International Journal of Disaster Risk Reduction 143 (2026) 106263

Contents lists available at ScienceDirect

International Journal of Disaster Risk Reduction

journal homepage: www.elsevier.com/locate/ijdrr

MAS-LAND: A multi-agent system for landslide detection and rapid response

Abdelkader Riche [a,*], Abdallah Zeggada [b], Francesco Caleca [a], Mohammed Alruqimi [c], Pierluigi Confuorto [a], Mawloud Guermoui [c], Silvia Bianchini [a], Veronica Tofani [a], Farid Melgani [c]

[a] Department of Earth Sciences, University of Florence, Via La Pira, 4, Florence, 50121, Italy
[b] Independent Researcher, Bari, Italy
[c] Department of Information Engineering and Computer Science, University of Trento, Via Sommarive, 9, Trento, I-38123, Italy

ARTICLE INFO

Keywords:
Landslide detection
Multi-agent systems
SegFormer
Emergency response
LLM-based reporting

ABSTRACT

Landslides require rapid, data-driven post-event assessment to support emergency response, yet most existing approaches focus on stand-alone landslide detection or mapping accuracy, with limited integration into operational decision-support workflows. This study presents MAS-LAND (Multi-Agent System for Landslide Detection and Rapid Response), a multi-agent, LLM-enhanced framework structured around three main stages: (1) post-event landslide detection, (2) infrastructure exposure assessment, and (3) automated reporting for civil protection operations. The framework is applied to the May 2023 Emilia–Romagna disaster and leverages very high-resolution post-event orthophotos and transformer-based semantic segmentation to identify newly occurred landslides. Among the evaluated models, SegFormer MIT-b2 achieved the strongest performance, detecting 603 of 806 landslides in the test dataset (Recall = 74.81%, F1 = 71.14%) and delivering robust multi-class segmentation results (macro F1 = 71.09%) despite pronounced class imbalance.

In the second stage, infrastructure exposure is assessed through automated building footprint extraction using a pretrained RT-DETR model and road-network data derived from OpenStreetMap, enabling deterministic prioritization of intervention needs via a transparent, rule-based decision matrix. In the third stage, a reporting agent powered by Llama-3.3-70B, operated in deterministic mode, synthesizes analytical outputs into standardized operational summaries tailored for civil protection use.

A key contribution of this work is the delivery of a fully integrated, post-event pipeline capable of transforming heterogeneous geospatial inputs into actionable intelligence within a few minutes of image acquisition and processing, as this ability is largely absent from previous landslide studies. The framework also produces structured, machine-readable outputs at every stage, forming a scalable knowledge base for systematic storage and reuse of event information. Overall, the results demonstrate the potential of multi-agent, hybrid AI–geospatial architectures to enhance situational awareness, improve reproducibility, and support timely decision-making of Civil Protection authorities during landslide emergencies.

* Corresponding author.
E-mail addresses: abdelkader.riche@unifi.it (A. Riche), abdellah.zeggada@gmail.com (A. Zeggada), francesco.caleca@unifi.it (F. Caleca), mohammed.alruqimi@unitn.it (M. Alruqimi), pierluigi.confuorto@unifi.it (P. Confuorto), mawloud.guermoui@unitn.it (M. Guermoui), silvia.bianchini@unifi.it (S. Bianchini), veronica.tofani@unifi.it (V. Tofani), farid.melgani@unitn.it (F. Melgani).

https://doi.org/10.1016/j.ijdrr.2026.106263
Received 2 April 2026; Received in revised form 19 May 2026; Accepted 15 June 2026
Available online 16 June 2026
2212-4209/

1. Introduction

Rainfall-induced landslides are among the most damaging natural hazards, frequently causing loss of life, disruption of transportation networks, and extensive damage to buildings and infrastructure. Following extreme rainfall events or earthquakes, civil protection authorities face significant challenges in rapidly identifying newly triggered landslides, understanding their potential impacts, and prioritizing emergency interventions. Field surveys remain essential for ground validation, yet they are time-consuming, resource-intensive, and often infeasible immediately after a disaster, particularly when failures are widespread or access is limited [1]. These constraints have motivated the increasing adoption of remote sensing, GIS, and artificial intelligence (AI) techniques to support rapid post-event landslide detection and mapping.

In the immediate aftermath of a landslide or other damaging event, civil protection authorities must rapidly determine where to intervene first, which infrastructures are threatened, and how to allocate limited operational resources under time pressure. In practice, this phase is often complicated by fragmented information, delayed field verification, limited accessibility, and the absence of unified tools capable of combining geospatial observations with intervention-oriented outputs. As a result, emergency management may rely heavily on manual interpretation, local experience, and sequential communication between services, which can slow decision-making during the most critical hours. This highlights a clear gap between post-event mapping products and the concrete operational needs of civil protection, particularly in relation to intervention prioritization, infrastructure impact assessment, and the production of structured information for rapid response. According to the European civil protection emergency-planning guidelines "https://civil-protection-knowledge-network.europa.eu/system/files/2025-01/emergency-planning-guidelines.pdf", post-event response is typically organized through a sequence of activation, mobilisation, and operations, including rapid situation assessment, deployment of emergency resources, evacuation or access-control measures where needed, and coordinated protection of critical infrastructure and essential services. However, in landslide emergencies these actions are still often supported by broad all-hazard procedures rather than by dedicated, landslide-specific operational protocols, which reinforces the need for tools that can translate post-event mapping into intervention-oriented information.

Recent advances in deep learning (DL) have substantially improved the automation and accuracy of landslide detection from high-resolution imagery. By learning hierarchical spatial representations directly from raw data, DL models reduce reliance on manual feature engineering and enable scalable mapping over large areas [2]. Convolutional Neural Networks (CNNs) have been widely applied for landslide detection and segmentation [3], though their generalization across diverse geomorphological settings remains challenging [4]. Progress in semantic segmentation architectures including U-Net variants, ResU-Net, and DeepLab has driven notable performance gains. For example [5], demonstrated the influence of patch size on U-Net performance, while [6] benchmarked FCN, U-Net, DeepLabv3+, and MFFENet on the CAS landslide dataset. Transformer-based models have further advanced the field [7]: evaluated SegFormer on the globally distributed GDCLD dataset [8], reported improved results on the CAS dataset, and [9] introduced a SAM-based fusion network (SAM-CFFNet) achieving state-of-the-art segmentation performance.

Beyond detection, some studies have begun exploring how textual interpretation and reporting can support landslide analysis [10]. proposed one of the earliest landslide captioning frameworks, combining U-Net with bi-temporal LSTMs to generate natural-language descriptions of landslides and nearby exposed elements. More recently [11], introduced a multimodal large language model trained on expert-transcribed explanations of landslide imagery, highlighting the potential of LLMs for domain-aware interpretation.

In parallel, the disaster-AI community has increasingly adopted multi-agent systems to manage complex, multi-step disaster-response workflows [12]. combined Case-Based Reasoning with multi-agent systems for flood response, while [13] proposed the Adaptive Disaster Interpretation (ADI) paradigm, enabling LLM-driven agents to orchestrate sequential interpretation tasks [14]. emphasized the role of intelligent agents in disaster mitigation under infrastructural constraints, and [15] introduced Disaster Copilot, a coordinated multi-agent architecture for real-time disaster management [16]. further advanced this direction with a retrieval-augmented multi-agent system for hazard adaptation. Within the landslide domain [17], introduced CC-GRMAS, a multi-agent framework structured around prediction, planning, and execution agents that integrates satellite observations and environmental signals to support real-time situational awareness, response planning, and proactive landslide disaster preparedness.

Despite these advances, a critical gap remains in early post-event landslide management. Most existing approaches focus on detection accuracy, susceptibility analysis, or delayed post-event assessment, offering limited support for the rapid, interactive decision-making required by Civil Protection authorities in the hours immediately following extreme rainfall. In particular, few systems integrate landslide detection with infrastructure impact analysis, accessibility constraints, and automated reporting within a unified operational workflow.

Deep learning-based landslide detection studies ([5–9]) primarily focus on segmentation accuracy without incorporating downstream decision-support components, in addition, these approaches formulate landslide detection as a binary segmentation problem, focusing on the presence or absence of landslides without explicitly distinguishing between different types of slope movements. While this formulation is effective for mapping purposes, it limits the ability to support operational decision-making, where different landslide types may require distinct response strategies. Meanwhile multi-agent disaster frameworks [12–17] emphasize coordination and reasoning but do not operate directly on high-resolution geospatial detection outputs nor provide infrastructure-level exposure assessment and automated reporting tailored to civil protection needs. This fragmentation highlights the absence of integrated systems capable of bridging detection outputs with operational decision-making.

To address this gap, this study proposes MAS-LAND, a multi-agent, LLM-enhanced decision-support framework specifically designed for post-event landslide detection and intervention prioritization. The framework was applied to the Emilia–Romagna region in northern Italy, selected as the study area because of the severe May 2023 event and the availability of very high-resolution post-

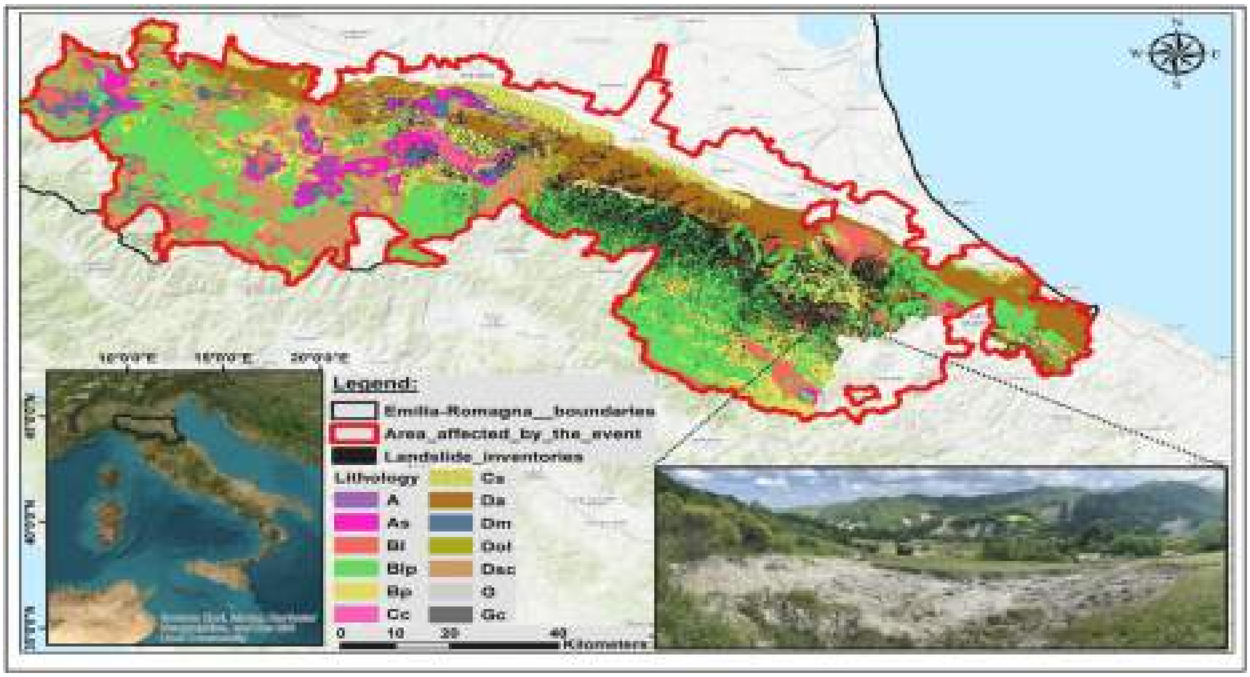


Fig. 1. Location of the study area. The area most affected by the event is displayed over the lithology classes. The left inset shows the location of Emilia–Romagna within Italy, while the lower right inset presents a representative field photo of the study area.

event data and detailed landslide inventories. MAS-LAND starts from post-event aerial imagery, applies deep learning-based segmentation to detect newly occurred landslides, and evaluates their operational relevance according to proximity to buildings and road infrastructure. A coordinated set of agents integrates landslide detection, infrastructure mapping, accessibility analysis, and automated reporting into a single, reproducible pipeline. In the test dataset, the framework delivered strong results, with SegFormer MIT-b2 detecting 603 of 806 landslides and providing robust multi-class segmentation performance considering the landslide types, while the multi-agent workflow successfully translated these outputs into prioritized intervention maps and standardized operational reports. Rather than relying on abstract risk scores, intervention relevance is determined using exposure-driven criteria aligned with real civil protection practices. Although demonstrated on the Emilia–Romagna case study, the workflow is designed to be transferable and can be adapted to future events and other geographic regions where suitable post-event imagery and infrastructure data are available.

2. Dataset and study area

In May 2023, the Emilia–Romagna region in northern Italy (as shown in Fig. 1) experienced two exceptional rainfall episodes that triggered widespread flooding and thousands of landslides. The first event, occurring between 1 and 3 May, delivered nearly 200 mm of rain within 48 h, followed by a second intense episode on 16–17 May, during which 200–250 mm of precipitation fell over largely overlapping areas [18,19]. These successive storms caused more than 80,000 landslides in areas classified as highly to extremely vulnerable in the map provided by Ref. [20]. The resulting cascading impacts were severe, with 17 fatalities and the displacement of approximately 36,600 people impacts largely driven by widespread flooding along with more than 600 road closures and extensive damage to buildings, agricultural land, and critical infrastructure. These consequences are documented by the European Climate Adaptation Platform"https://climate-adapt.eea.europa.eu/en/metadata/case-studies/mental-health-support-for-people-affected-by-floods-in-emilia-romagna". The dataset used in this study was cropped as patches tiles of 1000*1000 pixels with a spatial resolution of 0.2 m (as shown in Fig. 2) integrates post-event very–high-resolution orthophotos, topographic derivatives, environmental layers, and manually annotated landslide polygons [18]. Together, these sources provide a comprehensive representation of the terrain, surface conditions, and observed landslide impacts following the May 2023 Emilia–Romagna disaster.

