{g_i}, \cdots, x_{t_0t_j}^{g_i}, \cdots, x_{t_0t_n}^{g_i}$.

In the building damage monitoring of the Syrian civil war, each 0.5-m-resolution satellite image is cropped into nonoverlapping patches with a size of 120 × 120 pixels. Then each image patch is downsampled to 6 × 6 pixels as the image patch of 10-m- resolution images. The 10-m-resolution and 20-m-resolution bands of Sentinel-2 satellite images are used to monitor the Russia–Ukraine conflict urban destruction. They are cropped into image patches with different sizes (6 × 6 pixels for bands with a resolution of 10 m and 3 × 3 pixels for bands with a resolution of 20 m) as the sample data.

The bitemporal-image patch $x_{t_0t_j}^{g_i}$ is classified as a damaged building sample (positive sample) if g_i contains at least one complete destruction building in time t_j; otherwise, it is an undamaged building sample (negative sample). The dataset is expanded following the strategy similar to the study[5] (Supplementary Fig. 10). The bitemporal-image patches without labels in a grid are assigned to the classes according to the three principles. (1) For a grid without damaged buildings in time t_l, the bitemporal-image patches on the grid are also assigned to the undamaged category before t_l. (2) There are almost no urban reconstructions during or in a short period after the war. If the image patches are positive in t_l, all the patches in the same grid are also labeled as positive after t_l. (3) For the grids labeled as damaged or undamaged at both t_{l1} and t_{l2} ($t_{l1} < t_{l2}$), their bitemporal-image patches are correspondingly labeled damaged or undamaged at the two timestamps.

Multi-temporal patch sequence samples. The multi-temporal patch sequence samples are generated by labeling the completely damaged buildings with the same grid in multiple periods. Specifically, we sort annotated bitemporal-image patches by time within each grid $g_i (i = 1, 2, \cdots, k)$ to form the multi-temporal patch sequence samples PS^{g_i}. Each multi-temporal patch sequence sample is a four-dimensional patch matrix with a temporal dimension to store all annotated bitemporal-image patches $x_{t_0t_1}^{g_i}, x_{t_0t_2}^{g_i}, \cdot, x_{t_0t_n}^{g_i}$ in a specific grid g_i and a list that stores the corresponding labels $l_{t_1}^{g_i}, l_{t_2}^{g_i}, \cdots, l_{t_n}^{g_i}$. The length of the list is equal to the temporal dimension of the matrix.

TKDS-PtNet

As shown in Fig. 2, PtNet sequentially extracts the building damage features of each bitemporal-image patch from the image patch sequences, and a TCD uses the temporal context information of the image feature sequences to output a list of the predicted results. The workflow of the TKDS-PtNet model is similar to that of a natural language translation task, in which each bitemporal-image patch corresponds to a semantic vector in a sentence. Each of the bitemporal-image patches in the multi-temporal patch sequence is encoded into a semantic vector containing multi-temporal features through PtNet. The size and labeling time of class samples vary with regions, resulting in the multi-temporal patch sequences often having different lengths. Zero matrices are padded to make multi-temporal patch sequence samples have the same size as they are input into TKDS-PtNet. The predicted values in the corresponding positions of the zero matrices are excluded from the final results. The TCD outputs the detection results in chronological order by correlating the contextual relationships of these semantic vectors in the temporal dimension.

PtNet. PtNet is a classification/regression deep learning model (Fig. 2), consisting of a pixel embedding and a transformer-based classifier. The pixel embedding is composed of two ResNet blocks, an average pooling layer and a channel attention layer. It encodes the input bitemporal-image patch $x_{t_0t_j}^{g_i}$ into 36 semantic tokens T^1. The transformer-based classifier, consisting of two transformer blocks and a fully connected layer, uses a multi-head self-attention mechanism to extract the changing features of the damaged buildings before and

