Satellite-based Conflict Damage Detection: Siamese CNNs for Forced Displacement Planning in Ukraine

Taulant Matarova[1] and Diana Suleimenova[1][0000−0003−4474−0943]

Department of Computer Science, Brunel University of London, London, United Kingdom
taulantsyle@outlook.com, diana.suleimenova@brunel.ac.uk

Abstract. Forced displacement due to armed conflict is an escalating global challenge, with over 117 million people displaced worldwide in 2025. Traditional damage assessment methods are slow, resource-intensive, and often impractical in active conflict zones. We present a Siamese Convolutional Neural Network for binary building damage detection using freely available Sentinel-2 medium-resolution satellite imagery. Using pre- and post-conflict image pairs from the Copernicus Data Space Ecosystem and UNOSAT building damage annotations, we evaluate eight model variants across four ResNet backbone depths with a Multilayer Perceptron classification head, tested across twenty-two conflict-affected Ukrainian regions. Our best-performing model, a Siamese ResNet-101 with an MLP head, achieves an Area Under the ROC Curve of 0.911 and an Average Precision of 0.888, demonstrating strong detection capability without high-resolution commercial imagery. We analyse the model's utility for both outward and inward humanitarian displacement planning and discuss its potential integration with agent-based models for simulating forced displacement scenarios. Our results demonstrate that resource-efficient, publicly available satellite data can support data-driven humanitarian planning at scale.

Keywords: Siamese CNN · Damage Detection · Satellite Imagery · Population Displacement.

1 Introduction

Forced displacement is one of the defining humanitarian challenges of the twenty-first century. By June 2025, more than 117 million people were forcibly displaced worldwide due to armed conflict, violence, climate-related disasters, and socio-political instability [1]. The escalation of the war in Ukraine following the full-scale invasion of February 2022 illustrates the speed and scale at which contemporary conflicts generate displacement, uprooting millions of internally displaced persons and refugees, and placing severe pressure on humanitarian systems [2]. In such contexts, rapid, data-driven decision-making is essential for allocating resources, planning evacuation and relocation pathways, and preparing host regions.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_5

Effective response requires two complementary forms of operational planning: *outward displacement planning*, which addresses immediate population flight, and *inward displacement planning*, which supports safe return and resettlement [3,4]. Both depend critically on timely and accurate spatial information about building damage, as the distribution and severity of infrastructure destruction shape habitability, infrastructure functionality, and displacement patterns [5].

Traditional damage assessments rely on ground surveys and expert interpretation of imagery. Although reliable, these methods are slow, resource-intensive, and often infeasible in active conflict zones. Large-scale crises, such as the 2010 Haiti earthquake, required weeks of coordinated mapping efforts, highlighting the persistent challenge of timely assessment [6,7]. Remote sensing offers a scalable alternative by enabling observation without direct human access.

Very high-resolution (VHR) commercial imagery enables detailed building-level damage classification [8], but high costs and access constraints limit its operational utility for many humanitarian actors. Medium-resolution imagery from the European Space Agency's Sentinel-2 constellation provides a compelling alternative. It delivers multispectral imagery at 10 m ground sampling distance with a five-day revisit cycle and free, open access through the Copernicus Data Space Ecosystem [9,10]. Despite its coarser resolution, its temporal frequency, spectral richness, and zero-cost availability make it attractive for large-scale humanitarian monitoring.

Siamese Convolutional Neural Networks (CNNs) have emerged as a powerful approach for change detection in remote sensing, learning discriminative representations from paired pre- and post-event imagery through shared-weight architectures [11,12,13]. While prior studies have focused predominantly on VHR data [14,15], the application of Siamese CNNs to medium-resolution imagery remains underexplored. Moreover, systematic comparisons of backbone architectures and classification heads are limited, and integration with operational humanitarian workflows has not been comprehensively addressed.

To bridge these gaps, this paper presents a systematic evaluation of Siamese CNN architectures for binary building damage detection using freely available Sentinel-2 imagery, with explicit relevance to displacement planning. We (1) construct curated datasets from conflict-affected regions using Copernicus Emergency Management Service activations and United Nations Operational Satellite Applications Programme (UNOSAT) annotations [16], (2) benchmark eight model variants based on four ResNet backbones (ResNet-34, ResNet-50, ResNet-101, ResNet-152) with a Multilayer Perceptron classification head, (3) demonstrate competitive performance without reliance on costly VHR imagery, and (4) outline pathways for integrating automated damage mapping into displacement simulation and humanitarian decision-support systems, including the agent-based Flee simulation toolkit [17].

The remainder of the paper is structured as follows. Section 2 reviews the literature on displacement planning, remote sensing-based damage assessment, and Siamese CNN-based change detection. Section 3 describes the study area,

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_5

data sources, preprocessing pipeline, and the Siamese CNN approach with its model configurations. Section 4 presents the experimental results and discusses the humanitarian implications, limitations, and directions for future work. Finally, Section 6 concludes the paper.

