AGU ADVANCING EARTH AND SPACE SCIENCES

JGR Machine Learning and Computation

RESEARCH ARTICLE
10.1029/2025JH000857

Key Points:

- A multimodal landslide detection model with a Client/Server architecture utilizes multimodal remote sensing data from various data owners
- A distributed machine learning framework based on vertical federated learning is proposed to protect the privacy of geohazard data
- Experimental results across three study areas demonstrate that the proposed model maintains prediction accuracy while ensuring data privacy

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

X. Fan and X. Tang,
fxm_cdut@qq.com;
xc.tang@qq.com

Citation:

Tang, X., Lu, Z., Fan, X., Yan, X., Li, H., Jiang, S.-H., et al. (2026). FMLD: A vertical federated learning framework for privacy-preserving multimodal landslide detection. *Journal of Geophysical Research: Machine Learning and Computation*, *3*, e2025JH000857. https://doi.org/10.1029/2025JH000857

Received 28 JUL 2025
Accepted 11 JAN 2026

Author Contributions:

Conceptualization: Xiaochuan Tang, Filippo Catani
Data curation: Xuanmei Fan, Sansar Raj Meena
Formal analysis: Xiaochuan Tang, Filippo Catani
Funding acquisition: Xiaochuan Tang, Xuanmei Fan, Hualiang Li, Shui-Hua Jiang, Filippo Catani
Investigation: Zhong Lu, Xuanmei Fan, Filippo Catani

FMLD: A Vertical Federated Learning Framework for Privacy-Preserving Multimodal Landslide Detection

Xiaochuan Tang[1,2,3], Zhong Lu[2], Xuanmei Fan[1], Xiaochuang Yan[2], Huailiang Li[1], Shui-Hua Jiang[4], Sansar Raj Meena[3], Dongfen Li[2], Kushanav Bhuyan[1], Alessandro Novellino[5], Hongjun Li[2], and Filippo Catani[3]

[1]State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Chengdu, China, [2]College of Computer Science and Cyber Security, Chengdu University of Technology, Chengdu, China, [3]Department of Geosciences, University of Padova, Padova, Italy, [4]School of Infrastructure Engineering, Nanchang University, Nanchang, China, [5]British Geological Survey, Nottingham, UK

Abstract Landslides are among the most severe global geohazards posing a significant threat to human life and infrastructure. To support landslide detection and prediction, various geohazard monitoring approaches have been developed, such as optical remote sensing imagery, light detection and ranging, and ground-based sensors, generating vast volumes of landslide-related data. However, these data often involve privacy and security concerns for data owners. Institutions such as private companies, national space agencies, and geological survey departments are often reluctant to share geohazard data, which hinders the development of machine learning-based technologies for landslide assessment. To address this challenge, this study proposes a privacy-preserving framework named Federated Multimodal Landslide Detection (FMLD), based on vertical federated learning, enabling secure and collaborative model training across multiple institutions. FMLD integrates complementary multimodal data from different organizations, including optical imagery, digital elevation models, and hillshade maps, allowing the model to exploit the strengths of each modality while keeping raw data private. Extensive experiments conducted in three study areas show that FMLD achieves comparable landslide detection accuracy with its centralized counterpart. The proposed method effectively protects data privacy and security, thereby enhancing geological data sharing. This study demonstrates a practical pathway toward secure, collaborative, and knowledge-complementary artificial intelligence applications in geoscience.

Plain Language Summary Natural hazards cause significant damage and pose a serious threat to people around the world. Various sensors have been developed to monitor these events, generating tremendous data such as satellite images, ground deformation time-series, rainfall, etc. However, the owners of these data such as private companies and national space agencies, and geological survey departments are often reluctant to share sensitive data sets due to privacy and regulatory concerns, which leads to data islands. These islands significantly hinder the development of machine learning models for natural hazard assessment. Federated learning, a technique that enables joint model training without sharing raw data, offers a promising way to overcome this barrier. In this study, we introduce a vertical federated learning framework that combines complementary multimodal data from multiple organizations, such as optical images and hillshade maps, to improve landslide detection accuracy while preserving privacy. Experiments on three real-world regions confirm that the proposed method protects data privacy while maintaining strong predictive performance. This approach provides a feasible and effective path to encourage data sharing and advance data-driven natural hazard research.