2.1. Orthophoto images

Post-event aerial orthophotos were obtained from the Emilia–Romagna Geoportal. These four-band images (RGB + NIR) were captured in May 2023, only days after the disaster, with a spatial resolution of 0.2 m. The acquisition took place after the two major rainfall events that occurred between 1–3 May and 16–17 May 2023, and therefore the imagery reflects the cumulative post-event conditions following both triggering episodes. The dataset was commissioned specifically for emergency response and is publicly

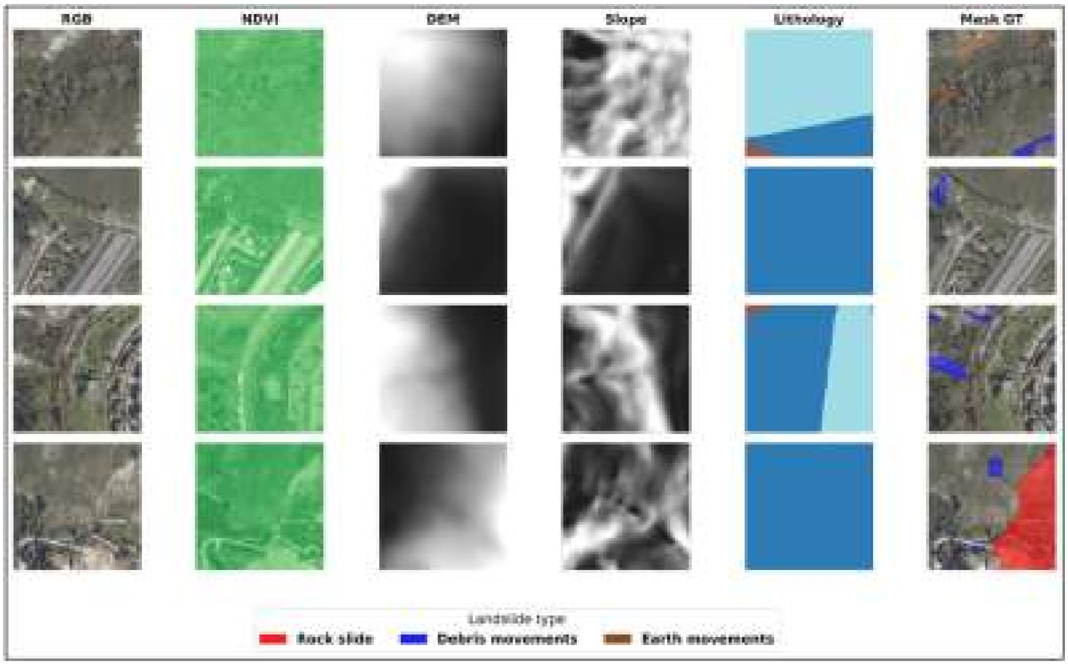


Fig. 2. Overview of the dataset used in this study, showing the RGB samples, their corresponding geo-auxiliary data, and the annotated landslide inventory.

accessible through the regional mapping platform: "https://geoportale.regione.emilia-romagna.it/catalogo/dati-cartografici/cartografia-di-base/immagini/layer-22" Their high spatial detail enables precise delineation of landslide boundaries and near-surface ground disturbance patterns.

2.2. Geo-layers factors

2.2.1. Digital elevation model (DEM)

A 5 m-resolution DEM from the Emilia–Romagna Geoportal https://geoportale.regione.emilia-romagna.it/catalogo/dati-cartografici/altimetria/layer-2 was resampled to 0.2 m to match the orthophotos. Elevation modulates gravitational potential, drainage pathways, and failure distribution. According to ISTAT https://www.istat.it/storage/rapporti-tematici/territorio2020/appendice.pdf?utm terrain classification, elevation was grouped into: (I) Lowland/Plains: <230 m, (II) Upper Plains: 230–<300 m, (III) Lower Hills: 300–<370 m, (IV) Upper Hills: 370–<700 m, (V) Mountains: ≥700 m.

2.2.2. Slope

Slope was calculated from the DEM as the planar gradient (°). Since terrain steepness controls shear stress and material storage, according [21] slope was categorized as follow: (I) Flat: <5°, (II) Gentle: 5–<10°, (III) Moderate: 10–<25°, (IV) Steep: 25–<35°, (V) Very Steep: ≥35°.

2.2.3. Normalized difference vegetation index (NDVI)

Vegetation cover was derived from the Red and NIR orthophoto bands using the standard NDVI formulation [22], expressed by the following equation: NDVI = (NIR − Red)/(NIR + Red). In our study the threshold of NDVI classes was according [23], as: (I) No vegetation: −1 to 0.199, (II) Low vegetation: 0.2 to 0.5, (III) Dense vegetation: 0.501 to 1.

2.2.4. Lithology

Lithological information according Regional geological map of the Emilia-Romagna region, scale 1:10,000 was rasterized to 0.2 m to match the orthophotos. Since lithology controls material strength, permeability, fabric, and failure style, units were grouped into the following engineering-based categories: A – Massive Rock Material, As – Stratified Rock Material, Bl – Alternations with Prevalent Lithoid Layers, Blp – Alternations with Balanced Lithoid & Pelitic Layers, Bp – Alternations with Prevalent Pelitic Layers, *Cc* – Weakly Cemented Clast-Supported Conglomerates/Breccias, *Cm* – Matrix-Supported Conglomerates/Breccias, Cs – Weakly Cemented Sands/Sandstones, Dm – Marls, Da – Clays, Marly Clays, Silty Clays, Dsc – Argillites ("Scaly Clays"), Dol – Clays with Primary Chaotic Structure, Eg – Predominantly Gravel Deposits, Egs – Gravel–Sand Mixtures, Es – Predominantly Sandy Deposits, Esl – Sandy–Silty

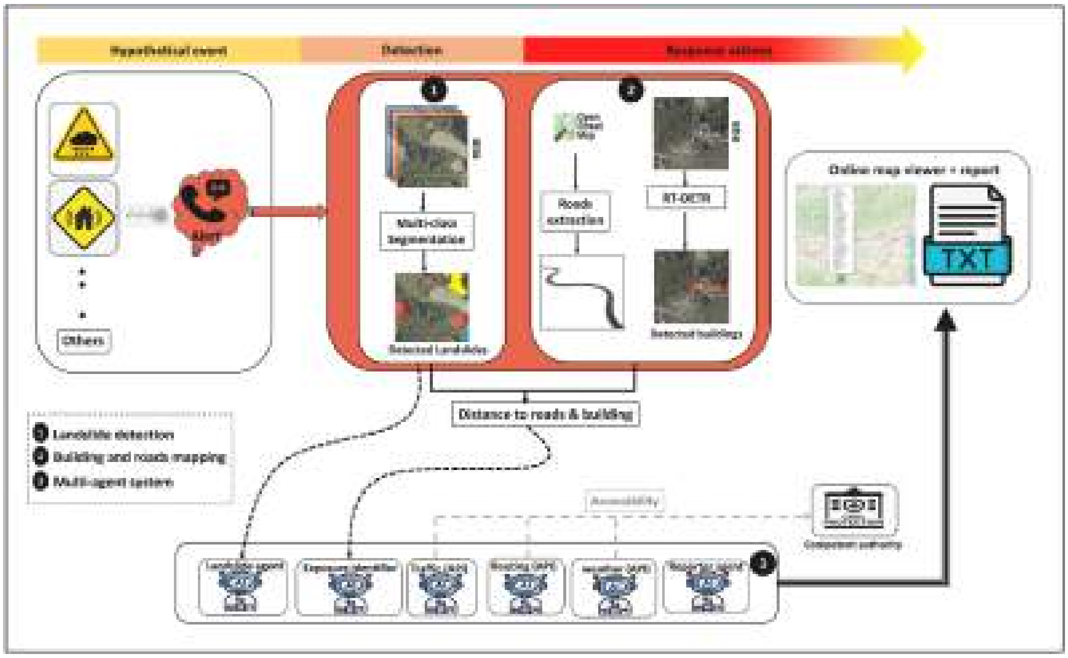


Fig. 3. Overview of the MAS-LAND pipeline. Following a triggering event, post-event RGB imagery is used for landslide segmentation, while buildings and roads are extracted to assess infrastructure exposure. A multi-agent system integrates detection, exposure analysis, accessibility, and LLM-based reporting to generate an interactive map and operational report for civil protection authorities.

Deposits, Ea – Predominantly Clayey Deposits, Elm – Predominantly Silty Deposits, El – Silty–Clayey Deposits, Em – Mixed-Grained Deposits, Ec – Cemented/Very Compact Detritus, Ea – Bedrock Weathering Mantle, FT – Intensely Fractured Rock/Cataclasite, G – Gypsum Rocks.

2.3. *Inventories (segmentation masks)*

Landslide inventory used as target mask for the segmentation procedure was obtained from Ref. [18], who manually mapped all failures triggered during the 2023 event. The dataset, available as polygon shapefiles through Zenodo https://zenodo.org/records/13742643, includes detailed landslide boundaries and type labels: DS1: Debris slide (high mobility), DS2: Debris slide (low mobility), DF1: Debris flow (long runout), DF2: Debris flow (limited runout), RS1: Fully developed rock-block slide, RS2: Incipient rock-block slide, ES: Earth slide, EF: Earth flow. For this study, and following the landslide classification by Ref. [24], classes were merged into broader categories based on shared material characteristics and rheological properties, as these factors provide a more robust basis for aggregation when using optical imagery: Rock slides: RS1, RS2, Debris movements: DS1, DS2, DF1, DF2, Earth movements: ES, EF. In this study we selected 10,103 mapped landslides corresponding to locations where post-event orthophoto imagery was available for model development. Among these, rock-slides account for 62 instances (50 for training, 8 for validation, and 4 for testing), representing 0.61% of the inventory; Debris movements comprise 7965 instances (6146 train, 1178 validation, 641 test), representing 78.9%; and Earth movements include 2076 instances (1631 train, 284 validation, 161 test), representing 20.55% of all mapped failures. This distribution reflects the geomorphological characteristics of the 2023 Emilia–Romagna disaster, where debris-related processes were dominant, and supports robust segmentation and downstream risk-modelling tasks.

3. Methodology

Following an extreme triggering event such as intense rainfall, earthquakes, or other destabilizing processes which motivates the acquisition of post-event aerial imagery, this study proposes an integrated methodological framework composed of three main stages (as shown in Fig. 3): (i) post-event landslide detection, supported by multi-class segmentation from RGB orthophotos to distinguish among different landslide types, (ii) automated landslide reporting using a large language model, and (iii) a multi-agent system that supports post-disaster assessment and intervention prioritization, accompanied by interactive geospatial maps. The triggering event serves solely as the contextual condition for activating the workflow and is not explicitly modelled as a computational stage. The segmentation stage relies exclusively on RGB imagery to ensure immediate applicability after a disaster, while auxiliary geo-environmental layers are incorporated in subsequent steps for interpretation, exposure analysis, and geotechnical characterization. The detailed methodology of each stage is described in the following sections.

3.1. Multi-class segmentation

To enable efficient model training on very high-resolution imagery, the data were partitioned into fixed-size patches of 1000 × 1000 pixels (200 × 200 m at 0.2 m spatial resolution). The selected patch size represents a compromise between preserving sufficient geomorphological context, enabling exposure identification, and maintaining computational feasibility. The mapped landslides in the dataset range from approximately 1 m^2 to more than 20,000 m^2 (average ≈ 1035 m^2; median ≈ 468 m^2), therefore the adopted patch extent allows most landslide bodies, surrounding terrain context, and nearby exposed elements such as roads and buildings to be represented within a single spatial window. Smaller patches could truncate large or elongated landslides and omit nearby exposed elements, whereas larger patches would substantially increase computational cost. Although explicit multi-scale training was not implemented, both the convolutional and transformer-based architectures employed in this study inherently incorporate hierarchical multi-scale feature extraction mechanisms.

3.1.1. U-net++ model

The original U-Net architecture [25] introduced an encoder–decoder structure with skip connections that preserve spatial information while capturing high-level semantics. U-Net++ [26] enhances these skip pathways by replacing them with densely connected convolutional blocks that gradually align encoder and decoder features, reducing the semantic gap. This refinement improves boundary precision, which is essential for delineating complex landslide shapes.

In this study, U-Net++ uses EfficientNet-B2 and EfficientNet-B3 encoders. EfficientNet architectures scale depth, width, and resolution uniformly [27], providing strong feature extraction efficiency. EfficientNet-B2 contains approximately 9 million trainable parameters, while EfficientNet-B3 includes around 12 million, enabling finer representation of textural variations in the high-resolution orthophotos.

3.1.2. SegFormer and its backbone variants

SegFormer [28] is a transformer-based architecture coupling a MiT (Mix Transformer) encoder with a lightweight MLP decoder. The encoder uses overlapping patch embeddings and hierarchical self-attention to capture both local textures and long-range spatial dependencies, making it suitable for heterogeneous terrain. The decoder fuses multi-scale features with minimal computational cost, achieving high-resolution segmentation with sharp boundaries.

SegFormer provides a set of backbone variants with increasing representational power. MiT-B0 contains around 3–4 million parameters and is optimized for speed. MiT-B2 and MiT-B3 increase capacity to approximately 27 million and 47 million, respectively. Larger variants such as MiT-B4 and MiT-B5 exceed 60–80 million parameters and offer stronger contextual modeling, albeit with higher computational requirements. This range allows evaluation across models from lightweight to high-capacity configurations.

3.1.3. Activation and loss function

For multi-class segmentation, pixel-wise class probabilities are produced using the softmax activation function,

$$p_{i,c} = \frac{e^{z_{i,c}}}{\sum_{k=1}^{C} e^{z_{i,k}}}, \tag{1}$$

where $z_{i,c}$ is the logit for pixel iii and class c, and C is the total number of classes. To handle class imbalance and improve boundary precision, a hybrid loss combining weighted cross-entropy and multi-class *Dice* loss is adopted:

$$l = 0.5\, l_{CE}^{weighted} + 0.5\, l_{Dice} \tag{2}$$

The weighted cross-entropy term was implemented using fixed class weights derived from the class distribution in the landslide inventory. Rock slides represent only 0.61% of the dataset, compared to 78.9% for debris movements and 20.55% for earth movements. To compensate for this imbalance, higher weights were assigned to underrepresented classes, resulting in the following configuration: background = 0.5, rock slides = 6.0, debris movements = 1.0, and earth movements = 2.5. The highest weight was assigned to rock slides due to their extreme rarity, while an intermediate weight was used for earth movements to account for their lower frequency relative to debris movements. This weighting strategy ensures that minority classes contribute effectively to the optimization process without destabilizing training. A label-smoothing factor of 0.05 was additionally applied to improve generalization.