2 Related Work

Forced displacement driven by armed conflict reflects a complex interaction of security conditions, economic pressures, and social networks [18,19]. As humanitarian actors seek scalable, data-driven decision-support tools, computational approaches to displacement modelling have gained prominence [20]. Agent-based models (ABMs) are particularly influential, with approaches such as Flee simulating refugee movement along transportation networks under conflict pressure across multiple crises [17,21], though these require spatially granular inputs often incomplete or delayed in operational contexts. Statistical and machine learning methods have similarly been applied using conflict event data, socioeconomic indicators, and satellite-derived variables [6,20,22], including CNN-based route accessibility assessment from satellite imagery [23] and broader ML enhancements to ABMs [24,25]. Nevertheless, research has focused primarily on outward displacement, with limited attention to return planning. Survey evidence from Ukraine highlights persistent uncertainty over housing damage and local safety [26], underscoring the importance of timely building damage information for both flight and return decisions.

Satellite remote sensing provides a scalable mechanism for such spatial evidence. Conflict damage assessment has relied on expert interpretation of optical and synthetic aperture radar imagery since the early 2000s [16], with deep learning substantially accelerating this process. Benchmarks such as xBD standardised large-scale datasets for building damage classification from VHR imagery, enabling convolutional and transformer-based architectures [8], though dependence on commercial VHR data constrains cost-efficiency and scalability. Medium-resolution imagery remains comparatively underexplored. Existing studies demonstrate the feasibility of combining Sentinel-1 and Sentinel-2 data or applying recurrent networks to multitemporal Sentinel-2 imagery for urban change detection [27,28,29], yet translating medium-resolution signals into actionable building-level damage estimates remains difficult [30].

Siamese neural networks, originally developed for signature verification and metric learning [11,12], offer a weight-sharing approach to change detection in paired imagery, enforcing consistent feature representations across co-registered pre- and post-event inputs. In remote sensing, these models demonstrate strong change-detection capability with further gains from recurrent components and attention mechanisms, primarily in VHR settings [14,15,31], and perform competitively against early-fusion approaches when labelled data are limited [13]. Backbone selection also shapes performance: ResNets enable deeper architectures through skip connections [32], with variants such as ResNet-50, -101, and -152 balancing accuracy and cost [33]. VGG-16 remains competitive in trans-

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_5

fer learning despite higher parameter counts [34], and comparative surveys emphasise backbone choice for downstream performance [35]. However, systematic benchmarking of Siamese CNNs with varied backbones for binary building damage detection using medium-resolution Sentinel-2 imagery in conflict settings remains absent.

Overall, although displacement modelling, satellite-based damage assessment, and deep learning are increasingly converging, an operational gap persists. The lack of systematically evaluated, resource-efficient damage detection approaches tailored to medium-resolution imagery and explicitly aligned with displacement planning workflows. This study addresses that gap by evaluating Siamese CNN-based building damage detection using Sentinel-2 imagery and examining its integration into humanitarian simulation and decision-support systems.

3 Methods

The ongoing conflict in Ukraine, which escalated to full-scale war in February 2022, has caused extensive damage to both urban and rural infrastructure [27]. To monitor and analyse this impact, Sentinel-2A Level-2A imagery was obtained from the Copernicus Data Space Ecosystem [10]. Each tile is approximately 1 GB and provides surface reflectance at 10 m ground sampling distance. Four spectral bands were retained for analysis: B02 (Blue, 490 nm), B03 (Green, 560 nm), B04 (Red, 665 nm), and B08 (Near-Infrared (NIR), 842 nm). The visible RGB bands enable full-colour composites, while the NIR band provides additional phenological information for the deep learning model, as healthy vegetation strongly reflects in NIR whereas damaged or disturbed surfaces do not [36]. Sentinel-1 imagery was deliberately excluded to limit the dataset size, making it more feasible for humanitarian organisations with restricted storage capacity.

Building damage labels were obtained from UNOSAT assessment shapefiles covering the 2022 Ukraine conflict [16], as individual buildings cannot be reliably distinguished at Sentinel-2 resolution through visual inspection alone. Only first-instance assessments with available shapefiles (`.shp`) were included; updated reports were excluded to avoid label inconsistencies. Pre-damage images were selected from any date prior to 24 February 2022, while post-damage images corresponded to the exact date of the report, preventing mislabelling of undamaged buildings as damaged. Images were visually inspected to minimise cloud coverage, though small residual clouds were retained to improve the model's robustness to atmospheric artefacts.

After data quality checks, 22 out of 25 conflict-affected regions were retained for modelling. These regions were selected to capture diverse conflict intensities and geographic contexts, maximising the representativeness of the trained models. Labels describe building-level damage assessments recorded during the 2022 Ukraine conflict. Table 1 summarises the complete dataset specification.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_5

3.1 Preprocessing and Exploratory Data Analysis

All preprocessing used the GDAL library [37] within a Conda environment [38]. Coordinate reference systems of each image pair and its shapefile were verified and reprojected where necessary. Bands were merged into a four-channel TIF file (RGB + NIR) and visually inspected in QGIS [39]. Three regions were removed: Borodyanka (insufficient before/after overlap), Sievierodonetsk (single damage label), and Moshchun (a single large contiguous mask block). Points in Mykolaiv that fell outside the image boundary were removed manually.