during/after the war from the semantic tokens. The length of 36 semantic tokens is compressed to 4 (T^2) by the fully connected layer. These semantic tokens are concatenated into a semantic vector $V_{t_0t_j}^{g_i}$ with a size of 144 and a multi-layer perceptron compressed $V_{t_0t_j}^{g_i}$ into a single classification result $C_{t_j}^{g_i}$. Each bitemporal-image patch $x_{t_0t_j}^{g_i}$ in a certain multi-temporal patch sequence $\mathrm{PS}^{g_i} = \{x_{t_0t_1}^{g_i}, x_{t_0t_2}^{g_i}, \cdots, x_{t_0t_n}^{g_i}\}$ is encoded into a vector $V_{t_0t_j}^{g_i}$, and TCD compresses the vector sequence $\mathrm{VS}^{g_i} = \{V_{t_0t_1}^{g_i}, V_{t_0t_2}^{g_i}, \cdots, V_{t_0t_n}^{g_i}\}$ into a classification probability vector $\mathrm{CS}'^{g_i} = \{C_{t_1}'^{g_i}, C_{t_2}'^{g_i}, \cdots, C_{t_n}'^{g_i}\}$.

We do not adopt pooling layers in PtNet, and set the kernel size of all convolution layers in the ResNet block to 1, to prevent mixing information from adjacent pixels during the encoding process. In this way, the detection results are only determined by the input image patches, and each semantic vector is encoded from a single pixel.

When Sentinel-2 satellite images are used as the input of the model, the multi-scale feature fusion head is constructed (Supplementary Fig. 11a) before pixel embedding, to fuse bands with 10 m and 20 m resolutions. The upscaling block (Supplementary Fig. 11b) in the multi-scale feature fusion head upscales 20-m-resolution bands to match the size of the 10-m-resolution ones. We stack them as the input of the pixel embedding. The upscaling block consists of a 2 × 2 deconvolution layer and a 1 × 1 convolution layer, and the number of feature channels remains unchanged during the upscaling process.

Temporal feature learning. In TKDS-PtNet, $n (n > 0)$ temporal bitemporal-image patches $x_{t_0t_1}^{g_i}, x_{t_0t_2}^{g_i}, \cdots, x_{t_0t_n}^{g_i}$ in a multi-temporal patch sequence PS^{g_i} represent n semantic vector features $V_{t_0t_1}^{g_i}, V_{t_0t_2}^{g_i}, \cdots, V_{t_0t_n}^{g_i}$ via PtNet. These n vectors are sequentially ordered in the temporal dimension and mapped to a sequence of the prediction results $\mathrm{CS}'^{g_i} = \{C_{t_1}'^{g_i}, C_{t_2}'^{g_i}, \cdots, C_{t_n}'^{g_i}\}$ through the TCD structure. The TCD structure consists of two one-dimensional convolutional layers with kernel sizes of 3, a padding of 1 and a rectified linear unit (ReLu) activation layer. One-dimensional convolution is a common method for correlating neighboring semantic vectors in natural language processing, which endows TCD with the ability to capture contextual features in the temporal dimension. This indicates that TKDS-PtNet can learn the features of the neighboring temporal images before and after the buildings are destroyed.

The state of building damage exhibits irreversible patterns in the temporal dimension, which is manifested in temporal monitoring as a constraint in which the number of transitions NT in the building damage state changes is no more than 1. As NT is equal to 1, the temporal sequence begins with an undamaged state. These patterns are encapsulated within the label vector of the temporal samples. The TCD in TKDS combines the temporal context of the features to capture local patterns in the temporal dimension. TKDS also incorporates TTV regularization, which enforces the temporal global constraints on the detection results. In this way, our temporal prediction is aligned more closely with the patterns of building damage in real-world scenarios.