1. Introduction

Landslides are among the most frequent and destructive geological hazards, posing significant threats to human life, infrastructure, and environmental stability in mountainous, seismic, and rainfall-prone regions. Rapid and accurate detection and mapping of landslides are fundamental to hazard assessment and risk mitigation in engineering geology. With the increasing volume and complexity of spatial-temporal data from remote sensing and in situ monitoring systems, Artificial Intelligence (AI), especially deep learning-based methods, has become instrumental in advancing landslide research (Ghorbanzadeh, Xu, Ghamisi, et al., 2022; Tehrani et al., 2022). These models have demonstrated strong capabilities in automating the extraction and classification of landslide-

Methodology: Xiaochuan Tang, Zhong Lu
Resources: Xiaochuan Tang, Sansar Raj Meena, Alessandro Novellino, Filippo Catani
Software: Zhong Lu, Xiaochuang Yan, Hongjun Li
Supervision: Filippo Catani
Validation: Xuanmei Fan, Kushanav Bhuyan, Alessandro Novellino
Visualization: Zhong Lu, Hongjun Li
Writing – original draft: Xiaochuan Tang
Writing – review & editing: Xiaochuan Tang, Xiaochuang Yan, Huailiang Li, Shui-Hua Jiang, Kushanav Bhuyan, Alessandro Novellino, Filippo Catani

related features from various data sources such as optical imagery (Dong et al., 2024; Tang et al., 2025), Light Detection and Ranging (LiDAR) (Farmakis et al., 2022; D. Li, Tang, et al., 2023), and other geospatial sources.

However, although many landslide inventory maps (i.e., landslide labels) are publicly available (Berti et al., 2025; Fan et al., 2019), the practical deployment of AI in landslide detection still faces a critical barrier, that is, the restricted accessibility and privacy sensitivity of high-resolution multimodal raw data. For example, Unmanned Aerial Vehicle (UAV) imagery from the public CAS (Xu et al., 2024) and GLGCD (Fang et al., 2024) data sets is not directly accessible, as it is subject to institutional or licensing restrictions. This raises an important question, how can we train AI models using data sets that contain only labels without access to the corresponding raw input data? Actually, landslide-related data are often generated through costly field instrumentation, proprietary aerial surveys, or strategic satellite missions. These raw data may include imagery of critical infrastructure, mining activity, or terrain features with national security implications (Zhu et al., 2023). As a result, the raw data are typically siloed across governmental agencies, commercial satellite operators, mining enterprises, and academic institutions. Legal frameworks such as the General Data Protection Regulation (GDPR) further impose strict constraints on data sharing and cross-border exchange (D. Li, Xie, et al., 2023). This creates a paradox in landslide modeling: the most informative raw data are often inaccessible for joint analysis, undermining the generalizability and robustness of AI-based hazard models.

The assumption of centralized, shareable landslide data sets that include both labels and raw data is frequently unrealistic in real-world applications. Different types of landslides require distinct forms of remote sensing data for effective detection (Casagli et al., 2023). For normal shallow landslides, publicly available optical satellite imagery is typically sufficient. However, other types of landslides often necessitate a combination of public and private remote sensing sources. For instance, in the rapid detection of earthquake-induced landslides, public satellite imagery is frequently limited by long revisit intervals and adverse weather conditions, which hinder timely acquisition of optical images over affected areas. In recent years, UAV imagery and commercial satellite data have been increasingly used for co-seismic landslide detection. Nonetheless, such data are not openly shared, which hinders the development of co-seismic landslide detection modeling. In the case of deep-seated landslides, vegetation cover and human engineering activities frequently obscure landslide boundaries, necessitating the use of high-resolution airborne LiDAR data and 3D optical remote sensing imagery. However, due to their high spatial resolution and the inclusion of sensitive topographic information, the sharing of such data is often constrained by legal and regulatory restrictions. These multimodal data are rarely managed by a single entity. Moreover, the vertical data split, where different institutions hold complementary feature sets for the same geographic regions, presents a unique challenge to collaborative modeling.