Table 1. Dataset and Sentinel-2 configuration details.

Field	Description
Dataset	Sentinel-2 Level-2A images
Resolution	10 m per pixel
Bands used	B02 (Blue, 490 nm), B03 (Green, 560 nm), B04 (Red, 665 nm), B08 (NIR, 842 nm)
Images	25 pre-conflict, 25 post-conflict
Labels	UNOSAT building damage assessment shapefiles
Portals	`https://browser.dataspace.copernicus.eu/`; `https://unosat.org/products`
Regions	Borodyanka*, Bucha, Volnovakha, Trostianets, Sumy, Shchastia, Rubizhne, Okhtyrka, Mykolaiv, Moshchun*, Melitopol, Makariv, Lysychansk, Kremenchuk, Kramatorsk, Kherson, Kharkiv, Avdiivka, Antonivka, Chernihiv, Hostomel, Irpin, Vorzel, Azovstal industrial, Sievierodonetsk*

* Removed during preprocessing (see Section 3.1).

Rasterising labels. Shapefile damage points were buffered by 50 m before rasterisation to reduce misalignment at 10 m pixel resolution. Binary masks assigned each pixel a value of 1 (damage) or 0 (no damage). Bounding boxes were computed from the image geotransform as:

$$x_{\max} = x_{\min} + W \times p_x \tag{1}$$

$$y_{\min} = y_{\max} + H \times p_y \tag{2}$$

where $(x_{\min}, y_{\max})$ is the top-left origin, W and H are image dimensions in pixels, and p_x, p_y are pixel sizes.

Patching and dataset construction. Each of the 22 valid region triplets (before image, after image and mask as shown in Figure 1) was converted into aligned 64×64 pixel patches covering 640×640 m on the ground, capturing local

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_5

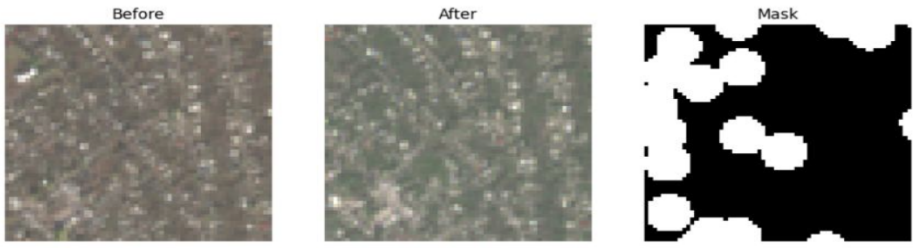

Fig. 1. Representation of the patch triplet from Avdiivka: (left) pre-conflict Sentinel-2A image acquired before 24 February 2022, (centre) post-conflict image acquired on the UNOSAT report date, and (right) corresponding binary damage mask (white indicates damage, black indicates no damage).

road and building context while limiting dilution of damage signals. A patch was labelled *damaged* if any mask pixel equalled 1. Up to 1,000 damaged and 1,000 non-damaged patches per region were stored in compressed `.npz` format (see Data Availability Statement for further details).

Exploratory analysis then identified 2,578 and 2,253 entirely black (invalid) patches in either the before or after channels, respectively, which were removed. The remaining 19,422 non-damaged patches were downsampled to 4,000. The 3,252 damaged patches were increased to 4,000 by augmentation of randomly selected patches via vertical flip, horizontal flip, and 90° rotation, adding 748 patches. 17 duplicate pairs were identified and removed across both classes, and the non-damaged patches were trimmed to match the damaged patches, yielding a balanced final dataset of 3,984 damaged and 3,984 non-damaged patches.

Figure 2 shows that 46.5 % of damaged patches had less than 5 % of their area covered by damage labels, motivating a deep learning approach over simple thresholding. To confirm this, principal component analysis (PCA) reduced the dimensionality of damaged images to 86 components explaining 80 % of variance. K-means clustering ($k = 2$) yielded a silhouette score of 0.1904, indicating minimal linear separability (Fig. 3). The dataset was split 60/20/20 into training, validation, and test subsets, which were used across the eight model variants for fair comparison.

3.2 Siamese CNN Architecture

The Siamese CNN passes before and after image patches independently through a shared ResNet backbone. The first convolutional layer, which expects 3 RGB channels, was extended to a fourth NIR channel by copying the layer and averaging the RGB weights, enabling the model to exploit NIR features from the RGB prior. The ResNet backbone captures multi-scale spatial representations inherently through its hierarchical convolutional stages: shallow layers encode fine-grained edge and texture features at sub-patch spatial scales, while deeper residual stages aggregate semantically richer representations at progressively coarser

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_5

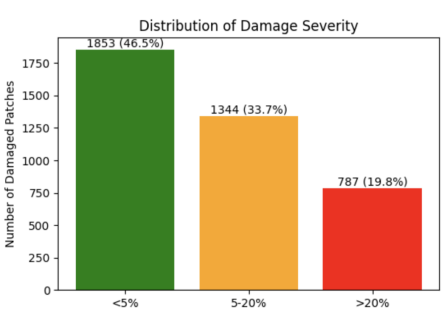

Fig. 2. Distribution of damage severity (percentage of pixels labelled as damaged) across damaged patches.