We combine the cross-entropy function with the TTV regularization term (equation (1)) to form the loss function (equation (2)).

$$\mathrm{TTV}\left(\mathrm{CS}'\right) = \varphi \frac{\sum_{i=1}^{\mathrm{bs}} K^{g_i}\,\mathrm{NT}^{g_i}}{\sum_{i=1}^{\mathrm{bs}} n^{g_i}}$$

$$\mathrm{NT}^{g_i} = \sum_{j=1}^{n^{g_i}} \left| C_{t_{j+1}}'^{g_i} - C_{t_j}'^{g_i} \right| \tag{1}$$

$$K^{g_i} = \begin{cases} 0, & \mathrm{NT}^{g_i} = 1 \text{ and } C_{t_1}'^{g_i} = 0 \\ 1, & \text{else} \end{cases}$$

$$\mathrm{TotalLoss} = \mathrm{CELoss}(C', l) + \mathrm{TTV}\left(\mathrm{CS}'\right) \tag{2}$$

where $\mathrm{CELoss}(C', l)$ is a cross-entropy loss to compute the mapping error loss between the damage predicted labels C' and ground truth l for each image patch. $\mathrm{TTV}(\mathrm{CS}')$ is a TTV regularization term to measure the deviation between a predicted binary sequence CS' and its damage pattern. NT^{g_i} is the sum of the absolute values of the first-order difference of the predicted sequence CS'^{g_i}, and represents the number of the transitions between the undamaged and damaged states in a predicted sequence. K^{g_i} is a piecewise constant. $\mathrm{TTV}(\mathrm{CS}')$ is the sum of NT^{g_i} divided by the length of multi-temporal patch sequence. φ is the weight coefficient, bs is the batch size and n^{g_i} is the length of CS'^{g_i}.

Elimination of pseudo-temporal regularity. In some intense conflict zones, the number of damaged buildings suddenly increases at a given time. Such patterns are usually unique in certain cities or regions. While knowing these patterns helps identify the building damage in specific areas, it may adversely affect the accuracy of the detection in other conflict zones. To prevent the model from learning the patterns inherent in the temporal sequence samples, the random copy and delete (RCD) strategy has been developed as a unique data augmentation method for expanding temporal sequence samples. RCD can randomly select, copy, insert and delete the patches within the temporal sequence samples.

Specifically, for the multi-temporal patch sequence samples, their regularity in the temporal dimension (for example, adjacent relationships among images, and positions of the images in the sequence) is usually determined by the acquisition and annotation time. As shown in Supplementary Fig. 8, to eliminate the pseudo-regularity, the RCD is introduced during the training procedure. Without changing the temporal sequence, some samples in the multi-temporal patch sequence are copied or removed to increase the diversity of the samples in the temporal dimension. The number and positions of the copied and deleted samples are randomly determined. We restrict the number of the copied samples N_{copy} and delete the samples N_{del} to prevent the length L of the samples in the multi-temporal patch sequence from being too long ($L > L_{\max}$, where $L_{\max}$ is the maximum length of the original multi-temporal patch sequence) or too short ($L > 5$). $L_{\mathrm{copy}} \le L$ and $L_{\mathrm{copy}} \le L_{\max} - L$ and $L + L_{\mathrm{copy}} - L_{\mathrm{del}} \ge 5$. The goal is to preserve the intrinsic temporal patterns of the sample sequences while increasing the diversity of the distribution of the damaged samples within them. This strategy enhances the detection performance and the generalizability of the proposed model.

SSDA for building damage detection

Incorporating SCL[26] and MMD[25], we develop an SSDA strategy based on the TKDS-PtNet model in the context of few-shot building damage detection (Supplementary Fig. 12). The aim is to improve the detection performance in the target city with few image samples. Labeled source domain samples ($\mathrm{PS_S}$), unlabeled target domain samples ($\mathrm{PS_{TU}}$) and labeled target domain samples ($\mathrm{PS_{TL}}$) are fed into TKDS-PtNet, where $\mathrm{PS_{TL}}$ is much smaller than $\mathrm{PS_S}$ or $\mathrm{PS_{TU}}$. TKDS-PtNet only predicts the results of $\mathrm{PS_S}$ and $\mathrm{PS_{TL}}$, and computes the TotalLoss (equation (2)) together with the corresponding label sequences ($\mathrm{LS_S}$ and $\mathrm{LS_{TL}}$).

The input features of the last layer in TCD are extracted and divided into patch features f, which match with each bitemporal-image patch x in the input PS. The MMD[25] loss is calculated according to the source domain features f_{S} and the target domain features $\{f_{\mathrm{TL}}, f_{\mathrm{TU}}\}$ (f_{TL} and f_{TU} respectively represent the features of labeled samples and unlabeled samples). In the feature space, a consistency constraint is introduced to bring the feature distributions of the two domain features closer.