The Dice component measures the overlap between predictions and ground truth:

$$\begin{aligned} Dice_c &= \frac{2\sum_i p_{i,c} y_{i,c} + \varepsilon}{\sum_i p_{i,c} + \sum_i y_{i,c} + \varepsilon'}, \\ l_{Dice} &= 1 - \frac{1}{C}\sum_{c=1}^{C} Dice_c \end{aligned} \tag{3}$$

Here, $p_{i,c}$ is the predicted probability for class c, $y_{i,c}$ is the one-hot encoded ground truth, the summation $\sum_i$ runs over all pixels, and ε is a small constant to avoid division by zero. This combined objective stabilizes learning, reinforces the contribution of rare classes through weighted cross-entropy, and improves segmentation coherence, particularly along complex landslide boundaries.

3.2. Landslide reporting

Automated textual reporting is performed using Llama-3.3-70B-Versatile (Meta AI, 2024), which transforms segmentation outputs and geo-attributes into structured summaries describing each detected landslide. The reporting agent operates only on attributes derived from earlier geospatial processing, without recomputing any physical or environmental variables. The reporting workflow includes.

3.2.1. Landslide size classification

Landslide area is computed for each polygon and classified According to the universal scheme proposed by Ref. [29]: Very small: $<10^2\,m^2$, Small: $<10^4\,m^2$, Medium: $<10^6\,m^2$, Large: $<10^8\,m^2$, Giant: $<10^{10}\,m^2$, Monster: $\geq 10^{10}\,m^2$.

3.2.2. Spatial localization and geo-environmental attribute extraction

For each detected landslide polygon, the centroid is computed to derive geographic coordinates (latitude and longitude), ensuring precise spatial traceability and direct compatibility with field operations and civil protection workflows. In parallel, a set of geo-environmental attributes is extracted and provided to the LLM for structured reporting. These descriptors include elevation and slope derived from the DEM, lithology determined as the dominant geological unit intersecting the landslide polygon, and vegetation conditions characterized through NDVI values computed within a 3 m buffer surrounding the landslide. The use of a buffer zone allows the environmental context to be captured, while avoiding spectral contamination from disturbed surfaces within the landslide body itself. Together, these spatial and geo-environmental attributes support geotechnically and environmentally informed landslide reporting.

3.2.3. Distance to buildings and roads

Building footprints are detected using the pretrained RT-DETR model finetuned for satellite roof detection: https://huggingface.co/Yifeng-Liu/rt-detr-finetuned-for-satellite-image-roofs-detection. Because the segmentation workflow operates on tiled orthophoto patches, building detection is performed independently on each image tile. Several confidence thresholds were tested, and 0.55 provided the best trade-off, effectively reducing false positives. The minimum distance between each landslide polygon and detected buildings is computed to infer structural exposure. Road networks were acquired from OpenStreetMap (OSM):" https://www.openstreetmap.org2". Although roads are not explicitly visible within some orthophoto tiles, the OSM shapefiles provide complete network coverage across the region. The minimum distance to roads is calculated for each landslide polygon, providing an estimate of potential transport disruption and accessibility constraints.

3.3. Multi-agent system

The multi-agent component of the framework is designed to coordinate the full post-disaster workflow following a triggering event capable of inducing landslides, such as intense rainfall or earthquakes. The occurrence of such an event motivates the acquisition of post-event aerial imagery, which serves as the primary input for downstream analysis.

Once new imagery becomes available, the multi-agent system orchestrates the sequential processing of landslide detection, exposure assessment, and intervention prioritization. The workflow transitions from automated landslide mapping toward evaluating potential impacts on exposed elements, including buildings, road networks, and critical infrastructure, thereby supporting operational decision-making in the immediate aftermath of an event. All interactions between agents are governed by a hybrid architecture that combines deterministic geospatial analysis with targeted LLM-supported reasoning.

This design ensures that safety-critical computations such as spatial proximity analysis and intervention prioritization remain transparent, reproducible, and fully auditable, while language-based agents are used exclusively for interpretation, coordination, and

Table 1
Proposed decision-support matrix for post-landslide intervention prioritization.

Priority Level	Distance to Building or Roads (m)	Slope & Debris Movement Influence	Operational Description
P1 – Very Urgent	0–1	Steep/Very Steep Slope & Debris confirm P1	- Immediate civil protection action, evacuations, road closures, emergency stabilization. - Information must also be shared with the population through official communication channels.
P2 – Urgent	1–2	Steep/Very Steep Slope & Debris → P1	- Urgent inspection by civil protection; rapid technical assessment; potential emergency works depending on field conditions.
P3 – High Priority	2–5	Steep/Very Steep Slope & Debris → P2	- Inspection in 1–3 days; evaluate runout; technical experts decide on any escalation.
P4 Medium Priority	5–20	Steep/Very Steep Slope & Debris → P3	- Medium-term mitigation planning and periodic UAV monitoring.
P5 – Low Priority	20–50	Steep/Very Steep Slope & Debris → P4	- Low threat; included in inventory; low-frequency monitoring.
P6 Monitoring Only	>50	Steep/Very Steep Slope & Debris remains P6	- Remote monitoring only (UAV, InSAR). Long-term hazard tracking.

structured reporting. MAS-LAND is implemented as a LangGraph-based deterministic multi-agent workflow in which each agent performs a predefined operational task. Agents are instantiated as task-specific Python functions registered as nodes within the LangGraph execution graph. Inter-agent communication is performed primarily through a shared workflow state object rather than peer-to-peer conversational message passing.

3.3.1. Supervisor Agent (LangGraph ReAct-Inspired orchestrator)

The Supervisor Agent manages workflow sequencing and state-aware coordination within a LangGraph-based deterministic multi-agent workflow. Unlike unconstrained autonomous ReAct systems, the Supervisor does not perform open-ended LLM reasoning or autonomous tool selection. Instead, it applies deterministic routing logic inspired by agentic task decomposition and reactive workflow orchestration. Given a user or system-level instruction, the supervisor.

✓ interprets workflow requirements,
✓ checks the shared workflow state,
✓ determines which operational agent must be executed next,
✓ dispatches execution to the corresponding agent,
✓ monitors completion states until all mandatory outputs are generated.

The execution sequence involves the Landslide Agent, Weather Agent, Traffic Agent, Route Agent, Exposure/Priority Agent, Reporter Agent, and Validation Agent.

Inter-agent communication is performed primarily through a shared LangGraph state object. A lightweight message state is additionally maintained within LangGraph for initial task instructions and simple execution-status tracking, each agent reads the required fields from the shared state, executes its specialized task, and returns updated fields to the workflow state.

This design preserves traceability, reproducibility, and operational reliability while avoiding unrestricted autonomous reasoning in safety-critical disaster-response scenarios. Within MAS-LAND, the only LLM-based narrative synthesis is performed by the Reporter Agent, while all safety-critical computations, intervention priorities, and exposure assessments remain deterministic and rule-based.

3.3.2. Landslide Agent (deterministic geomorphological processor)

This agent processes each landslide instance by ingesting DEM-derived elevation, slope, NDVI, lithology, building/road proximity, landslide type, and area. It performs all deterministic classifications including vegetation status, elevation class, size class, and lithological group and exposure elements.

3.3.3. Environmental agents (API-based)

The environmental agents supply the system with real-time or near-real-time environmental information that supports post-event reassessment. Their outputs feed directly to the exposure elements and the accessibility conditions from the competent authorities to the occurred landslides that need immediate interventions.

✓ **Weather Agent:** The Weather Agent retrieves meteorological variables, including temperature, wind conditions, current rainfall, and short-term rainfall forecasts for the coming hours, from the Open-Meteo API (https://open-meteo.com/) These parameters provide essential contextual information for civil protection teams, supporting situational awareness during field operations by characterizing evolving weather conditions that may affect access, safety, and the feasibility of intervention activities.
✓ **Traffic Agent:** The Traffic Agent provides real-time mobility conditions essential for assessing terrain accessibility immediately after a landslide event. Using the TomTom Traffic API "https://developer.tomtom.com/", the agent retrieves live traffic flow metrics, including current travel speeds, free-flow speed baselines, congestion levels, and road segment status (open, restricted, or blocked). These data allow the system to rapidly evaluate the operational feasibility of accessing affected areas. Because these mobility metrics fluctuate dynamically, this agent operates strictly in real-time or near real-time, supporting emergency planning and route prioritization rather than long-term monitoring. The outputs of the Traffic Agent are fed directly into the route-planning logic to adjust accessibility analysis according to current mobility conditions.
✓ **Route Agent:** The Route Agent is responsible for computing the fastest and most reliable access paths between emergency centers and detected landslide locations, incorporating both the static road network and the real-time mobility information provided by the Traffic Agent. In accordance with the Italian hierarchical disaster-response structure of Civil Protection National System, the starting points for route computation correspond to the municipality-level emergency coordination hubs, known as Centro Operativo Comunale (COC) translated as the Municipal Operations Center. Each COC location is represented by its geographic coordinates (latitude and longitude) and serves as the primary dispatch point for ground-based emergency response.

Given the COC coordinates and the landslide centroid, the Route Agent computes alternative access routes using APIs capable of integrating travel-time estimation with traffic-aware pathfinding. By considering road connectivity, and current congestion levels, the agent identifies the fastest-access route as well as backup routes in case of blocked roads. This supports timely decision-making during the immediate post-disaster window, ensuring that responders can quickly assess accessibility constraints and determine the optimal intervention trajectory.

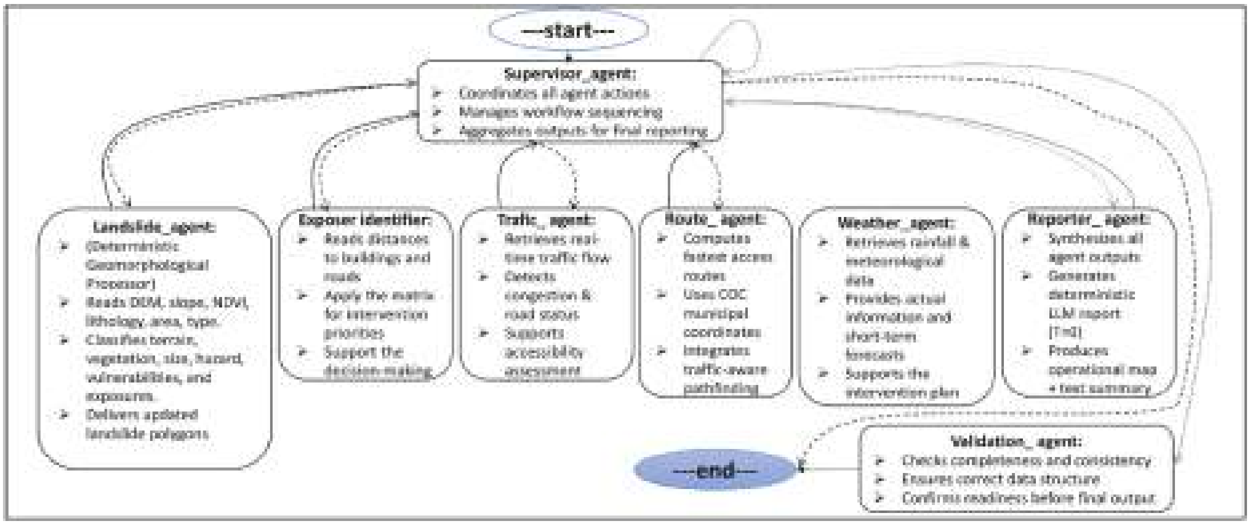


Fig. 4. Supervisor-driven workflow of the multi-agent system, where the Supervisor coordinates and reacts to exposure updates, orchestrates all agents, and the Validator ensures completeness before generating the final operational report.

3.3.4. *Exposure identifier and intervention priority*

The Exposure Identifier agent provides a structured spatial characterization of each detected landslide by analysing its relationship with exposed elements in the surrounding environment. Using the segmented landslide polygons as input, the agent evaluates spatial proximity to critical infrastructure, specifically building footprints extracted through the RT-DETR model and road networks obtained from OpenStreetMap (OSM). These proximity-based indicators allow the system to determine whether a detected landslide potentially affects structures, transportation corridors, or access routes relevant for civil protection operations.

Within the framework, the role of this agent is purely descriptive and analytical. It does not estimate risk or economic loss, nor does it determine emergency intervention actions. Instead, it provides objective exposure information that supports downstream decision-making and reporting. Intervention relevance is assessed exclusively through deterministic spatial relationships such as distance to buildings and roads derived from geospatial analysis of post-event data.

Because no detailed or standardized technical rules currently exist for post-landslide intervention prioritization, this study proposes a set of transparent and reproducible decision guidelines, summarized in Table 1. These guidelines classify intervention priority based on exposure to buildings and roads, local slope conditions, and the potential influence of debris on infrastructure. They are intended as a practical support tool for organizing field inspections and emergency response activities, while remaining flexible and readily adaptable should official technical standards or more detailed intervention protocols become available in the future. Within the multi-agent system, the Exposure Identifier therefore serves as the central component for linking landslide detection to infrastructure impact assessment and intervention prioritization.

3.3.5. *Reporter Agent (LLM-based synthesis)*

The Reporter Agent integrates and synthesizes the outputs generated by all upstream agents to form a comprehensive and operationally useable summary for Civil Protection authorities. Drawing on the, segmentation outputs, and accessibility analyses, the Reporter constructs a structured narrative describes the characteristics of each identified landslide, and contextualizes these observations with the most recent rainfall and meteorological conditions.

In addition, it incorporates information from the Traffic and Route Agents to highlight relevant accessibility constraints that may affect field presence or emergency movements. The synthesis is performed using Llama-3.3-70B, operated with a temperature setting of $T = 0$, to minimize stochastic variability during text generation and reduce the likelihood of hallucinated or inconsistent statement.

Determinism in MAS-LAND is achieved architecturally rather than by relying solely on the LLM temperature setting. All safety-critical outputs are computed before the Reporter Agent is invoked. The Landslide Agent deterministically extracts geospatial and terrain attributes, while the Exposure/Priority Agent assigns intervention priorities and recommended actions using a fixed rule-based decision matrix. The Reporter Agent is restricted to transforming these structured outputs into readable operational summaries and does not modify authoritative fields such as coordinates, distances, priority levels, deployment requirements, or recommended actions.

As a result, identical static inputs produce identical structured operational outputs, while minor wording variations may still occur in the LLM-generated narrative if the provider updates the underlying model. Similarly, external operational agents relying on evolving third-party services (e.g., weather or routing APIs) may produce slightly different contextual information over time as underlying data sources and service versions are updated. This design preserves operational reliability, traceability, and auditability for safety-critical disaster-response applications.

Table 2
Binary landslide segmentation performance (%).