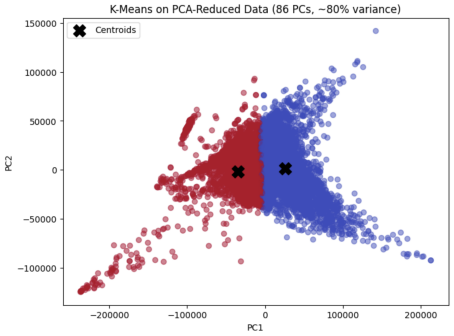

Fig. 3. K-means clustering on PCA-reduced features of damaged patches, showing minimal cluster separation (silhouette score = 0.1904).

scales. The 64×64 pixel patch (640×640 m on the ground) defines the operational inference scale, bridging the 10 m pixel resolution of Sentinel-2 imagery to the neighbourhood-block level at which damage assessments are produced. This patch-to-neighbourhood scale coupling is a deliberate design choice that positions the model's outputs for direct integration with city- and regional-scale displacement simulations. The preprocessing pipeline, implemented in Python using NumPy and Rasterio within a Conda environment, operates across these scales. GDAL reprojection and band merging are performed at the full-tile scale ($\sim 100 \times 100$ km), binary mask rasterisation at the 10 m pixel scale, and patch extraction and `.npz` serialisation at the 64×64 pixel scale. The final fully connected classification layer was removed.

After extracting features f_{before} and f_{after}, a 4D combined vector is constructed:

$$f_{\text{combined}} = \left[f_{\text{before}};\ f_{\text{after}};\ |f_{\text{before}} - f_{\text{after}}|;\ f_{\text{before}} \odot f_{\text{after}} \right] \tag{3}$$

where $\odot$ denotes element-wise product, capturing individual context, magnitude of change, and feature similarity. These combined features are passed to a Multilayer Perceptron (MLP) head with Gaussian Error Linear Unit (GELU) activation [40]:

$$\text{GELU}(x) = 0.5\, x \left(1 + \tanh\left[\sqrt{\tfrac{2}{\pi}} \left(x + 0.044715\, x^3 \right) \right] \right) \tag{4}$$

and `BatchNorm1d` [41] after each linear layer, defined as:

$$y = \frac{x - \text{E}[x]}{\sqrt{\text{Var}[x] + \varepsilon}} \cdot \gamma + \beta \tag{5}$$

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_5

where $\mathrm{E}[x]$ and $\mathrm{Var}[x]$ are the batch mean and variance, ε is a small constant for numerical stability, and γ, β are learnable affine parameters. The output is a single logit with a damage threshold of 0.5 (sigmoid output). Eight architectures were evaluated across ResNet backbone depths (34, 50, 101, 152) and MLP head configurations, as listed in Table 2.

Table 2. Incremental improvement results. Model numbers correspond to training iterations, backbone and MLP head configurations are listed. Dropout rates are shown in parentheses.

Model	ResNet	MLP Head	Accuracy
0	34	$2048\,(0.5) \to 1024\,(0.2) \to 256$	0.8338
1	50	$8192\,(0.5) \to 1024\,(0.2) \to 256$	0.8369
2	50	$8192\,(0.5) \to 4096\,(0.2) \to 1024\,(0.2) \to 256$	0.8331
3	101	$8192\,(0.5) \to 4096\,(0.2) \to 1024\,(0.2) \to 256$	**0.8507**
4	101	$8192\,(0.6) \to 4096\,(0.2) \to 1024\,(0.2) \to 256$	0.8344
5	152	$8192\,(0.6) \to 4096\,(0.2) \to 1024\,(0.2) \to 256$	0.8375
6	152	$8192\,(0.5) \to 4096\,(0.2) \to 1024\,(0.2) \to 256$	0.8319
7	101	$8192\,(0.5) \to 4096\,(0.2) \to 1024\,(0.2) \to 256^*$	0.8394

*Model 7 uses backbone learning rate 10^{-3} (vs. 10^{-4} for all others).

3.3 Training Setup

Data loading, training, and evaluation used PyTorch [42] with a fixed random seed. Sentinel-2 surface reflectance values (stored scaled by 10,000) were divided by 10,000, then z-score normalised per channel using training-set statistics:

$$z = \frac{x - \mu}{\sigma + \epsilon}, \quad \epsilon = 10^{-8} \tag{6}$$

where μ and σ are per-channel mean and standard deviation computed on training data only. Normalisation was applied across all splits via the `PairsDataset` class, which supplies batched (before, after, label) triplets to PyTorch `DataLoader` objects.