We also separate f_{TL} and f_{S} to features of damaged samples $\{f_{\mathrm{TL_1}}, f_{\mathrm{S_1}}\}$ and features of damaged samples $\{f_{\mathrm{TL_0}}, f_{\mathrm{S_0}}\}$, and compute feature prototypes for the four categories ($Pf_{\mathrm{TL_1}}, Pf_{\mathrm{TL_0}}, Pf_{\mathrm{S_1}}, Pf_{\mathrm{S_0}}$). The feature prototypes are normalized to the unit hypersphere, and the distances between each prototype and the other three are computed as a measure of similarity according to the distance between vectors on the hypersphere. Based on whether they come from the images with the same

class, an SCL loss[26] is calculated. This enables the model to map patches from two domains with the same class into a feature space and extract discriminative features.

Validations

We use the georeferenced bombing events obtained from LiveUAmap (http://liveuamap.com) to externally validate the predicted results. We captured a total of 1,042 bombing events recorded by LiveUAmap in the six Syrian cities from 2011 to 2018 and the four Ukrainian cities from 24 February to 1 October 2022. The prediction results are obtained by TKDS-PtNet using the LiveUAmap records which include over 3.3 million predictions and all missing values are replaced with zero. A two-way fixed effects model is used to estimate the association of the bombing events with the building damage prediction results (equation (3)).

$$D_{gt} = \mathrm{Bomb}_{gt} + \theta_g + \mu_t \nu_c + \epsilon_{gt} \quad (3)$$

where D_{gt} is the binary prediction value of the building damage, and g and t represent the grid and capture time of the input image patches, respectively. Bomb_{gt} is a binary indicator that denotes whether there has been a bombing event before time t at grid g. θ_g and $\mu_t \nu_c$ are the grid fixed effects and time–city fixed effects. ϵ_{gt} is the error term.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The original labels of the completely destroyed buildings are publicly available via UNITAR at https://www.unitar.org/maps (ref. 27). The 0.5-resolution satellite images in 2009 and 2018 are available via Google Earth at https://www.google.com/earth (ref. 37). The sentinel-2 satellite images are available from https://dataspace.copernicus.eu/. The bombing events are available from http://liveuamap.com. The WorldPop population dataset is available from https://hub.worldpop.org/. Building footprints in Syria and Ukraine are available from https://www.openstreetmap.org/ and via GitHub at https://github.com/microsoft/GlobalMLBuildingFootprints (ref. 32).

Code availability

The source codes are available on GitHub at https://github.com/Houzy116/TKDS-PtNet (ref. 38).