Model	Backbone	Precision	Recall	f1	Iou	Miou
Unet++	**(efficientnet-b2, 512)**	67.85	61.29	64.41	47.5	70.91
	(efficientnet-b3, 640)	72.41	67.21	68.84	51.56	73.24
SegFormer	**MIT_b0**	64.66	72.31	68.27	51.83	73.07
	MIT_b1	64.7	75.01	67.12	50.2	72.74
	MIT_b2	67.94	74.81	**71.14**	**55.31**	**75.64**
	MIT_b3	72.15	71.5	70.1	55.41	75.24
	MIT_b4	62.97	**77.04**	70.2	53.18	73.36
	MIT_b5	**73.84**	68.94	70.14	54.27	75.15

The Reporter Agent operates using a constrained system prompt designed to generate standardized Civil Protection operational summaries exclusively from structured upstream outputs. To improve reproducibility and clarify the prompt-engineering strategy, the Reporter Agent prompt template and deterministic routing logic are provided in Appendix A.

3.3.6. *Validation Agent (rule-based consistency checking)*

A deterministic validation module confirms the presence and structure of key report sections, checks completeness, and ensures formatting consistency.

This rule-based approach is intentionally chosen over LLM-based validation to avoid hallucinations and ensure operational reliability.

Taken together, the individual agents operate within a coordinated, supervisor-driven architecture that ensures coherent execution of all post-event detection, exposure analysis, reporting tasks, and interactive map generation. As illustrated in Fig. 4, the Supervisor Agent orchestrates the workflow by dispatching instructions to the environmental agents (Weather, Traffic, and Route), the analytical agents (Landslide Agent and Exposure Identifier), and the LLM-based Reporter Agent. Following the availability of post-event imagery, the Supervisor activates the Landslide Agent to perform segmentation-based detection and the Exposure Identifier to evaluate spatial relationships between detected landslides and nearby buildings and road infrastructure. In parallel, the Traffic and Route Agents assess real-time accessibility conditions, providing essential mobility and access constraints relevant for field operations. Once all agents have completed their respective tasks, the Supervisor aggregates their outputs and forwards them to the Reporter Agent, which synthesizes a structured operational report and corresponding interactive map outputs. A final Validation Agent ensures completeness, internal consistency, and coherence of the generated products before the workflow terminates.

Within the implemented architecture, the Supervisor routes execution through predefined operational agents using deterministic state-aware workflow control rather than unconstrained autonomous LLM reasoning. This agent-centric interaction pattern guarantees that each component contributes its specialized function while maintaining full coordination under a unified supervisory logic, enabling traceable, reproducible, and operationally focused decision support across the entire post-event landslide management pipeline.

The framework was implemented in Python (v3.10+) using a modular multi-agent architecture orchestrated with LangGraph, while landslide segmentation relied on SegFormer models fine-tuned in PyTorch via the HuggingFace ecosystem. Natural-language reporting was powered by the Groq Llama-3.3-70B-versatile model. Deterministic computations for exposure was carried out using NumPy and Pandas and GeoPandas. PDF operational forms and interactive geospatial visualizations were generated using ReportLab and Folium, respectively.

4. Results

This section presents the outcomes obtained at each stage of the proposed framework, following the operational sequence of landslide detection, infrastructure exposure analysis, and multi-agent coordination. Results are reported to evaluate the performance of individual components as well as the consistency and effectiveness of the integrated workflow when applied to the May 2023 Emilia–Romagna landslide crisis. Each subsection introduces the objective of the corresponding stage before discussing its quantitative and qualitative outputs.

4.1. *Landslide segmentation*

The first stage of the framework focuses on post-event landslide detection and multi-class delineation using the 0.20 m RGB orthophotos acquired following the May 2023 events. All models were trained on the manually mapped inventories, and performance was evaluated on an independent test set containing 806 ground-truth landslides. Two evaluation configurations were explored: (i) binary segmentation, where all mapped landslides were treated as a single class, and (ii) multi-class segmentation, introduced here as no existing studies have considered this aspect, in which the three consolidated movement types (Rock slides, Debris movements, and Earth movements) were mapped separately.

Table 3
Multiclass landslide segmentation performance (%).

Model	Class	Pre	Rec	f1	Iou	OA	Pre macro	Rec macro	f1 macro	Pre micro	Rec micro	f1 micro	Miou
SegFormer MIT_b2	Rock slide	78.41	60.5	68.12	51.42	87.41	71.02	71.14	71.09	87.7	87.42	87.4	57.2
	Debris movements	91.71	94.2	92.71	86.5								
	Earth movements	45.91	62.9	52.04	35.71								

4.1.1. *Binary segmentation results*

Binary detection performance is summarized in Table 2. Transformer-based architectures outperformed convolutional U-Net++ variants across all metrics. SegFormer MIT-b2 achieved the best overall detection performance, with Precision = 67.94%, Recall = 74.81%, F1 = 71.14%, IoU = 55.31%, and mIoU = 75.64%. Instance-level detection was derived by applying connected component analysis (8-connectivity) to binary masks (landslide vs background), where each connected component was treated as an individual landslide and counted as detected if it overlapped with a ground-truth instance. SegFormer MIT-b2 successfully detected 603 of the 806 landslides in the test dataset. This represents the highest detection rate among all evaluated models. While some small or low-contrast failures were missed, especially in densely vegetated terrain, the model consistently captured medium- and large-scale events with high completeness.

Although U-Net++ with the EfficientNet-b3 backbone achieved a reasonable F1-score of 68.84%, it was clearly outperformed by all mid- and large-scale SegFormer variants (MIT-b2, MIT-b3, MIT-b4, MIT-b5), which consistently exceeded the 70% F1 threshold. Even the lighter MIT-b0 variant (F1 = 68.27%) performed comparably to U-Net++ while using substantially fewer parameters. These results confirm that transformer-based encoders deliver superior performance overall, providing more complete landslide outlines, improved spatial coherence, and enhanced discrimination of complex runout geometries compared with CNN-based U-Net++ models.

4.1.2. *Multi-class segmentation results*

Given its superior performance in the binary segmentation stage, SegFormer MIT-b2 was selected as the reference architecture for the subsequent three-class segmentation task (

Table 3). After identifying this model as the most effective detector of landslides, we extended the evaluation to the multi-class setting using the same backbone. The distribution of landslides in the test dataset is highly imbalanced Rock slides: 4, Earth movements: 161, and Debris movements: 641, a factor that strongly influenced class-specific performance and must be considered when interpreting the results.

Among the three movement types, debris movements the most numerous and morphologically distinctive class achieved the strongest results, with an F1-score of 91.71% and an IoU of 86.5%, reflecting the model's ability to capture their clear spectral textures and well-defined geometric patterns. Rock slides, although represented by only four instances in the test dataset, reached an F1-score of 68.12% and IoU of 51.42%; while the limited sample size reduces statistical confidence, the model still demonstrated a reasonable ability to discriminate rock-slide features despite extreme class imbalance. In contrast, earth movements (161 instances) recorded lower performance (F1 = 52.04%, IoU = 35.71%), consistent with their diffuse boundaries, shallow deformation morphology, and weak radiometric contrast in RGB orthophotos. These challenges are illustrated in Fig. 5, where the last row sample presents representative earth movement cases that are either partially delineated or misclassified as debris movements due to their similar spectral and morphological characteristics. Despite these variations, the overall multi-class performance remains robust, with a macro F1 of 71.09%, micro F1 of 87.40%, and overall accuracy (OA) of 87.41%, indicating that the selected model provides reliable operational mapping capacity even under heavily imbalanced class distributions. A clear qualitative distinction emerges when visually comparing the segmentation outputs across models. U-Net++ variants tend to generate false positives in vegetated or shadowed regions and frequently miss parts of the true landslide outlines, particularly along narrow debris tracks or diffuse earth-movement boundaries. Their predictions generally appear fragmented and less faithful to the reference inventory. In contrast, SegFormer MIT-b2 consistently produces the most coherent and accurate delineations, capturing fine-scale runout morphology, narrow scarps, and subtle slope breaks with markedly fewer misclassifications. The model also avoids the over-segmentation observed in CNN-based architectures, reinforcing the quantitative advantages demonstrated in Fig. 5.

4.2. *Multi-agentic system for priority assessment and real-time response*

Following a triggering event such as intense rainfall or an earthquake, the multi-agent system proceeds directly to the impact-assessment stage by analysing newly detected landslides using the SegFormer MIT-b2 model applied to post-event orthophotos. The pretrained RT-DETR roof-detection model is used to extract building footprints, while the corresponding road network is retrieved from OpenStreetMap (OSM). These infrastructure layers are spatially intersected with the segmented landslide polygons to evaluate whether detected failures fall within the critical proximity thresholds defined in the proposed intervention matrix (Table 1). Based on these deterministic spatial relationships specifically distance to buildings, distance to roads, and associated slope movement conditions the system assigns each landslide an intervention priority level. The resulting outputs are visualised through an interactive map viewer

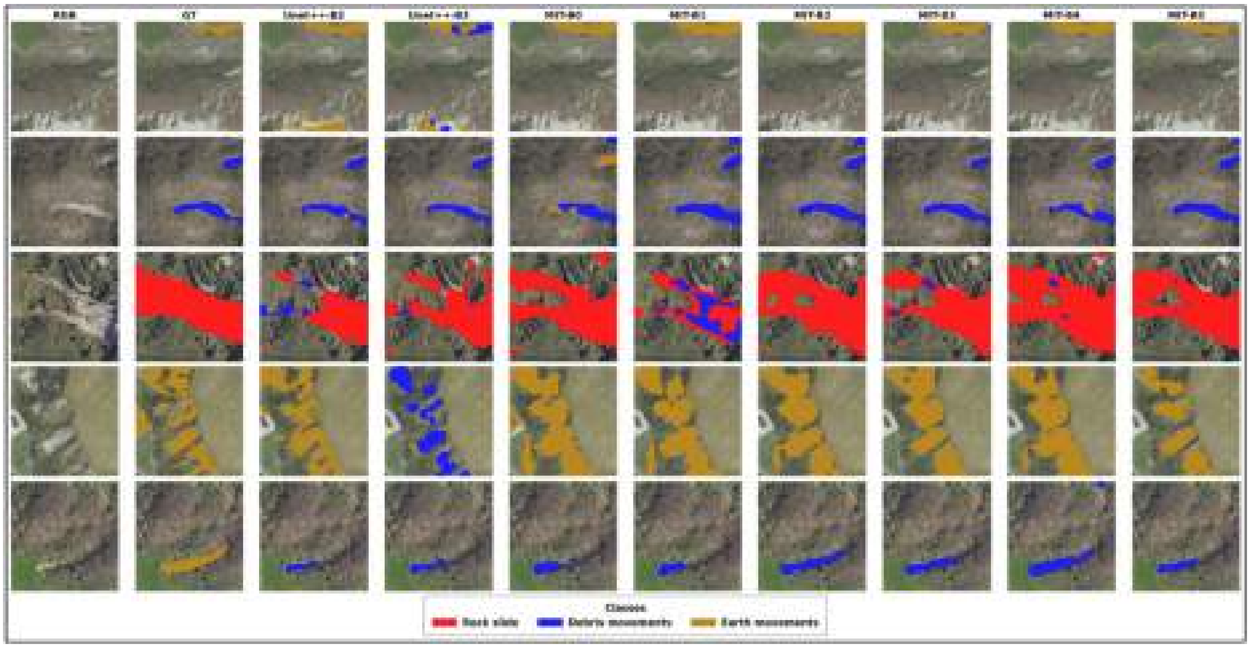


Fig. 5. Visual comparison with ground-truth (GT) annotations showing the superior delineation performance of SegFormer MIT-b2 relative to U-Net++ and other SegFormer variants.

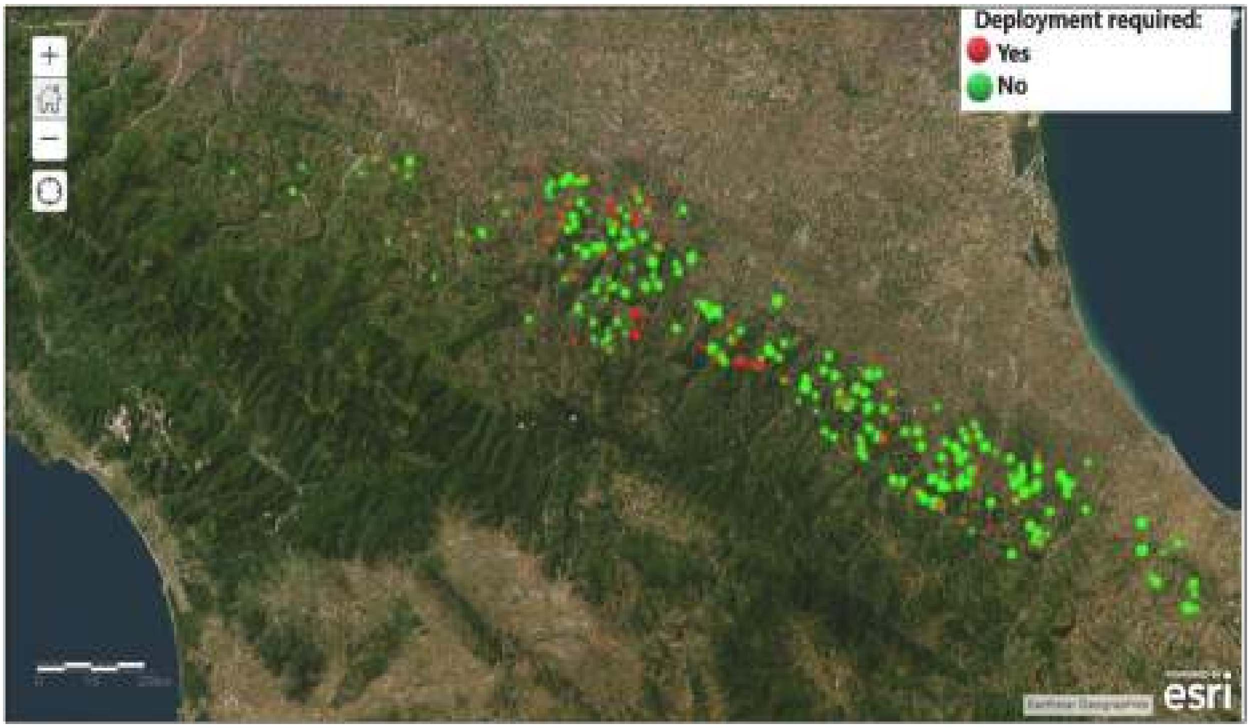


Fig. 6. Interactive map viewer generated by the multi-agent system, displaying detected landslides classified by intervention requirement (Yes/No) based on building and road proximity and decisionsupport criteria.