Training ran for up to 40 epochs with early stopping (patience = 5, minimum improvement = 10^{-6}), meaning training halts if validation loss does not improve by at least 10^{-6} over 5 consecutive epochs [42]. The loss function was `BCEWithLogitsLoss` [41]:

$$\ell(x, y) = -\big[y \cdot \log \sigma(x) + (1 - y) \cdot \log(1 - \sigma(x))\big], \quad \sigma(x) = \frac{1}{1 + e^{-x}} \tag{7}$$

The Adam optimiser [43] used separate learning rates for the backbone (10^{-4}) and MLP head (10^{-3}), with L2 weight decay of 10^{-4} and gradient clipping at a maximum norm of 1.5. All models were trained on a GPU.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_5

3.4 Evaluation Metrics

All models were trained and evaluated on the same pre-split dataset, with a fixed random seed to ensure a fair comparison. They were assessed using standard classification metrics, including Precision, Recall, F1 score, Area Under the Receiver Operating Characteristic (AUROC) curve, and Average Precision (AP) from the precision-recall curve. For a binary class i, these metrics are defined as:

$$Precision_i = \frac{TP_i}{TP_i + FP_i}, \quad Recall_i = \frac{TP_i}{TP_i + FN_i} \tag{8}$$

$$F1_i = 2 \cdot \frac{Precision_i \cdot Recall_i}{Precision_i + Recall_i} \tag{9}$$

4 Results

All models triggered early stopping at epoch 6 (except Model 7 at epoch 13), indicating rapid overfitting given the dataset size and suggesting that backbone capacity is constrained by the availability of labelled training data. Table 3 summarises performance across all eight variants ranked by AP. Model 3, a Siamese ResNet-101 with a four-layer MLP head, achieved the highest AP (0.888), AUROC (0.911), and test accuracy (0.851), indicating the strongest discrimination between damaged and non-damaged patches across all evaluated configurations.

Model 3's confusion matrices (Fig. 4) confirm high damaged-patch recall (D-Recall = 0.907) alongside accurate non-damaged classification (ND-Recall = 0.794), with a slight bias toward the damaged class attributable to training-time oversampling. The ROC curve (Fig. 5a) lies consistently above all other models, confirming the highest AUROC, while the precision-recall curve (Fig. 5b) shows a performance advantage that widens notably at recall values above 0.80, indicating an optimal balance of sensitivity and specificity for patch-level building damage detection.

Table 3. Model performance ranked by average precision (AP), Area Under the Curve (AUC), and Accuracy (ACC). ND = no-damage class, D = damage class, P = precision, R = recall.

Model	AP	AUC	ACC	ND-P	ND-R	ND-F1	D-P	D-R	D-F1
3	**0.888**	**0.911**	**0.851**	0.895	0.794	0.842	0.815	0.907	0.859
2	0.878	0.900	0.833	0.865	0.789	0.825	0.806	0.877	0.840
6	0.876	0.906	0.832	0.866	0.785	0.824	0.804	0.878	0.839
7	0.874	0.899	0.839	0.872	0.795	0.832	0.812	0.883	0.846
0	0.872	0.896	0.834	0.850	0.811	0.830	0.819	0.857	0.838
1	0.870	0.902	0.837	0.882	0.778	0.827	0.801	0.896	0.846
5	0.869	0.904	0.838	0.883	0.778	0.827	0.802	0.897	0.847
4	0.865	0.897	0.834	0.869	0.788	0.826	0.806	0.881	0.842

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_5

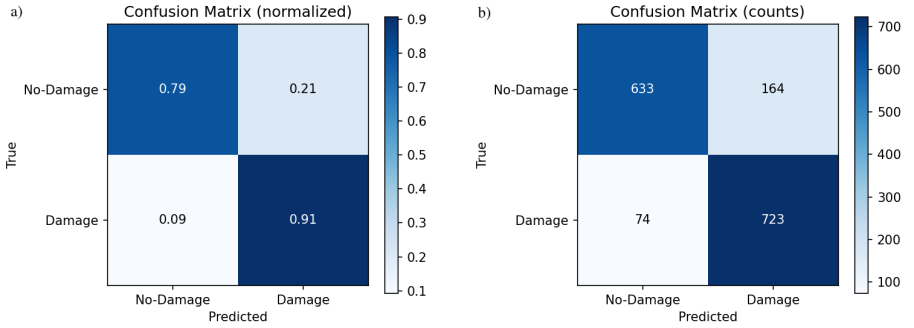

Fig. 4. Confusion matrices for Model 3: (a) row-normalised per-class recall rates (D-Recall = 0.907, ND-Recall = 0.794); (b) absolute prediction totals across No-Damage and Damage classes.

The model exhibited robustness to seasonal variation, successfully distinguishing genuine damage signals from spectral differences attributable to vegetation and phenological change between pre- and post-conflict acquisitions. The underperformance of deeper ResNet-152 variants relative to ResNet-101 is most plausibly attributed to dataset scale: larger parameter counts are harder to fine-tune without overfitting on limited labelled data, corroborated by the rapid early-stopping behaviour observed across all variants.