In this context, Federated Learning (FL) has emerged as a promising privacy-preserving framework, allowing decentralized model training without exchanging raw data (McMahan et al., 2017). FL operates under three main paradigms, that is, Horizontal FL (HFL), where different clients share feature space but differ in samples; Vertical FL (VFL), where clients share samples but differ in features; and Federated Transfer Learning (FTL), where both sample and feature spaces differ (Yin et al., 2021). Most existing applications of FL in landslide research have focused on the HFL setting. For example, Tang et al. (2024) proposed a horizontal federated learning model to identify landslides from remote sensing images. Joshi et al. (2023) applied federated learning for landslide monitoring and prediction. Yang et al. (2023) introduced an HFL approach to predict landslide displacement using ground-based monitoring data. Elmoulat et al. (2021) proposed an HFL framework for training AI models in a landslide early warning system. X. Zhang et al. (2023) utilized an HFL model for general remote sensing image segmentation, with application to a landslide data set. By contrast, VFL directly targets the common feature-disjoint but sample-aligned setting in geohazards, where different institutions hold complementary modalities over the same terrain, making it particularly suitable for privacy-aware multimodal landslide detection (Fang et al., 2024; Xu et al., 2024). For instance, high-resolution optical imagery may be available to a national space agency, while LiDAR data and geotechnical parameters may be collected by local geological survey departments. These data sets represent distinct feature spaces over the same terrain. A VFL framework enables each party to contribute to a global landslide detection model without sharing its proprietary data. This approach not only aligns with privacy regulations but also reflects the operational reality of data acquisition in engineering geology and remote sensing. In contrast, the VFL framework, which is well-suited for scenarios where institutions hold complementary geohazard data for the same geographic regions but are unable to share it directly, remain largely unexplored in the geohazard domain.

This study proposes a Federated Multimodal Landslide Detection (FMLD) framework based on vertical federated learning, aimed at addressing the data-sharing bottleneck in landslide detection modeling. To the best of our knowledge, this paper represents one of the first attempts to study the privacy issue in landslide detection using vertical federated learning. The framework is particularly suited to feature-disjoint but sample-aligned data settings, which are typical in multimodal geohazard data sets managed by different institutions. The proposed FMLD framework constructs a distributed Client/Server architecture, where each client represents a distinct data owner managing a unique modality (e.g., optical imagery, Digital Elevation Models (DEMs), hillshade maps). The server coordinates encrypted feature fusion and model updates while ensuring that raw data remain local. An incentive mechanism is also designed to quantify each client's contribution to the overall model performance. The main contributions of this study are as follows: (a) A vertical federated learning strategy is proposed for privacy-preserving landslide detection using multimodal data, addressing the feature-disjoint but sample-aligned nature of landslide data sets across institutions. (b) A federated learning architecture with a tailored task initialization protocol is designed for landslide modeling, incorporating an incentive-based contribution evaluation algorithm to promote fair and effective participation among clients. (c) The effectiveness of the proposed framework is evaluated in three study areas, demonstrating its feasibility in terms of accuracy, privacy preservation, and scalability.

The remainder of this paper is shown below. Section 2 reviews related work on landslide modeling and federated learning. Section 3 details the FMLD framework, including model architecture and training protocol. Section 4 presents the experimental results and analysis. Section 5 discusses the broader implications and limitations. Section 6 concludes the study and describes future research directions.