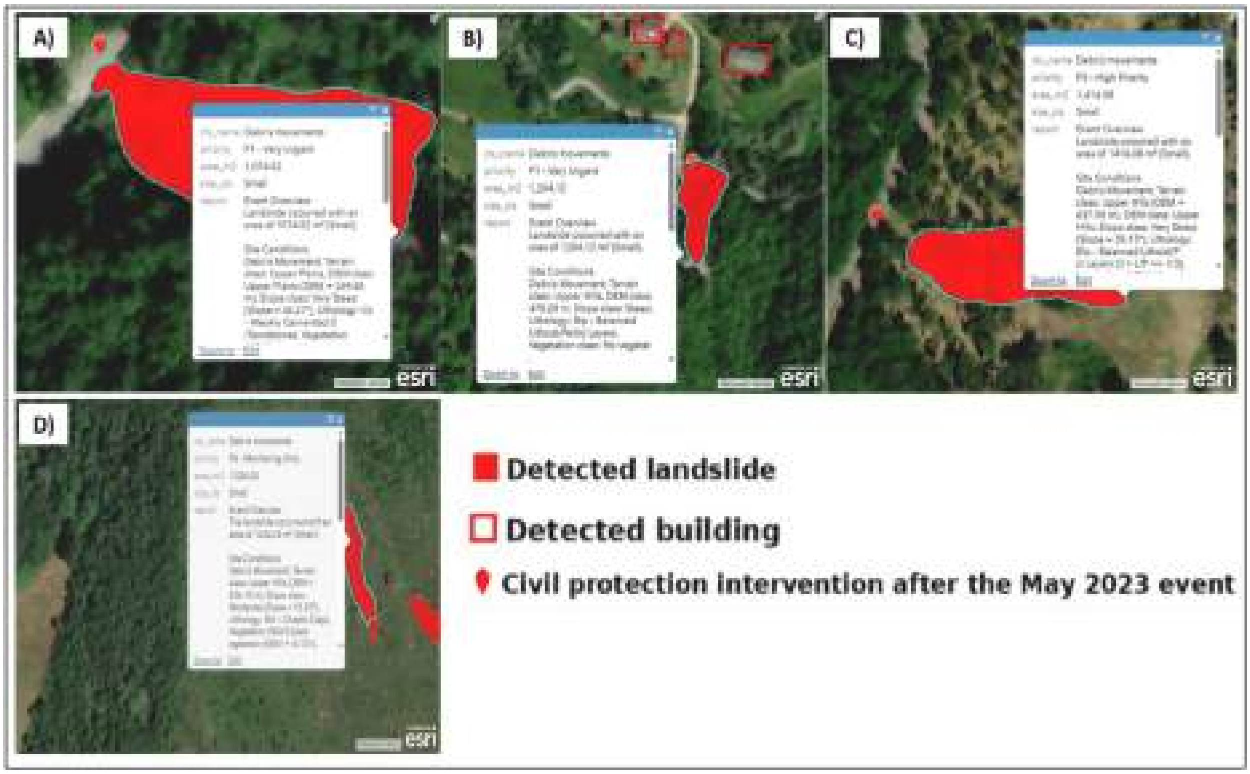


Fig. 7. Representative examples of landslide priority classes assigned by the proposed framework from documented post-event cases of the May 2023 event: A–B) P1 – Very Urgent, C) P3 – High Priority, and D) P6 – Monitoring Only.

(Fig. 6), where landslides requiring deployment are highlighted in red and those not requiring action are shown as non-critical in green.

Applying the full workflow to the test dataset resulted in 973 classified landslide cases, of which 113 required intervention and 860 did not. The distribution across priority levels was as follows: 63 landslides classified as P1 – Very Urgent, 9 as P2 – Urgent, 41 as P3 – High Priority, 87 as P4 – Medium Priority, 12 as P5 – Low Priority, and 761 as P6 – No Immediate Intervention Required.

This interface allows emergency operators to rapidly inspect the geographic distribution of priority events and understand, at a glance, which locations may involve structural damage or road blockage. By converting raw geospatial detections into operationally meaningful intervention indicators, the multi-agent architecture delivers a coherent, real-time decision-support layer that directly supports field coordination and strategic emergency management. The final output is provided through an online interactive web map, which enables dynamic exploration of detected landslides, intervention priorities, and exposed infrastructure and is accessible,at: (https://www.arcgis.com/home/item.html?id=edf943a9dd634a629b2758e208a46cef).

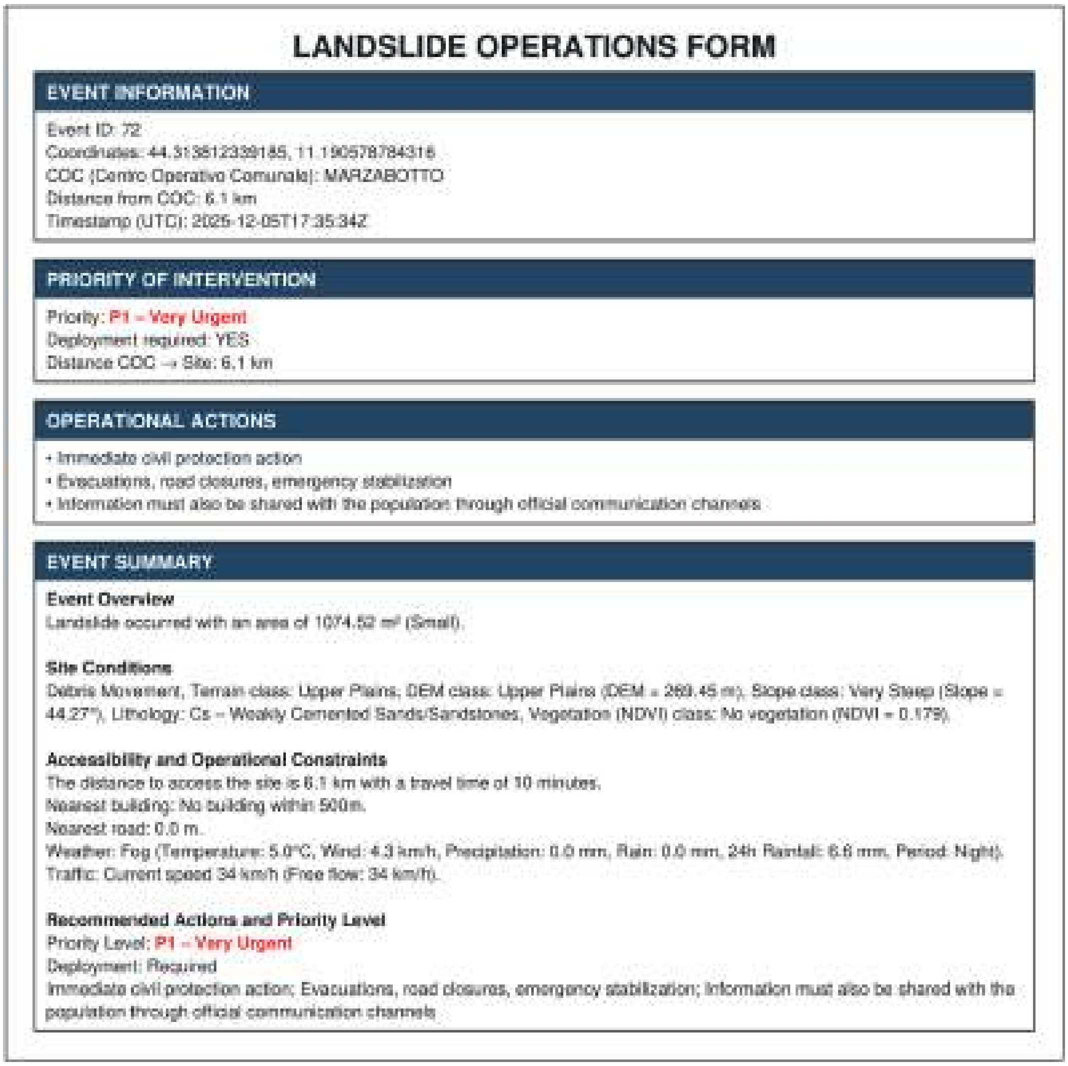
A representative P1 – Very Urgent case is shown in Fig. 7 panel **A)**, where the detected landslide is located in close proximity to exposed infrastructure. In this situation, the system identifies a potentially critical threat condition and generates a report supporting immediate intervention.

A second example of a P1 – Very Urgent classification is presented in panel **B)**, where the detected landslide is associated with nearby buildings. The spatial relationship between the mapped failure and the exposed elements leads the framework to classify the event among the most critical cases requiring rapid operational response.

Panel **C)** illustrates a P3 – High Priority case, in which the detected landslide affects exposed elements but with a lower degree of urgency than in the previous examples. Although intervention is still recommended, the generated report indicates a less critical priority level.

In contrast, panel **D)** presents a representative non-intervention case, where the detected landslide is not spatially associated with buildings or critical infrastructure. Under these conditions, the system assigns a P6 – Monitoring Only priority, indicating that no immediate field deployment is required and that remote monitoring is sufficient at the current stage.

For validation, the intervention priorities generated for the cases reported in panels **A)–C)** were compared with documented emergency responses carried out by the Emilia–Romagna Civil Protection after the May 2023 event, as reported in the official post-disaster records. The correspondence between the critical cases identified by the system and the locations where emergency interventions were actually conducted supports the practical reliability of the proposed approach. Conversely, the non-critical case shown in panel **D)** is consistent with a situation in which no operational deployment was required, further confirming the ability of the



LANDSLIDE OPERATIONS FORM

EVENT INFORMATION

Event ID: 72
Coordinates: 44.313812339185, 11.190578784316
COC (Centro Operativo Comunale): MARZABOTTO
Distance from COC: 6.1 km
Timestamp (UTC): 2025-12-05T17:35:34Z

PRIORITY OF INTERVENTION

Priority: P1 – Very Urgent
Deployment required: YES
Distance COC → Site: 6.1 km

OPERATIONAL ACTIONS

- Immediate civil protection action
- Evacuations, road closures, emergency stabilization
- Information must also be shared with the population through official communication channels

EVENT SUMMARY

Event Overview
Landslide occurred with an area of 1074.52 m² (Small).

Site Conditions
Debris Movement, Terrain class: Upper Plains, DEM class: Upper Plains (DEM = 269.45 m), Slope class: Very Steep (Slope = 44.27°), Lithology: Cs – Weakly Cemented Sands/Sandstones, Vegetation (NDVI) class: No vegetation (NDVI = 0.179).

Accessibility and Operational Constraints
The distance to access the site is 6.1 km with a travel time of 10 minutes.
Nearest building: No building within 500m.
Nearest road: 0.0 m.
Weather: Fog (Temperature: 5.0°C, Wind: 4.3 km/h, Precipitation: 0.0 mm, Rain: 0.0 mm, 24h Rainfall: 6.6 mm, Period: Night).
Traffic: Current speed 34 km/h (Free flow: 34 km/h).

Recommended Actions and Priority Level
Priority Level: P1 – Very Urgent
Deployment: Required
Immediate civil protection action; Evacuations, road closures, emergency stabilization; Information must also be shared with the population through official communication channels

Fig. 8. Example of the automated landslide operations form generated by the multi-agent system.

framework to distinguish hazardous from non-critical events. Overall, these results demonstrate the potential of the system to support rapid, near-real-time prioritization during future rainfall-induced landslide emergencies.

In addition to the spatial prioritization results, the framework automatically generates a structured operational report for each detected landslide, as illustrated in Fig. 8. Each report is organized into four main sections: Event Information, which summarizes the identification and location of the event; Priority of Intervention, which records the assigned urgency level and deployment requirement; Operational Actions, which lists the recommended response measures; and Event Summary, which provides a concise synthesis of the main landslide characteristics, site conditions, accessibility constraints, and suggested actions. This standardized format translates the analytical output of the framework into an operational product that can be directly interpreted by civil protection actors and emergency managers.

The complete MAS-LAND workflow was evaluated on a workstation equipped with an NVIDIA GeForce RTX 3090 GPU (36 GB VRAM) using the Emilia–Romagna case study. SegFormer inference over the complete post-event test dataset required approximately 56 s, while the RT-DETR building-detection stage required approximately 206 s. Environmental API queries (Open-Meteo and TomTom Routing/Traffic APIs) introduced only minor latency per landslide instance, whereas the remaining analytical agents consisted primarily of lightweight deterministic programmatic operations. Including operational report synthesis by the Reporter Agent, the full workflow required approximately 5–8 min for the complete scenario containing 973 detected landslide instances, depending mainly on external API latency and LLM inference speed. These results demonstrate the operational feasibility of the proposed framework for near-real-time post-event landslide assessment and reporting.

To ensure traceability and future reuse, these reports are saved individually as PDF files for every detected landslide. In this way, the system not only supports immediate emergency-response interpretation, but also progressively builds a structured documentary database of landslide events. Such a database can serve as the foundation for a retrieval-augmented generation (RAG) system in future

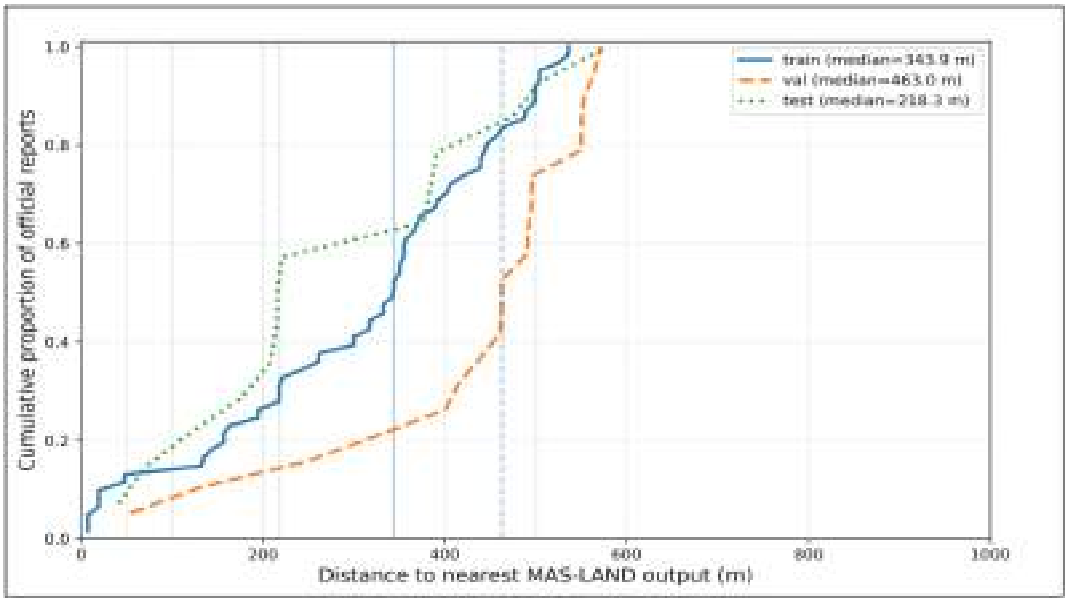


Fig. 9. Spatial agreement between official landslide reports and MAS-LAND output.

developments, enabling a chatbot to access past reports, retrieve relevant event-specific information, and support interactive querying for operational analysis and decision support. standardized format translates the analytical output of the framework into an operational product that can be directly interpreted by civil protection actors and emergency managers.