5 Discussions

5.1 Displacement Planning Applications

In the immediate aftermath of conflict escalation, outward displacement planning requires rapid identification of unsafe areas. The model's damage probability map, generated at 64×64-pixel resolution (approximately 0.41 km^2 footprints), provides a spatially continuous raster suitable for prioritising evacuation corridors, identifying safe passage routes, and allocating emergency resources. Due to training-time oversampling, the model exhibits a slight bias toward the damaged class, producing fewer false negatives at the expense of some false positives, an operationally appropriate trade-off in emergency contexts where underestimating danger carries higher costs than issuing precautionary alerts. A single Sentinel-2 scene covering roughly 100 km $\times$ 100 km can be processed within hours of acquisition, enabling near-real-time spatial intelligence infeasible through ground-based surveys in active conflict zones, while the five-day revisit frequency supports near-continuous monitoring of damage dynamics.

For inward displacement and return planning, spatially comprehensive damage assessments are equally critical. UNHCR intention surveys in Ukraine (2022–2024) consistently identify housing damage and physical safety as primary barriers to return [26]. Housing condition is also a strong predictor of reintegration

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_5

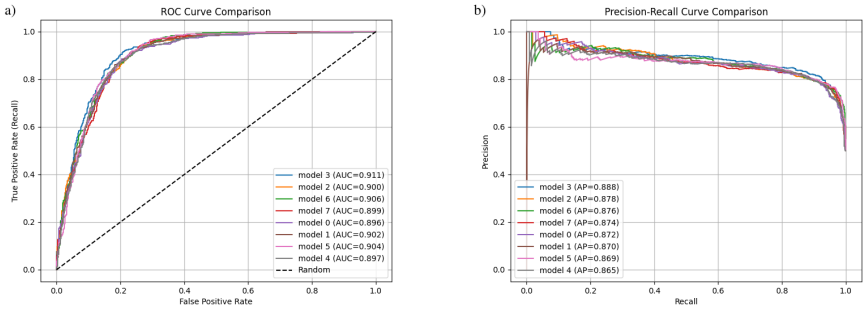

Fig. 5. Evaluation curves for all models: (a) ROC curves, with Model 3 achieving the highest AUROC = 0.911; (b) precision-recall curves, with Model 3 achieving the highest AP = 0.888.

success [44], while delayed return imposes measurable economic burdens on host communities [45]. The Siamese CNN model provides a regularly updatable, large-scale building damage assessment that can complement ground-based housing surveys and inform return facilitation strategies. Given the model's false positive rate, outputs should be treated as an initial spatial screening tool rather than definitive household-level evidence, used to prioritise secondary field verification, the approach can substantially reduce the cost and time of systematic damage assessment, enabling coverage of conflict-affected territories that would otherwise exceed humanitarian resource capacity.

5.2 Integration with Agent-based Displacement Models

A natural extension of this work is integration with an agent-based displacement simulation toolkit, namely Flee [21]. The damage probability maps generated by the Siamese CNN could serve as dynamic data layers within ABM simulations, providing near real-time information about building damage intensity that influences agent decision-making and route accessibility within the simulation environment. This integration would address one of the primary data limitations of current ABM-based displacement models: the difficulty of obtaining timely, spatially granular information about conflict-affected infrastructure. Conceptually, the approach extends the framework proposed by Suleimenova et al. [23], who identify CNNs applied to satellite imagery as a mechanism for updating route accessibility parameters in Flee simulations. In addition, the potential for continuous integration, drawing inspiration from real-time detection frameworks [46], could enable ABMs to ingest dynamically evolving damage layers as new Sentinel-2 imagery becomes available, rather than relying on periodic static snapshots.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_5

5.3 Limitations and Future Work

Several limitations constrain the generalisability and operational readiness of the presented approach. The model was trained and evaluated exclusively on Ukrainian conflict data, and transfer to regions with substantially different settlement patterns, building typologies (e.g., informal settlements, mud-brick construction), or environmental backgrounds may degrade performance and requires explicit cross-theatre validation. Pre- and post-conflict image pairs acquired under different seasonal and atmospheric conditions can introduce spectral artefacts that confound damage signals. Z-score normalisation mitigates some radiometric variation, but residual seasonal and phenological differences, particularly between summer pre-event and winter post-event acquisitions, may influence predictions. A further source of noise arises from the resolution mismatch between UNOSAT annotations derived from VHR commercial imagery and the Sentinel-2 patches used for training: the 50 m spatial buffering applied to convert building-level annotations to patch-level labels introduces boundary noise and may cause undamaged patches near damaged areas to be mislabelled. At the granularity of a 64×64-pixel Sentinel-2 patch (approximately $0.41\,\mathrm{km}^2$), assessments are necessarily at neighbourhood-block rather than individual building level, limiting utility for fine-grained return planning at the household scale. Finally, although the preprocessing pipeline applies cloud masking via the Sentinel-2 Scene Classification Layer, prolonged cloud cover can introduce temporal gaps in damage monitoring that may be operationally critical.