2. Related Works

2.1. Machine Learning-Based Landslide Detection

Automatic landslide detection refers to the development of data-driven models to identify landslides from remote sensing data. In the following, we provide a brief review of relevant works in landslide detection, comprehensive surveys can be found in Casagli et al. (2023); Mohan et al. (2021); Tehrani et al. (2022). Machine learning techniques have been widely used in landslide detection, such as random forest (Chen et al., 2014), convolutional neural network (Liu et al., 2023; Lu et al., 2023), transformer network (Lv et al., 2023; Tang et al., 2022) and graph neural network (W. Li et al., 2023; Wei et al., 2023). These models heavily rely on the availability of landslide-related geospatial data. The rapid growth of such data, including optical imagery, DEMs, and other remote sensing sources, provides a promising foundation for the development of machine learning-based landslide detection. Various types of landslide monitoring instruments have generated a large amount of heterogeneous data with different spatial, temporal, and spectral resolutions (Casagli et al., 2023). Space-borne optical images are widely used in landslide detection. Meena et al. (2023) used Landsat images to build a globally distributed landslide data set and test the performance of deep learning-based landslide detection methods. Airborne optical images have the advantage of high resolution, which has been used in rapid mapping of earthquake-triggered landslides (Dai et al., 2023; Tang et al., 2022). LiDAR data has the advantage of removing vegetation cover, which has been used to identify forested landslides (Fang et al., 2022). Synthetic Aperture Radar (SAR) is an active remote sensing technology that remains effective even under cloud cover and supports continuous operation day and night. Multi-temporal SAR images with image classification and segmentation methods are proposed for landslide detection (Nava et al., 2022, 2024). InSAR generates time-series data for ground deformation, which can be used to measure the movement of landslides. J. Cai et al. (2023) used time-series InSAR and Faster RCNN to identify active landslides over wide areas. Multimodal landslide detection attracts increasing attention (Ghorbanzadeh, Xu, Ghamisi, et al., 2022; Xu et al., 2024). Typically, these data are collected by different organizations, and the raw data are not allowed to directly share with others. There are several justifications for refraining from sharing data., for example, industrial competition (Xu et al., 2023), privacy protection policy and national security (Yang et al., 2023).

Some efforts have been made to promote data sharing in landslide detection. For example, Ghorbanzadeh, Xu, Zhao, et al. (2022) proposed a Landslide4Sense data set and used it to organize a landslide detection competition (Ghorbanzadeh, Xu, Zhao, et al., 2022). Landslide4Sense consists of multimodal data, including multispectral images, digital elevation model (DEM), and slope data. Meena et al. (2023) proposed a landslide detection data set named HR-GLDD. It contains Planet images and landslide mapping labels of 10 different study areas. Xu et al. (2024) proposed the CAS Landslide Data set. It consists of Sentinel-2 images and unmanned aerial vehicle images from nine regions. Fang et al. (2024) introduced a coseismic landslide data set called GDCLD, which consists of multi-source high-resolution remote sensing images. Most of the existing landslide detection data sets use publicly available monitoring data with some newly annotated landslides. While these data sets can serve as benchmarks for evaluating various landslide detection models, they are often insufficient to support practical applications of different regions and scales. Regional-scale landslide detection depends on the special characteristics of the study area, whereas global-scale landslide detection models must account for diverse conditions. Therefore, incorporating confidential landslide monitoring data is essential for developing generalizable landslide detection models.

Data silos have become a significant bottleneck in landslide detection. How to address the privacy and security issues associated with landslide detection has become an urgent concern. This paper introduced federated learning to address this challenge.

2.2. Vertical Federated Learning

Vertical federated learning is a federated learning framework that enables multiple organizations to collaboratively train a shared model using complementary features of the same set of samples without exposing their raw data. This paradigm is well-suited for integrating multimodal data distributed across institutions, where each client contributes a distinct feature modality such as optical imagery or elevation data, and the global model is trained on the aggregated representations without direct data exchange. By preserving data privacy while leveraging heterogeneous information sources, VFL facilitates the development of more robust machine learning models and has attracted increasing attention in recent years for privacy-sensitive applications. Liu et al. (2024) studied the basic framework, communication efficiency, effectiveness, security, and incentive of FTL. Hardy et al. (2017) proposed a federated learning method for vertically partitioned data, which is considered to be one of the earliest works on vertical federated learning (Saadati et al., 2024; Yang et al., 2019). Wang et al. (2024) proposed an optimization approach for privacy and communication efficiency in VFL. He et al. (2024) proposed a hybrid self-supervised learning framework for VFL, which made better use of unaligned and unlabeled samples. Feature inference (Luo et al., 2021) and label inference (Fu et al., 2022) attacks are proposed against VFL. Lu et al. (2022) proposed a trustful incentive mechanism for VFL.