4.3. System-level validation

Beyond conventional segmentation metrics, an additional system-level validation was conducted to evaluate the operational relevance of MAS-LAND outputs using independently collected emergency landslide reports from the Emilia–Romagna Regional Geoportal. The official dataset corresponds to the first regional mapping of hydrogeological instability reports associated with the May 2023 Emilia–Romagna event and was published on 01/06/2023 by the Regional Civil Protection authorities through the Emilia–Romagna Geoportal (https://geoportale.regione.emilia-romagna.it/catalogo/dati-cartografici/informazioni-geoscientifiche/zone-a-rischio-naturale/layer).

According to the regional authority, the dataset represents a preliminary mapping of hydrogeological instability reports collected during the emergency period and does not constitute a complete landslide inventory. Each point corresponds either to a territorial signalization received from local stakeholders or to rapid field observations performed by geologists of the Regional Civil Protection Agency. Consequently, the published locations represent approximate signalization points rather than precise landslide polygon boundaries. This characteristic introduces an inherent positional uncertainty that must be considered during validation.

To account for this limitation, the validation was performed using spatial proximity analysis rather than strict polygon overlap. First, all official signalization points were spatially associated with the training, validation, and independent test subsets through buffered patch-footprint intersections. Subsequently, nearest-distance analysis was performed between each official report and the closest MAS-LAND detected landslide output. In addition, a buffered-point analysis was conducted using a representative buffer derived from the average mapped landslide-object size estimated from the segmentation masks. This approach allowed the evaluation to incorporate the uncertainty associated with emergency signalization coordinates. The resulting spatial agreement analysis is presented in Fig. 9. The cumulative distribution curves show that MAS-LAND outputs maintain strong spatial consistency with independently collected emergency reports across all subsets. The independent test subset achieved the strongest agreement, with a median nearest distance of approximately 218 m between official reports and MAS-LAND outputs. Furthermore, 85.7% of the official reports located in the independent test subset were positioned within 500 m of a detected MAS-LAND landslide output. The training subset showed a median distance of approximately 344 m, whereas the validation subset exhibited a larger median distance of approximately 463 m, mainly due to the lower number of available official reports in that subset.

Importantly, the official signalization dataset became publicly available approximately 13 days after the last major rainfall-triggered event of May 2023, reflecting the time required for traditional territorial reporting and rapid field-survey compilation. In contrast, once post-event satellite imagery becomes available, MAS-LAND is capable of generating the complete operational workflow including landslide detection, exposure analysis, intervention-priority assessment, and automatic report generation in only a few minutes for the entire study scenario. This comparison demonstrates the potential of the proposed framework to substantially accelerate the first-stage situational assessment process during large-scale hydrogeological emergencies. Overall, these results provide quantitative evidence that MAS-LAND is not limited to generating visually plausible outputs, but is capable of identifying operationally relevant instability zones that spatially agree with independently collected emergency observations.

5. Discussion

Recent advances in disaster-oriented artificial intelligence increasingly emphasize multi-agent architectures to manage the complexity of real-world emergency scenarios, where multiple data streams, analytical tasks, and operational constraints must be handled simultaneously. Prior studies have demonstrated the potential of agent-based systems to improve coordination and interpretation in disaster contexts, including flood early warning and response [12], autonomous multi-step scene understanding from remote sensing imagery [13], and decision-support platforms integrating heterogeneous information systems under infrastructural and socioeconomic constraints [14]. More recent frameworks such as Disaster Copilot [15] and MARSHA [16] further illustrate a shift toward language-enhanced, orchestrated agent ecosystems capable of supporting complex disaster-management workflows. Within the landslide domain, CC-GRMAS [17] represents an important step toward agent-based coordination yet remains primarily focused on forecasting and situational awareness rather than post-event operational response.

Building on this emerging body of work, the framework presented in this study advances the state of the art by explicitly targeting the early post-event phase, in which civil protection authorities require rapid, interpretable, and actionable information derived from newly acquired imagery. By integrating automated landslide detection, infrastructure exposure analysis, and LLM-based synthesis within a unified multi-agent architecture, the proposed workflow transforms raw geospatial outputs into operationally meaningful decision-support products. Within the experimental setting adopted here, SegFormer MIT-b2 provided the most stable overall performance, yielding a favorable balance between detection accuracy and boundary delineation under strong class imbalance and heterogeneous terrain conditions. However, this result should be interpreted with caution and not as evidence of universal superiority. The suitability of a segmentation model depends on several factors, including image spatial resolution, landslide size and typology, landscape complexity, vegetation cover, and the geomorphological characteristics of the study area. In particular, the present assessment was conducted using very high resolution orthophotos (20 cm), which offer a level of spatial detail rarely available in routine large-scale emergency monitoring. Under more commonly accessible satellite data, such as Sentinel-2 or PlanetScope imagery, model behavior may differ substantially, and alternative architectures or tailored training strategies may prove more appropriate. Therefore, the selection of the segmentation backbone should be regarded as context-dependent, with SegFormer MIT-b2 representing the most suitable solution for the specific data and conditions examined in this study. This validation justified its adoption as the reference backbone within MAS-LAND. To further investigate the robustness and transferability of the proposed framework beyond the original Emilia–Romagna orthophoto setting, additional experiments were conducted under different spatial-resolution and geographic conditions. Appendix B presents a direct inference experiment on the Lushan landslide dataset using WorldView imagery at 0.5 m spatial resolution, where the same pretrained SegFormer and RT-DETR models were applied without additional fine-tuning. This experiment was designed to evaluate the transferability of MAS-LAND to an unseen geographic region and independent event scenario. In addition, Appendix C introduces a lower-resolution operational “*Plan B*″ configuration of MAS-LAND based on PlanetScope imagery (3 m spatial resolution) over the Emilia–Romagna event. In this scenario, the segmentation backbone was fine-tuned specifically for PlanetScope imagery, while the RT-DETR building-detection stage was replaced by OpenStreetMap (OSM)-derived building footprints to simulate conditions where very-high-resolution imagery and automated building extraction are not available. These complementary experiments were included to reduce potential biases associated with a single-event, single-resolution evaluation setting and to assess the adaptability of MAS-LAND under heterogeneous acquisition conditions. In lower-resolution scenarios where high-resolution orthophotos are not available, alternative data sources must be considered for infrastructure extraction. One practical option is the use of OpenStreetMap (OSM) building data in place of automated detection models such as RT-DETR. While this substitution enables continued operation of the framework, it may introduce performance limitations due to positional inaccuracies, incomplete building coverage, and temporal inconsistencies inherent to crowd-sourced data. As a result, the accuracy of exposure assessment and subsequent prioritization may be reduced, particularly in rural or rapidly evolving environments. These limitations should be considered when transferring the framework to coarser-resolution imagery, where both landslide detection and infrastructure characterization are inherently more uncertain. Nevertheless, the PlanetScope-based experiment demonstrated that the proposed multi-agent workflow can still produce operationally relevant prioritization outputs under reduced spatial detail, although with lower spatial agreement relative to the very-high-resolution orthophoto scenario.

Furthermore, because no standardized or officially codified technical procedure currently exists for post-landslide intervention prioritization based on automated geospatial analysis, this study introduces a transparent and reproducible priority-intervention matrix grounded in deterministic spatial proximity rules. The matrix translates distances to buildings and roads, together with slope movement characteristics, into clearly defined operational priority levels. While designed to align with practical civil protection reasoning, the proposed matrix remains adaptable and can be readily updated or replaced should formal institutional guidelines become available in the future. Agent coordination ensures that all analytical steps are executed through deterministic, transparent geospatial procedures, while language models are used exclusively for structured interpretation and reporting.

A key contribution of the framework is its ability to automatically generate a standardized “Landslide Operations Form” for each detected failure, formatted for immediate use by civil protection authorities. A major added value of the proposed architecture is its ability to produce structured, machine-readable information at every stage of the workflow. Landslides detection, environmental attributes, exposure metrics, and natural-language intervention summaries are stored systematically, forming a progressively expanding event-specific and region-specific knowledge base. In this respect, the framework is designed to be scalable and transferable, and its modular structure allows application to other geographic contexts and emergency scenarios beyond the May 2023 case study. The workflow can also be adapted to different image sources, including not only very high-resolution orthophotos but also more operationally accessible satellite imagery such as Sentinel-2 or PlanetScope, depending on data availability and the scale of analysis. This flexibility supports its potential implementation over larger areas, provided that adequate computational resources and data-processing infrastructures are available. This structured database provides the foundation for future integration of Retrieval-Augmented Generation (RAG) systems, in which a conversational agent could retrieve past events, summarize historical interventions, provide context-aware explanations, and assist civil protection personnel during training or real-time operations. However, such RAG capabilities are not part of the current implementation and are considered a prospective extension of the framework.

In addition, the reliability of the intervention-related outputs was assessed through comparison with documented emergency actions carried out by the Civil Protection after the May 2023 event, providing an initial validation not only of the detection stage but also of the operational relevance of the generated priority assignments and recommendations. Thus, beyond its operational function, the workflow establishes the foundations for a long-term, knowledge-driven conversational assistant tailored to landslide risk management.

Regarding the limitations, the primary constraint of the proposed framework lies in the inherent uncertainty of image-based landslide detection, which may result in a limited number of false negatives and false positives. In particular, a small subset of very small failures, low-contrast earth movements, or landslides partially obscured by vegetation, shadows, or acquisition artifacts may not be consistently identified in post-event orthophotos, leading to occasional missed detections. A quantitative analysis of false negatives confirms that missed detections are strongly associated with small-scale landslides. Specifically, 47.40% of false negatives correspond to areas smaller than 100 m^2, and 72.08% fall below 200 m^2, with a median size of 105.88 m^2 as shown in Fig. 10. These results indicate that the effective detection limit of the segmentation framework lies in the range of small landslides, while medium and large-scale events are consistently identified. Similarly, a few geomorphic features or surface disturbances may be erroneously classified as landslides. Such limitations are not unique to the May 2023 case study; instead, they reflect broader challenges inherent to optical image-based automated segmentation and may therefore also emerge when the framework is transferred to different geographic settings or event scenarios. An additional limitation concerns the reproducibility of MAS-LAND in contexts where very high-resolution imagery, such as 0.2 m orthophotos or imagery close to 1 m resolution, is not available. In many operational settings, the most accessible post-event data may come from satellite platforms such as PlanetScope (3 m) or Sentinel-2 (10 m). Although these datasets are highly valuable for rapid and low-cost response over large areas, their lower spatial resolution can reduce the accuracy of the detection stage, particularly for small landslides, narrow runout zones, and precise boundary delineation. This limitation also propagates to the exposure-analysis stage, where reduced image detail may weaken the identification of affected infrastructure, especially buildings. In such cases, one practical alternative would be to replace image-based roof detection with pre-existing building datasets, for example by directly extracting building footprints from OpenStreetMap, which may partly compensate for the lower performance of the RT-DETR model under reduced-resolution imagery. Accordingly, while MAS-LAND remains transferable to new events and regions, its performance is expected to depend on the spatial resolution and completeness of the available post-event data. While such inaccuracies are relatively limited and do not compromise the operational relevance of the system which is designed to support rapid prioritization rather than exhaustive inventory generation, they underline the potential value of integrating complementary data sources, such as multi-temporal imagery or radar observations, in future developments. Additionally, the workflow depends on the availability of timely post-event aerial or satellite imagery, a constraint that may vary across operational settings and should be considered when transferring the approach to different geographic or institutional contexts.

Future developments will focus on expanding both the real-time data streams and the reasoning capabilities integrated within the multi-agent environment. Incorporating social media posts, crowdsourced imagery, and citizen call reports could provide additional situational cues especially where cloud cover or acquisition delays affect remote sensing. Higher-frequency SAR data, UAV-based rapid mapping, or integration of hydrological sensor networks may further reinforce the post-trigger monitoring capacity. Enhancing the LLM agents to provide uncertainty-aware reasoning, cross-modal validation, and adaptive learning will support more sophisticated operational guidance. Ultimately, these extensions aim to evolve the current system into a comprehensive, multi-modal disaster intelligence framework capable of supporting not only emergency response, but also preparedness, training, and long-term resilience-building.

6. Conclusion

This study presented an integrated multi-agent, LLM-enhanced framework for post-event landslide detection and operational decision support, demonstrated using the May 2023 Emilia–Romagna disaster as a real-world test case. By combining post-event aerial imagery, deep-learning–based landslide segmentation, infrastructure exposure analysis, and structured automated reporting, the proposed workflow transforms heterogeneous geospatial inputs into coherent and actionable intelligence tailored for early post-disaster operations. At the core of the framework, the SegFormer MIT-b2 model proved highly effective for landslide detection and multi-class segmentation from RGB orthophotos, achieving robust performance despite strong class imbalance and complex terrain conditions. The integration of automated building detection using a pretrained RT-DETR model and road-network extraction from OpenStreetMap enabled a reproducible and objective assessment of infrastructure exposure. These spatial relationships were subsequently translated into operational intervention priorities through a transparent and rule-based decision-support matrix, aligning closely with civil protection practices. The Reporter Agent further synthesized all analytical outputs into standardized, deterministic operational reports designed for immediate field use, ensuring interpretability, traceability, and reproducibility.