Several promising directions for future work emerge from these limitations. Expanding training data to additional conflict regions, including Syria, Yemen, Gaza, and the Iran-Iraq border region, where sustained conflict has produced large-scale infrastructure damage and Sentinel-2 imagery is freely available, would enable rigorous cross-theatre generalisation assessment and is a prerequisite for operational deployment beyond Ukraine. The consequences of regional conflict dynamics on Iran and Persian Gulf states represent a further candidate application context given the absence of systematic medium-resolution damage detection studies in these areas. Integrating Sentinel-1 SAR data as a complementary input modality [27] would mitigate the cloud-cover limitation and improve damage signal robustness. Architectural enhancements including attention mechanisms, recurrent classification heads for temporal feature integration, and multi-scale feature pyramids offer clear pathways for improving performance on medium-resolution imagery [31,15]. Finally, end-to-end integration of the damage detection pipeline with ABM displacement simulations represents an important applied research target with direct humanitarian impact.

6 Conclusion

This paper presented a Siamese CNN approach for binary building damage detection using freely available Sentinel-2 medium-resolution satellite imagery, motivated by the operational needs of humanitarian displacement planning in conflict-affected regions. Eight model variants combining four ResNet backbone

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_5

architectures with a four-dimensional combined feature representation and MLP classification heads were systematically evaluated on 7,968 balanced patches spanning 22 Ukrainian cities and regions annotated with UNOSAT damage assessments. The best-performing Siamese ResNet-101 achieved an AUROC of 0.911 and AP of 0.888 on the held-out test split, demonstrating strong damage detection capability from publicly available medium-resolution data without reliance on costly commercial VHR imagery. Architectural analysis revealed that ResNet-101 with a four-layer MLP head provides the optimal capacity-regularisation balance at this dataset scale, with deeper ResNet-152 backbones yielding diminishing returns due to limited training data. We further analysed the operational utility of the resulting damage probability maps for both outward and inward humanitarian displacement planning and discussed integration pathways with agent-based displacement simulations, establishing a compelling foundation for resource-efficient, scalable conflict damage assessment using open satellite data.

Data Availability Statement. The code for all eight model variants and hardware specifications are available at `https://github.com/TMatarova/S2_urban_damage_detector_siamese`. The dataset and pretrained weights for the best-performing model are additionally archived on Zenodo at `https://zenodo.org/records/19441105`.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. UNHCR: Figures at a Glance. UNHCR, Geneva (2025). Available at: `https://www.unhcr.org/about-unhcr/overview/figures-glance`
2. IOM: Ukraine Internal Displacement Report. General Population Survey. International Organization for Migration (2024). Available at: https://dtm.iom.int/reports/ukraine-internal-displacement-report
3. Black, R., Gent, S.: Sustainable return in post-conflict contexts. Int. Migr. **44**(3), 15–38 (2006)
4. Özerdem, A., Payne, L.: Ethnic minorities and sustainable refugee return and reintegration in Kosovo. Confl. Secur. Dev. **19**(4), 403–425 (2019)
5. Ramazani, R., et al.: War, displacement, and the best location for temporary sheltering: A qualitative study. BMC Public Health **22**(1), 2099 (2022)
6. Lu, X., Bengtsson, L., Holme, P.: Predictability of population displacement after the 2010 Haiti earthquake. Proc. Natl. Acad. Sci. USA **109**(29), 11576–11581 (2012)
7. Winkler, N.E., et al.: Critical failings in humanitarian response: A cholera outbreak in Kumer Refugee Camp, Ethiopia. BMJ Glob. Health **9**(12), e015585 (2024)
8. Gupta, R., et al.: Creating xBD: A dataset for assessing building damage from satellite imagery. In: Proc. IEEE/CVF CVPR Workshops (2019)
9. European Space Agency: Sentinel-2 User Handbook. ESA Standard Document, Issue 1 Rev. 2 (2015). Available at: `https://sentinels.copernicus.eu/documents/247904/685211/Sentinel-2_User_Handbook`