VFL has been successfully applied in several domains, such as recommendation systems (Atarashi & Ishihata, 2021), finance (Kang et al., 2022), vehicle (Teimoori et al., 2022), wireless communication (Shi et al., 2023), and multimodal tasks (Qi et al., 2023; Yan et al., 2024). However, the application of federated learning in the geosciences remains largely unexplored. A review of security challenges in geoscience and remote sensing, including adversarial attacks, backdoor attacks, federated learning, uncertainty, and explainability can be found in Xu et al. (2023). In remote sensing, federated have been applied to remote sensing data collection, fusion, annotation, image classification and segmentation. Fadlullah and Kato (2021) proposed an asynchronous federated learning approach for remote sensing over integrated terrestrial-aerial-space networks. Gao et al. (2020) proposed a federated region learning framework for urban environment sensing. Liu et al. (2020) applied federated learning to for UAV-based aerial-ground air quality monitoring. Lee (2022) applied federated reinforcement learning for aerial remote sensing. D. Li, Xie, et al. (2023) introduced vertical federated learning for multi-satellite and multimodal fusion of remote sensing data. Coca et al. (2020) proposed a federated system for collaborative labeling of remote sensing data sets. W. Cai et al. (2023) introduced horizontal federated learning for hyperspectral image classification. Zhu et al. (2023) proposed a federated learning method for remote sensing image classification with dishonest majority. Tam et al. (2021) applied federated learning to real-time remote sensing image classification. X. Zhang et al. (2023) introduced a horizontal federated learning model for general remote sensing image segmentation. They also applied the proposed method to a landslide data set. Z. Zhang et al. (2023) introduced differential privacy to address the white-box membership inference attacks in remote sensing image recognition.

In natural hazards assessment, federated learning has been applied to landslides, earthquakes, and forest fires. Tang et al. (2024) proposed a horizontal federated learning method named FedLD for landslide detection. FedLD was able to accurately identify landslides while protecting data privacy. Yang et al. (2023) applied federated learning for predicting landslide displacement. Joshi et al. (2023) used federated learning for landslide monitoring and prediction. Elmoulat et al. (2021) introduced a horizontal federated learning approach for developing

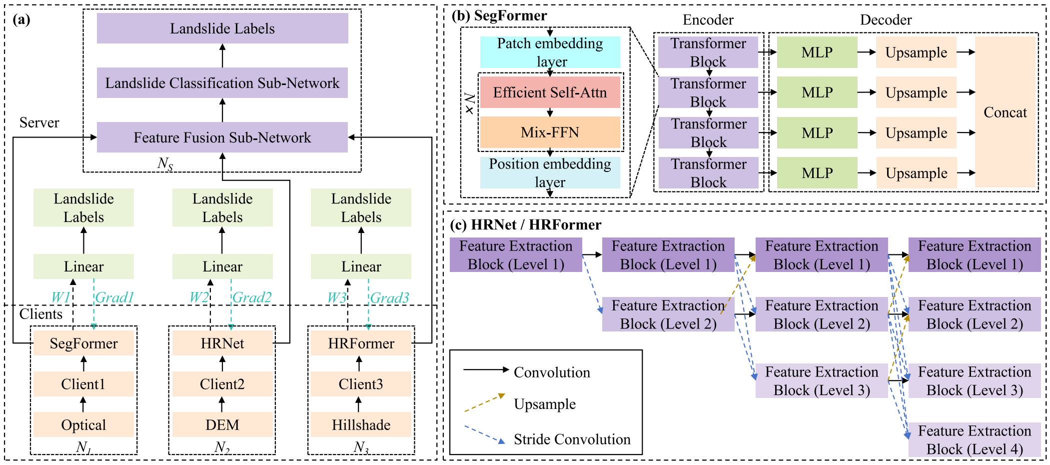

Figure 1. The framework of the federated multimodal landslide detection method.

landslide early warning system. Tehseen et al. (2021) proposed a federated learning method for earthquake prediction. Fadlullah and Kato (2021) proposed an asynchronous federated learning model for smart remote sensing with application to forest fire detection. Despite its potential, the integration of federated learning into landslide detection research remains underexplored. This highlights a gap between current data sharing practices in landslide studies and the opportunities offered by federated learning frameworks. In this study, we proposed a vertical federated learning framework for landslide detection that enables the integration of multimodal landslide-related data and supports collaborative modeling across institutions.