References

1. Piccoli, G. B., Brunori, G., Gesualdo, L. & Kalantar-Zadeh, K. The impact of the Russian–Ukrainian war for people with chronic diseases. *Nat. Rev. Nephrol.* **18**, 411–412 (2022).
2. Aiken, E., Bellue, S., Karlan, D., Udry, C. & Blumenstock, J. E. Machine learning and phone data can improve targeting of humanitarian aid. *Nature* **603**, 864–870 (2022).
3. Smythe, I. S. & Blumenstock, J. E. Geographic microtargeting of social assistance with high-resolution poverty maps. *Proc. Natl Acad. Sci. USA* **119**, 2120025119 (2022).
4. Wagner, Z. et al. Armed conflict and child mortality in Africa: a geospatial analysis. *Lancet* **392**, 857–865 (2018).
5. Mueller, H., Groeger, A., Hersh, J., Matranga, A. & Serrat, J. Monitoring war destruction from space using machine learning. *Proc. Natl Acad. Sci. USA* **118**, 2025400118 (2021).
6. Dando, B. D. E. et al. Identifying attacks in the Russia–Ukraine conflict using seismic array data. *Nature* **621**, 767–772 (2023).
7. Lubin, A. & Saleem, A. Remote sensing-based mapping of the destruction to Aleppo during the Syrian civil war between 2011 and 2017. *Appl. Geogr.* **108**, 30–38 (2019).
8. Kahraman, F., Imamoglu, M. & Ates, H. F. Battle damage assessment based on self-similarity and contextual modeling of buildings in dense urban areas. In *Proc. 2016 IEEE International Geoscience and Remote Sensing Symposium* 5161–5164 (IEEE, 2016).
9. Nabiee, S., Harding, M., Hersh, J. & Bagherzadeh, N. Hybrid u-net: semantic segmentation of high-resolution satellite images to detect war destruction. *Mach. Learn. Appl.* **9**, 100381 (2022).
10. Nex, F., Duarte, D., Tonolo, F. G. & Kerle, N. Structural building damage detection with deep learning: assessment of a state-of-the-art CNN in operational conditions. *Remote Sensing* **11**, 2765 (2019).
11. Tilon, S., Nex, F., Kerle, N. & Vosselman, G. Post-disaster building damage detection from Earth observation imagery using unsupervised and transferable anomaly detecting generative adversarial networks. *Remote Sensing* **12**, 4193 (2020).
12. Wu, C. et al. Building damage detection using u-net with attention mechanism from pre- and post-disaster remote sensing datasets. *Remote Sensing* **13**, 905 (2021).
13. Kalantar, B., Ueda, N., Al-Najjar, H. A. H. & Halin, A. A. Assessment of convolutional neural network architectures for earthquake-induced building damage detection based on pre- and post-event orthophoto images. *Remote Sensing* **12**, 3529 (2020).
14. Janalipour, M. & Mohammadzadeh, A. Building damage detection using object-based image analysis and anfis from high-resolution image (case study: Bam earthquake, Iran). *IEEE J. Select. Top. Appl. Earth Obs. Remote Sensing* **9**, 1937–1945 (2016).
15. Gueguen, L. & Hamid, R. Large-scale damage detection using satellite imagery. In *Proc. 2015 IEEE Conference on Computer Vision and Pattern Recognition* 1321–1328 (IEEE, 2015).
16. Fujita, A. et al. Damage detection from aerial images via convolutional neural networks. In *Proc. 2017 15th IAPR International Conference on Machine Vision Applications* 5–8 (IAPR, 2017).
17. Dong, L. & Shan, J. A comprehensive review of earthquake-induced building damage detection with remote sensing techniques. *ISPRS J. Photogramm. Remote Sensing* **84**, 85–99 (2013).
18. Abdi, G. & Jabari, S. A multi-feature fusion using deep transfer learning for earthquake building damage detection. *Can. J. Remote Sensing* **47**, 337–352 (2021).
19. Stramondo, S., Bignami, C., Chini, M., Pierdicca, N. & Tertulliani, A. Satellite radar and optical remote sensing for earthquake damage detection: results from different case studies. *Int. J. Remote Sensing* **27**, 4433–4447 (2006).
20. Oksuz, K., Cam, B. C., Kalkan, S. & Akbas, E. Imbalance problems in object detection: a review. *IEEE Trans. Pattern Anal. Mach. Intell.* **43**, 3388–3415 (2021).
21. Aung, T. S. Satellite analysis of the environmental impacts of armed-conflict in Rakhine, Myanmar. *Sci. Total Environ.* **781**, 146758 (2021).
22. Wind, B. & Ibrahim, B. The war-time urban development of Damascus: how the geography- and political economy of warfare affects housing patterns. *Habitat Int.* **96**, 102109 (2020).
23. Yuksel, A. S. Cities at war: global insecurity and urban resistance. *Ethnic Racial Stud.* **44**, 527–529 (2021).
24. Weinthal, E. & Sowers, J. Targeting infrastructure and livelihoods in the West Bank and Gaza. *Int. Affairs* **95**, 319–340 (2019).
25. Schölkopf, B., Platt, J. & Hofmann, T. A Kernel Method for the Two-Sample-Problem. *Advances in Neural Information Processing Systems 19*, 513–520 (MIT Press, 2007).
26. Khosla, P. et al. Supervised contrastive learning. *Adv. Neural Inform. Process. Syst.* **33**, 18661–18673 (2020).