A key strength of the proposed workflow lies in its ability to generate structured, machine-readable outputs at every stage from landslide polygons and exposure indicators to intervention priorities and operational reports. Over time, these outputs naturally

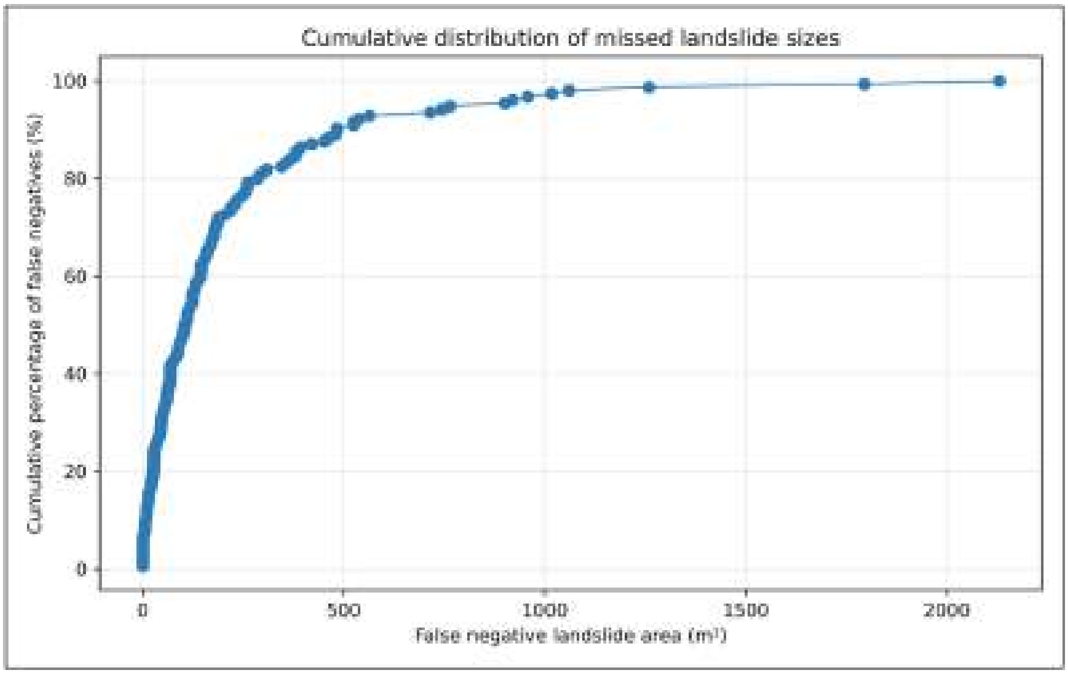


Fig. 10. Cumulative distribution of false-negative (FN) landslide areas.

accumulate into an event-based knowledge repository, providing a solid foundation for future Retrieval-Augmented Generation (RAG) systems and enabling the development of domain-specific conversational tools. In the current civil protection scenario, MAS-LAND can support operational activities by assisting personnel, improving preparedness through the organization of past event knowledge, and facilitating real-time decision-making during emergencies through rapid and standardized interpretation of newly detected landslides and their potential impacts on exposed elements. Based on the experimental evaluation, the end-to-end MAS-LAND pipeline processes input data and generates operational reports within approximately [5–8 min] depending on the size of the area and computational configuration, providing a quantifiable indication of its suitability for time-critical response scenarios. In this way, the framework contributes to faster prioritization, more consistent intervention assessment, and a more efficient allocation of emergency-response resources. While the framework demonstrates strong operational relevance, minor limitations remain, particularly related to occasional false negatives and false positives in image-based landslide detection, especially for very small or visually ambiguous failures. Future work will focus on strengthening detection robustness by integrating complementary data sources such as SAR imagery, UAV-based rapid mapping, social media content, and citizen reports, as well as extending the multi-agent architecture to support uncertainty-aware reasoning and cross-modal validation. Overall, the proposed system provides a robust, extensible, and transferable foundation for next-generation landslide disaster intelligence, with the potential to significantly improve the speed, consistency, and reliability of post-event assessment and intervention support across diverse operational contexts.

CRediT authorship contribution statement

Abdelkader Riche: Conceptualization, Data curation, Methodology, Software, Validation, Visualization, Writing – original draft. **Abdallah Zeggada:** Conceptualization, Data curation, Methodology, Software, Validation, Visualization, Writing – original draft. **Francesco Caleca:** Conceptualization, Methodology, Writing – review & editing. **Mohammed Alruqimi:** Data curation, Methodology, Software. **Pierluigi Confuorto:** Conceptualization, Data curation, Methodology, Supervision, Validation, Writing – review & editing. **Mawloud Guermoui:** Conceptualization, Methodology, Software, Writing – review & editing. **Silvia Bianchini:** Conceptualization, Data curation, Methodology, Supervision, Validation, Writing – review & editing. **Veronica Tofani:** Conceptualization, Methodology, Supervision, Validation, Writing – review & editing. **Farid Melgani:** Conceptualization, Methodology, Supervision, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Abdelkader Riche reports article publishing charges was provided by University of Florence Department of Earth Sciences. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This study was carried out within the Space It Up project funded by the Italian Space Agency, ASI, and the Ministry of University and Research, MUR, under contract n. 2024-5-E.0 – CUP n. I53D24000060005.

Appendix A

Appendix A provides the constrained prompt template used by the Reporter Agent within MAS-LAND. Unlike the other agents in the framework, which operate through deterministic geospatial processing or external API queries, the Reporter Agent uses a controlled LLM-based synthesis stage to transform structured operational outputs into standardized Civil Protection summaries. As illustrated in Figure A1, the prompt structure is intentionally constrained to minimize hallucinations, preserve authoritative geospatial and operational attributes, and ensure consistent formatting of the generated reports.

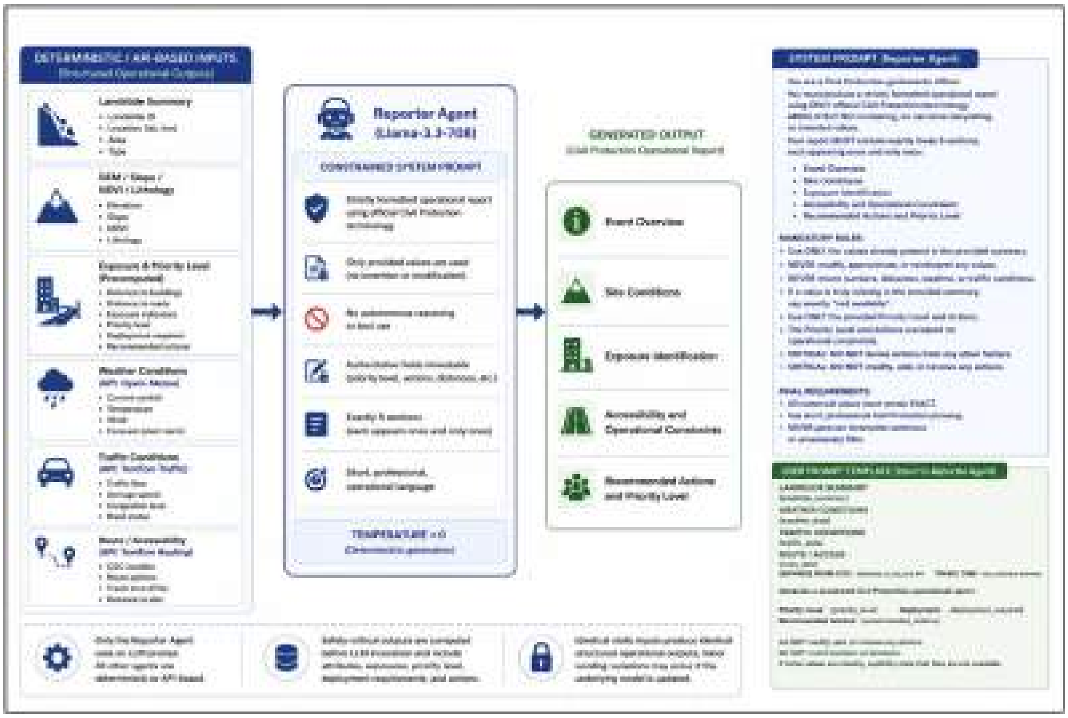


Fig. A 1. Constrained prompt architecture and prompt template of the MAS-LAND Reporter Agent for standardized Civil Protection operational report generation.

Appendix B. Transferability Experiment on the Lushan Landslide Dataset

To further evaluate the geographic transferability of the proposed framework, an additional experiment was conducted using the Lushan landslide subset of the GDCLD dataset [7], available at (https://zenodo.org/records/13612636). The dataset is based on 0.5 m WorldView imagery acquired after the Mw 5.9 Lushan earthquake that occurred in China on 1 June 2022. In this experiment, the framework was applied directly in an inference-only configuration without additional fine-tuning or retraining.

The quantitative results obtained on the Lushan dataset are summarized in Table B1, while representative examples of the generated MAS-LAND outputs are presented in Figure B1. The model achieved an IoU of 52.19%, F1-score of 65.20%, and an overall accuracy of 95.64%, indicating that the framework preserved a stable detection capability despite the geographic and geomorphological differences relative to the original training conditions. Notably, these results were obtained on an earthquake-triggered landslide event, whereas the original training scenario was based on heavy-rainfall-induced landslides, suggesting a reasonable degree of transferability across different triggering mechanisms.

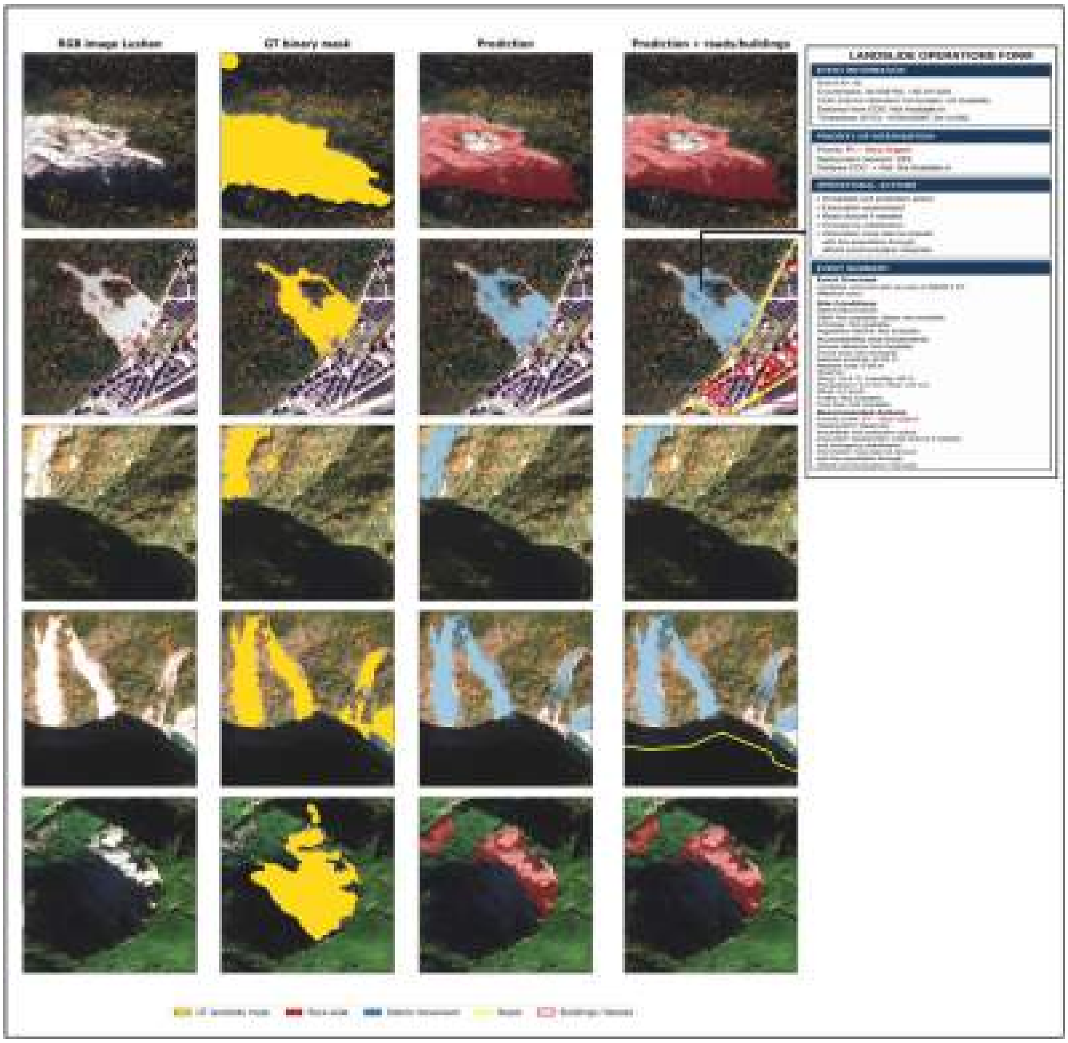


Fig. B 1. Representative MAS-LAND outputs generated on the Lushan GDCLD dataset under a direct inference.

Table B 1
Segmentation performance of the inference-only MAS-LAND configuration on the Lushan subset

Model	P	R	f1	Iou	Miou	OA
SegFormer MITb2	61.20	66.61	65.20	52.19	61.49	95.64

Representative examples further demonstrate that MAS-LAND successfully generated operational prioritization outputs integrating detected landslides, exposed infrastructure, and automated intervention reports. These results suggest that the proposed framework preserves a reasonable degree of transferability across geographically distinct post-event scenarios, although some reduction in segmentation precision and boundary agreement is expected without region-specific adaptation. In several generated operational forms, some environmental and operational fields were automatically reported as “Not Available”, particularly for DEM-derived layers (slope), NDVI, lithology information, and Civil Protection Operational Center (COC) accessibility indicators such as travel time and access conditions. This occurred because these auxiliary geospatial layers and routing services were not available for the Lushan study area. As defined in the operational prompt rules presented in Appendix A, the framework was designed to explicitly return the term “Not Available” whenever a required information source is missing, ensuring transparent reporting and avoiding the generation of unsupported or hallucinated operational information.

Appendix C. PlanetScope-Based Operational Scenario (3 m Resolution)

Table C 1
Quantitative segmentation performance of the SegFormer MIT-b2 model under the PlanetScope 3 m (%)

Model	Evaluation	Class	P	R	f1	Iou	Miou	OA	P macro	R macro	f1 macro	P micro	R micro	F1 micro
SegFormer MIT b2	**Binary**	(0, 1)	66.17	67.09	68.36	51.93	63.24	92.58	#	#	#	#	#	#
	Multiclass	Rock slide	49.22	26.29	31.02	10.39	30.49	80.27	69.03	41.58	48.92	82.32	46.90	61.74
		Debris movements	87.85	62.71	69.45	51.48								
		Earth movements	70.03	35.73	46.28	29.61								

To investigate the adaptability of MAS-LAND under lower-resolution operational conditions, an additional experimental scenario was conducted using PlanetScope imagery with a spatial resolution of 3 m over the Emilia–Romagna event area. In this configuration, the SegFormer MIT-b2 backbone was fine-tuned using multi-class PlanetScope landslide patches, while the infrastructure-exposure component of the workflow was adapted by replacing RT-DETR building extraction with OpenStreetMap (OSM)-derived building footprints. This experiment represents an operational “*Plan B*″ configuration of MAS-LAND intended for situations where very-high-resolution orthophotos and dedicated building-detection models are unavailable.