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_5

10. Copernicus Data Space Ecosystem: Copernicus Data Space Ecosystem (2025). Available at: https://dataspace.copernicus.eu
11. Bromley, J., Bentz, J.W., Bottou, L., Guyon, I., LeCun, Y., Moore, C., Säckinger, E., Shah, R.: Signature verification using a "Siamese" time-delay neural network. In: Series in Machine Perception and Artificial Intelligence, pp. 25–44. World Scientific (1994)
12. Chopra, S., Hadsell, R., LeCun, Y.: Learning a similarity metric discriminatively, with application to face verification. In: Proc. IEEE CVPR, vol. 1, pp. 539–546 (2005)
13. Daudt, R.C., Le Saux, B., Boulch, A.: Fully convolutional Siamese networks for change detection. In: Proc. IEEE ICIP, pp. 4063–4067 (2018)
14. Yang, L., Chen, Y., Song, S., Li, F., Huang, G.: Deep Siamese networks based change detection with remote sensing images. Remote Sens. **13**(17), 3394 (2021)
15. Wu, C., et al.: Building damage detection using U-Net with attention mechanism from pre- and post-disaster remote sensing datasets. Remote Sens. **13**(5), 905 (2021)
16. UNITAR-UNOSAT: Ukraine: Building Damage Assessment Maps (Collection). United Nations Satellite Centre, Geneva (2022). Available at: https://unosat.org/products
17. Suleimenova, D., Bell, D., Groen, D.: A generalized simulation development approach for predicting refugee destinations. Sci. Rep. **7**, 13377 (2017)
18. Schon, J.: Motivation and opportunity for conflict-induced migration: An analysis of Syrian migration timing. J. Peace Res. **56**(1), 12–27 (2018)
19. Urbański, M.: Comparing push and pull factors affecting migration. Economies **10**(1), 21 (2022)
20. Hoffmann Pham, K., Luengo-Oroz, M.: Predictive modelling of movements of refugees and internally displaced people: Towards a computational framework. J. Ethn. Migr. Stud. **49**(2), 408–444 (2022)
21. Ghorbani, M., Suleimenova, D., Jahani, A., Saha, A., Xue, Y., Mintram, K., Anagnostou, A., Tas, A., Low, W., Taylor, S.J.E., Groen, D.: Flee 3: Flexible agent-based simulation for forced migration. J. Comput. Sci. **81**, 102371 (2024)
22. Huynh, B.Q., Basu, S.: Forecasting internally displaced population migration patterns in Syria and Yemen. Disaster Med. Public Health Prep. **14**(3), 302–307 (2019)
23. Suleimenova, D., Xue, Y., Tas, A., Low, W.: AI-enhanced agent-based modelling approach for forced displacement predictions. In: ICCS 2025 Workshops. LNCS, vol. 15911, pp. 79–86. Springer, Cham (2025)
24. Dehkordi, M.A.E., Lechner, J., Ghorbani, A., Nikolić, I., Chappin, É.: Using machine learning for agent specifications in agent-based models and simulations: A critical review and guidelines. J. Artif. Soc. Soc. Simul. **26**(1), 9 (2023)
25. Monti, R.P., Dasgupta, T., Cattaneo, M., Pan, W.: On learning agent-based models from data. Sci. Rep. **13**(1), 9268 (2023)
26. UNHCR: Lives on Hold: Intentions and Perspectives of Refugees, Refugee Returnees and IDPs from Ukraine #5. UNHCR, Geneva (2024). Available at: https://data.unhcr.org/en/documents/details/106738
27. 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. **14**(24), 6239 (2022)
28. Papadomanolaki, M., Vakalopoulou, M., Karantzalos, K.: Detecting urban changes with recurrent neural networks from multitemporal Sentinel-2 data. In: Proc. IEEE IGARSS, pp. 4971–4974 (2019)

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_5

29. Papadomanolaki, M., Vakalopoulou, M., Karantzalos, K.: Urban change detection based on semantic segmentation and fully convolutional LSTM networks. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. V-2-2020, 541–547 (2020)
30. Sticher, V., Wegner, J.D., Pfeifle, B.: Toward the remote monitoring of armed conflicts. PNAS Nexus **2**(6) (2023)
31. Chen, H., Wu, C., Du, B., Zhang, L., Wang, L.: Change detection in multisource VHR images via deep Siamese convolutional multiple-layers recurrent neural network. IEEE Trans. Geosci. Remote Sens. **58**(4), 2848–2864 (2020)
32. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proc. IEEE CVPR, pp. 770–778 (2016)
33. Nagpal, P., Bhinge, S.A., Shitole, A.: A comparative analysis of ResNet architectures. In: Proc. IEEE SMART GENCON (2022)
34. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: Proc. ICLR (2015)
35. Elharrouss, O., Akbari, Y., Almaadeed, N., Al-Maadeed, S.: Backbones-review: Feature extractor networks for deep learning and deep reinforcement learning approaches. Comput. Sci. Rev. **53**, 100645 (2024)
36. National Aeronautics and Space Administration: Reflected Near-Infrared Waves. NASA Science (2010). Available at: `https://science.nasa.gov/ems/08_nearinfraredwaves`
37. GDAL/OGR Contributors: GDAL/OGR Geospatial Data Abstraction Software Library. Open Source Geospatial Foundation (2023). Available at: `https://gdal.org`
38. Conda Developers: Managing environments. Conda documentation (2025). Available at: `https://docs.conda.io/projects/conda/en/stable/user-guide/tasks/manage-environments.html`
39. QGIS Development Team: QGIS Geographic Information System. Open Source Geospatial Foundation (2023). Available at: `https://qgis.org`
40. Lee, M.: Mathematical analysis and performance evaluation of the GELU activation function in deep learning. J. Math. (2023)
41. PyTorch. PyTorch documentation (2025). Available at: `https://docs.pytorch.org`
42. Paszke, A., et al.: PyTorch: An imperative style, high-performance deep learning library. In: Advances in Neural Information Processing Systems, vol. 32, pp. 8024–8035 (2019)
43. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv:1412.6980 (2014)
44. Zavisca, J.R., Mitchneck, B., Gerber, T.P.: Housing and integration of internally displaced persons: The case of Ukraine in 2018. Front. Hum. Dyn. **5** (2023)
45. Verme, P., Schuettler, K.: The impact of forced displacement on host communities: A review of the empirical literature in economics. J. Dev. Econ. **150**, 102606 (2021)
46. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time object detection. In: Proc. IEEE CVPR, pp. 779–788 (2016)

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_5