27. United Nations Institute for Training and Research Operational Satellite Applications Programme. *UNITAR* https://www.unitar.org/maps/ (accessed 25 April 2023).
28. Liveuamap. *Liveuamap* https://www.liveuamap.com/ (accessed 25 April 2023).
29. Selvaraju, R.R. et al. Grad-cam: visual explanations from deep networks via gradient-based localization. In *Proc. 2017 IEEE International Conference on Computer Vision (ICCV)* 618–626 (IEEE, 2017).
30. University of Southampton WorldPop. *WorldPop* https://hub.worldpop.org/ (accessed 20 September 2023).
31. OpenStreetMap Contributors OpenStreetMap. *OpenStreetMap* https://www.openstreetmap.org/ (accessed 20 September 2023).
32. Bing Maps GlobalMLBuildingFootprints. *GitHub* https://github.com/microsoft/GlobalMLBuildingFootprints (accessed 20 September 2023).
33. Sticher, V., Wegner, J. D. & Pfeifle, B. Toward the remote monitoring of armed conflicts. *Proc. Natl Acad. Sci. Nexus* **2**, 181 (2023).
34. Besley, T. & Mueller, H. Estimating the peace dividend: the impact of violence on house prices in Northern Ireland. *Am. Econ. Rev.* **102**, 810–833 (2012).
35. Burke, M., Hsiang, S. M. & Miguel, E. Climate and conflict. *Annu. Rev. Econom.* **7**, 577–617 (2015).
36. Manacorda, M. & Tesei, A. Liberation technology: mobile phones and political mobilization in Africa. *Econometrica* **88**, 533–567 (2020).
37. Google Earth. *Google Earth* https://www.google.com/earth (accessed 25 April 2023).
38. TKDS-PtNet. *GitHub* https://github.com/Houzy116/TKDS-PtNet (accessed 25 March 2024).

Acknowledgements

This work was supported by the National Natural Science Foundation of China under grant numbers 41925006 (L.Z.), 42293272 (L.Z.) and 42201368 (Y.Q.).

Author contributions

Z.H., Y.Q., L.Z. and C.Z. conceptually designed the study; Z.H., Y.Q., Q.Y. and X.Y. performed the research; Z.H., Y.Q., L.Z., J.L., F.W., A.Z., Z.C., H.T., Y.W., J.R. and S.L. drafted the manuscript; Y.Z., Z.H., X.L., Y.L., S.P. and X.M. analyzed the data.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44284-024-00060-6.

Correspondence and requests for materials should be addressed to Liqiang Zhang or Chenghu Zhou.

Peer review information *Nature Cities* thanks Ali Darvishi Boloorani, Stergios-Aristoteles Mitoulis and Valerie Sticher for their contribution to the peer review of this work.

nature portfolio

Corresponding author(s): Liqiang Zhang, Chenghu Zhou

Last updated by author(s): Mar 13, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☐	A description of all covariates tested
☐	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Stata V18
Data analysis	Stata V18, pytorch 1.12.0, numpy 1.19.1, sklearn 1.1.3, geopandas 0.8.1, GDAL 3.0.2, Resnet, swin-transformer, STS

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The original labels of the completely destroyed buildings are publicly available at https://www.unitar.org/maps.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	This study does not involve this issue.
Reporting on race, ethnicity, or other socially relevant groupings	This study does not involve this issue.
Population characteristics	This study does not involve this issue.
Recruitment	This study does not involve this issue.
Ethics oversight	This stduy does not involve this issue.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	
Data exclusions	
Replication	
Randomization	
Blinding	