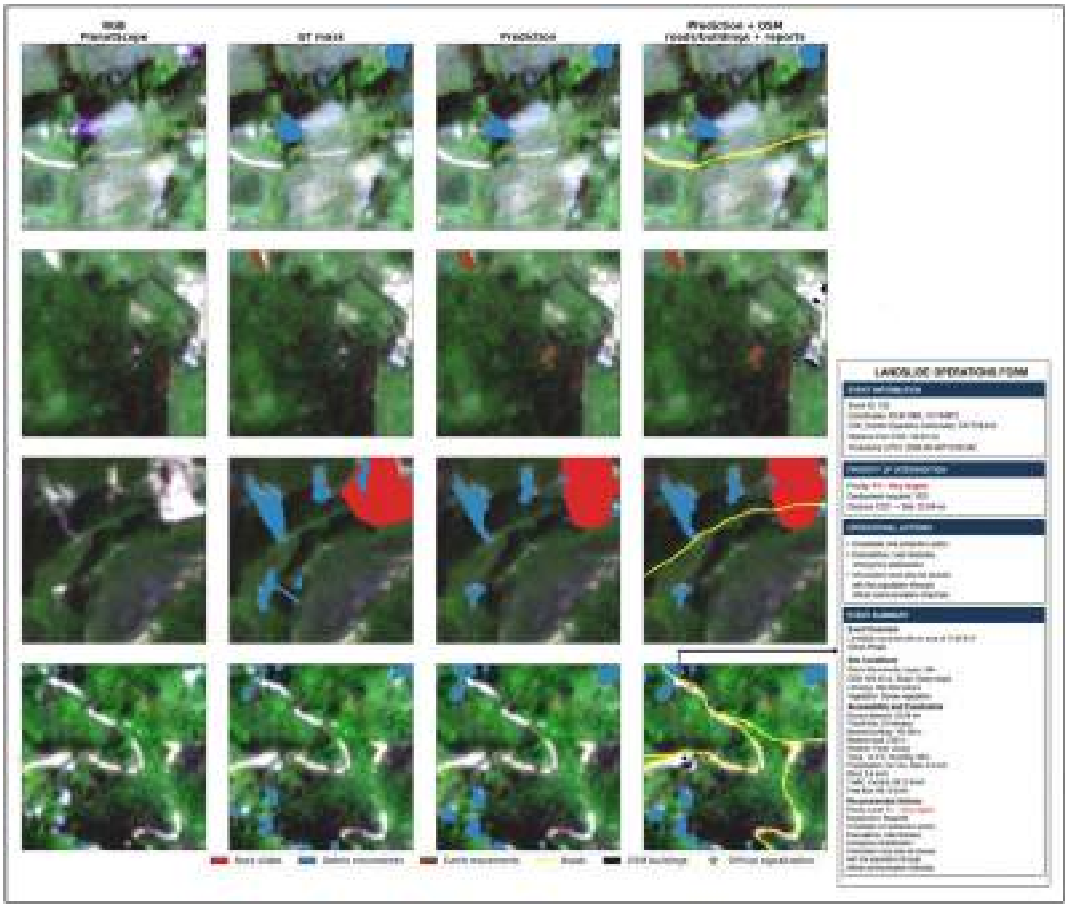


Fig. C 1. Spatial agreement between official intervention signalization points and MAS-LAND PlanetScope outputs.

The quantitative segmentation results obtained for the PlanetScope scenario are reported in Table C1. Despite the reduced spatial resolution, the model preserved the capability to detect the main landslide categories, particularly debris movements, while maintaining operationally meaningful prioritization outputs. Figure C1 presents representative examples of the multi-class segmentation results and the corresponding MAS-LAND operational outputs integrating roads, OSM buildings, official reports, and generated intervention priorities.

To further evaluate the operational consistency of the framework under lower-resolution conditions, the generated MAS-LAND outputs were spatially compared against the official intervention signalization points released by the Emilia–Romagna authorities. The validation was restricted to the PlanetScope test coverage area and a surrounding 500 m buffer to avoid spatial bias outside the observed region. The resulting spatial agreement analysis, presented in Figure C2, showed that all official reports located within the evaluated coverage area were associated with at least one MAS-LAND detection within 500 m, with a median distance of 86.3 m. These results suggest that the proposed framework retains operational relevance even under coarser-resolution imagery conditions, although with lower thematic and spatial detail compared to the original 20 cm orthophoto configuration.

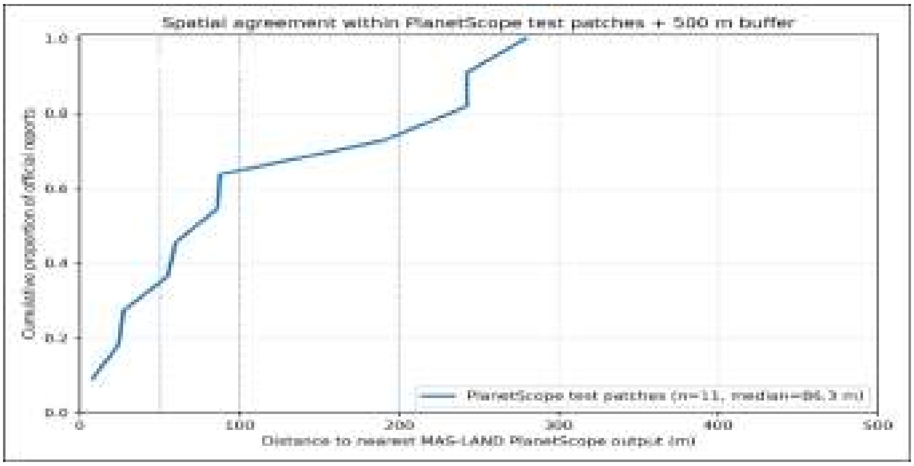


Fig. C 2. Representative PlanetScope-based MAS-LAND outputs and generated intervention priorities under the 3 m resolution operational scenario.

Data availability

Data will be made available on request.

References

[1] F. Ardizzone, M. Cardinali, A. Carrara, F. Guzzetti, P. Reichenbach, Impact of mapping errors on the reliability of landslide hazard maps, Nat. Hazards Earth Syst. Sci. 2 (2002) 3–14, https://doi.org/10.5194/nhess-2-3-2002.

[2] Q. Zhang, T. Wang, Deep learning for exploring landslides with remote sensing and Geo-Environmental data: Frameworks, progress, challenges, and opportunities, Remote Sens. 16 (2024) 1344, https://doi.org/10.3390/rs16081344.

[3] H. Wang, L. Zhang, K. Yin, H. Luo, J. Li, Landslide identification using machine learning, Geosci. Front. 12 (2021) 351–364, https://doi.org/10.1016/j.gsf.2020.02.012.

[4] G. Cheng, Z. Wang, C. Huang, Y. Yang, J. Hu, X. Yan, Y. Tan, L. Liao, X. Zhou, Y. Li, S. Hussain, M. Faisal, H. Li, Advances in deep learning recognition of landslides based on remote sensing images, Remote Sens. 16 (2024) 1787, https://doi.org/10.3390/rs16101787.

[5] L.P. Soares, H.C. Dias, G.P.B. Garcia, C.H. Grohmann, Landslide segmentation with deep learning: evaluating model generalization in rainfall-induced landslides in Brazil, Remote Sens. 14 (2022) 2237, https://doi.org/10.3390/rs14092237.

[6] Y. Xu, C. Ouyang, Q. Xu, D. Wang, B. Zhao, Y. Luo, CAS landslide dataset: a large-scale and multisensor dataset for deep learning-based landslide detection, Sci. Data 11 (2024) 12, https://doi.org/10.1038/s41597-023-02847-z.

[7] C. Fang, X. Fan, X. Wang, L. Nava, H. Zhong, X. Dong, J. Qi, F. Catani, A globally distributed dataset of coseismic landslide mapping via multi-source high-resolution remote sensing images, Earth Syst. Sci. Data 16 (2024) 4817–4842, https://doi.org/10.5194/essd-16-4817-2024.

[8] H.M. Syed, M.M. Oghaz, L.B. Saheer, Semantic Segmentation for Landslide Detection Using Segformer, Proceedings of the Artificial Intelligence XLI (December 2024) 17–19. Cambridge, UK.

[9] L. Xi, J. Yu, D. Ge, Y. Pang, P. Zhou, C. Hou, Y. Li, Y. Chen, Y. Dong, SAM-CFFNet: Sam-based cross-feature fusion network for intelligent identification of landslides, Remote Sens. 16 (2024) 2334, https://doi.org/10.3390/rs16132334.

[10] Wenqi Cui, X. He, M. Yao, Z. Wang, J. Li, Y. Hao, W. Wu, H. Zhao, X. Chen, Wei Cui, Landslide image captioning method based on semantic Gate and Bi-Temporal LSTM, ISPRS Int. J. GeoInf. 9 (2020) 194, https://doi.org/10.3390/ijgi9040194.

[11] K. Areerob, V. Nguyen, X. Li, S. Inadomi, T. Shimada, H. Kanasaki, Z. Wang, M. Suganuma, K. Nagatani, P. Chun, T. Okatani, Multimodal artificial intelligence approaches using large language models for expert-level landslide image analysis, Comput. Aided Civ. Infrastruct. Eng. 40 (2025) 2900–2921, https://doi.org/10.1111/mice.13482.

[12] N.A.M. Rashid, Z.Z. Abidin, Z.A. Abas, Integrating multi-agent System and case-based reasoning for flood early warning and response system, Int. J. Adv. Comput. Sci. Appl. 15 (2024), https://doi.org/10.14569/IJACSA.2024.0151250.

[13] Z. Liu, D. Zhao, B. Yuan, RescueADI: adaptive disaster interpretation in remote sensing images with autonomous agents. https://doi.org/10.48550/arXiv.2410.13384, 2024.

[14] D. Saravanan, D.D. Joseph, D.A. Sreeram, D.C.M. Subrahmanyam, D.K. Murty, D.M. Kumar, Using information systems and intelligent agents during natural occurrences, Int. J. Environ. Sci. 11 (2025).

[15] B. Li, J. Ma, K. Yin, Y. Xiao, C.-W. Hsu, A. Mostafavi, Disaster management in the era of agentic AI systems: a vision for collective human-machine intelligence for augmented resilience. https://doi.org/10.48550/arXiv.2510.16034, 2025.

[16] Y. Xie, B. Jiang, T. Mallick, J. Bergerson, J.K. Hutchison, D.R. Verner, J. Branham, M.R. Alexander, R.B. Ross, Y. Feng, L.-A. Levy, W.J. Su, C.J. Taylor, MARSHA: multi-agent RAG system for hazard adaptation, Npj Clim. Action 4 (2025) 70, https://doi.org/10.1038/s44168-025-00254-1.

[17] M. Panchal, Y.-J. Chen, S. Parkash, CC-GRMAS: a multi-agent graph neural system for spatiotemporal landslide risk assessment in high Mountain Asia. https://doi.org/10.48550/arXiv.2510.20875, 2025.

[18] M. Berti, M. Pizziolo, M. Scaroni, M. Generali, V. Critelli, M. Mulas, M. Tondo, F. Lelli, C. Fabbiani, F. Ronchetti, G. Ciccarese, N. Dal Seno, E. Ioriatti, R. Rani, A. Zuccarini, T. Simonelli, A. Corsini, RER2023: the landslide inventory dataset of the May 2023 Emilia-Romagna meteorological event, Earth Syst. Sci. Data 17 (2025) 1055–1074, https://doi.org/10.5194/essd-17-1055-2025.

[19] M.F. Ferrario, F. Livio, Rapid mapping of landslides induced by heavy rainfall in the Emilia-Romagna (Italy) Region in May 2023, Remote Sens. 16 (2023) 122, https://doi.org/10.3390/rs16010122.

[20] F. Caleca, L. Lombardo, S. Steger, H. Tanyas, F. Raspini, A. Dahal, C. Nefros, M.C. Mărgărint, V. Drouin, M. Jemec-Auflič, A. Novellino, M. Tonini, M. Loche, N. Casagli, V. Tofani, Pan-European landslide risk assessment: from theory to practice, Rev. Geophys. 63 (2025), https://doi.org/10.1029/2023RG000825.

[21] P.K. Sikdar, S. Chakraborty, Enakshi Adhya, P.K. Paul, Land use/land cover changes and groundwater potential zoning in and around raniganj coal mining area. Bardhaman District, West Bengal – a GIS and Remote Sensing Approach, 2004.

[22] J.W. Rouse, R.H. Haas, J.A. Schell, D.W. Deering. Monitoring Vegetation Systems in the Great Plains with ERTS, NASA, Washington, DC, USA, 1973.

[23] H. Hashim, Z. Abd Latif, N.A. Adnan, Urban vegetation classification with NDVI threshold value method with very high resolution (vhr) pleiades imagery, Int. Arch. Photogram. Rem. Sens. Spatial Inf. Sci. XLII-4/W16 (2019) 237–240, https://doi.org/10.5194/isprs-archives-XLII-4-W16-237-2019.

[24] O. Hungr, S. Leroueil, L. Picarelli, The varnes classification of landslide types, an update, Landslides 11 (2014) 167–194, https://doi.org/10.1007/s10346-013-0436-y.

[25] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in: N. Navab, J. Hornegger, W.M. Wells, A.F. Frangi (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Lecture Notes in Computer Science, Springer International Publishing, Cham, 2015, pp. 234–241, https://doi.org/10.1007/978-3-319-24574-4_28.

[26] Z. Zhou, M.M.R. Siddiquee, N. Tajbakhsh, J. Liang, UNet++: a nested U-Net architecture for medical image segmentation. https://doi.org/10.48550/arXiv.1807.10165, 2018.

[27] M. Tan, Q.V. Le, EfficientNet: rethinking model scaling for convolutional neural networks. https://doi.org/10.48550/arXiv.1905.11946, 2020.

[28] E. Xie, W. Wang, Z. Yu, A. Anandkumar, J.M. Alvarez, P. Luo, SegFormer: simple and efficient design for semantic segmentation with transformers. https://doi.org/10.48550/arXiv.2105.15203, 2021.

[29] S.T. McColl, S.J. Cook, A universal size classification system for landslides, Landslides 21 (2024) 111–120, https://doi.org/10.1007/s10346-023-02131-6.