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We propose a new Temporal-Knowledge-guided Detection Scheme (TKDS), which can monitor urban destructions with either highor medium-resolution satellite imagery. A pre-war image and each of the images available during the war are stacked to form a group of bitemporal image patches. Then, the bitemporal-image patches at a given location are sorted by date to construct a multitemporal patch sequence as input to TKDS. To extract the semantic vector representation of each bitemporal-image patch, a pixelbased transformer network (PtNet) is developed. A temporal total variation (TTV) regularization function incorporating buildingdamaged temporal patterns is utilized as a temporal global constraint to guide PtNet to learn the damage patterns. TKDS further uses a temporal convolution decoder to obtain the contextual relationship of the bitemporal-image patches in the temporal dimension based on the semantic vectors extracted by PtNet, and generates the detection results.
Research sample	We use the samples of damaged buildings in ten cities provided by the United Nations Operational Satellite Applications Program (UNOSAT), which include the representative severe destruction areas in Syria and Ukraine.
Sampling strategy	In this study, satellite images of six cities in Syria and four cities in Ukraine annotated by the UNOSAT were used to create samples. A total of 1,049,033 patch samples from Syria and 287,376 patch samples from Ukraine were transformed into 79,895 and 40,970 multi-temporal samples, respectively. 70% of all the multi-temporal samples were randomly selected for training, while the remaining 30% were used for testing.
Data collection	United Nations Operational Satellite Applications Program (UNOSAT) of the United Nations Institute for Training and Research used 0.5m-resolution satellite images from Google Earth and other data sources such as OSM data to visually identify and label multiple instances of building damage in the six Syrian cities and four Ukrainian cities. The testing samples are not used to train the model. Thus, the evaluation is reliable.
Timing	The training and validation data sets in the six cities are generated using 81 satellite images with red-green-blue and three-band

Timing	channels obtained from Google Earth during 2009-2018. We obtained all Sentinel-2 images with cloud coverage less than 90% in Mariupol and Volnovakha cities between February 24, 2022, and July 1, 2022, in Rubizhne and Sievierodonetsk cities between February 24, 2022, and October 1, 2022, and those in the corresponding areas from September 10, 2021 are taken as pre-war images.
Data exclusions	We only use annotations of completely destroyed buildings provided by the United Nations to create positive samples. As a result, our detection results do not account for buildings with other levels of damage.
Non-participation	No participants were involved in the study.
Randomization	We adopt a two-way fixed effects model and external bombing events to validate the reliability of the destruction detection results. This model includes two sets of fixed effects: The grid fixed effects are utilized to control for the fixed influences experienced by buildings in different locations over the long term, such as housing structures, geological conditions, etc. The time -city fixed effects are used to control for the specific shocks experienced by each city at different times, such as earthquakes, disasters, or policy impacts such as building renovations.

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	
Research sample	
Sampling strategy	
Data collection	
Timing and spatial scale	
Data exclusions	
Reproducibility	
Randomization	
Blinding	

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	
Location	
Access & import/export	
Disturbance	

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Validation

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)

Authentication

Mycoplasma contamination

Commonly misidentified lines
(See ICLAC register)

Palaeontology and Archaeology

Specimen provenance

Specimen deposition

Dating methods

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals

Wild animals

Reporting on sex

Field-collected samples

Ethics oversight

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	
Study protocol	
Data collection	
Outcomes	

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☐	☐	Public health
☐	☐	National security
☐	☐	Crops and/or livestock
☐	☐	Ecosystems
☐	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☐	☐	Demonstrate how to render a vaccine ineffective
☐	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☐	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☐	☐	Increase transmissibility of a pathogen
☐	☐	Alter the host range of a pathogen
☐	☐	Enable evasion of diagnostic/detection modalities
☐	☐	Enable the weaponization of a biological agent or toxin
☐	☐	Any other potentially harmful combination of experiments and agents

Plants

Seed stocks	This study does not involve this issue.
Novel plant genotypes	This study does not involve this issue.
Authentication	This study does not involve this issue.

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	
Files in database submission	
Genome browser session (e.g. UCSC)	

Methodology

Replicates	
Sequencing depth	
Antibodies	
Peak calling parameters	
Data quality	
Software	

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	
Instrument	
Software	
Cell population abundance	
Gating strategy	

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	
Design specifications	
Behavioral performance measures	

Acquisition

Imaging type(s)

Field strength

Sequence & imaging parameters

Area of acquisition

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software

Normalization

Normalization template

Noise and artifact removal

Volume censoring

Statistical modeling & inference

Model type and settings

Effect(s) tested

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference

(See Eklund et al. 2016)

Correction

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Graph analysis

Multivariate modeling and predictive analysis