3. Method

3.1. Problem Definition

Denote the data set as $\mathcal{D} = \{X_1, X_2, \ldots, X_p, Y\}$, where $\mathcal{X} = \{X_i\}_{i=1}^{p}$ represents the landslide-related data maintained by different data owners, $Y = \{y_k\}_{k=1}^{n}$ is handcraft label of landslides, p is the number of modalities, and n is the number of landslide samples. The landslide data $X_i = \{x_{ik}\}_{k=1}^{n}$ is a type of multimodal data, such as optical images, LiDAR-derived DEM, and Hillshade map. Assume the landslide data are held by different organizations. The landslide detection problem is how to use the data set $\mathcal{D}$ to develop a machine learning model for identifying landslides automatically.

3.2. Overall Framework

To address the challenges of restricted data sharing and multimodal data integration in landslide detection, this paper proposes a new landslide detection approach named FMLD. Vertical federated learning is first introduced to the field of multimodal landslide detection. It enables multiple institutions, each holding complementary types of landslide-related remote sensing data, such as optical images, DEM, and hillshade maps to collaboratively train a landslide detection model without directly sharing their raw data. This is particularly important in engineering geology, where high-resolution remote sensing data and field-based geotechnical information are often confidential, region-specific, and costly to acquire. As illustrated in Figure 1, the FMLD framework adopts a Client-Server (C/S) architecture, consisting of a central server and multiple clients. Each client represents an organization that holds a specific modality of landslide data, such as optical images, LiDAR-derived DEM, and hillshade maps. The primary function of each client is to extract meaningful landslide features from their local data. The server aggregates these features and performs the landslide detection task based on the integrated multimodal information. This architecture ensures that the computational load is balanced. Clients handle data preprocessing and local model training using their respective data sets, while the server integrates the intermediate features to update the global landslide detection model. This approach reduces the volume of intermediate data transmission, protects data privacy, and improves training efficiency through distributed computing. Additionally, the C/S structure enhances the scalability of the system, as new data sources or institutions can be incorporated into the system by adding new clients without requiring changes to the overall framework.

The architecture of FMLD is a distributed neural network, also known as split neural network (SplitNN) (Thapa et al., 2022). SplitNN partitions a neural network into multiple sub-networks. FMLD adopts this paradigm to transform the centralized multimodal landslide detection model (D. Li, Tang, et al., 2023) into a decentralized framework. In this architecture, each client (e.g., imagery provider, LiDAR provider, or DEM holder) holds a subnetwork of the model that processes its local data modality. Each modality is handled by a dedicated encoder tailored to its specific data type. This modular design allows for the seamless integration of new clients with additional modalities without requiring a complete redesign of the system. The server receives intermediate feature embeddings from all clients and contains its own sub-network to perform multimodal feature fusion and landslide prediction. Notably, only compact feature embeddings are transmitted between clients and the server, rather than raw images or DEM data, thereby enhancing privacy and reducing communication overhead.

Algorithm 1 depicts the clients and server of FMLD. In the rest of this section, we will show the design details of FMLD.

Algorithm 1. The design details of FMLD.

Input: The client $C_i (i = 1, 2, 3)$ has a data source X_i. The server S has a set of landslide labels Y.
Output: A landslide detection model M.

// *Server*
Initialize the task using the Task Initialization Protocol
Construct a split neural network (N_s, N_1, N_2, N_3)
for $i = 1$ to 3 **do**
  Send network N_i to client C_i
end for
while There exists C_i not finish feature learning **do**
  Client C_i send feature map Z_{ij} to server S
  Use landslide lables Y to compute loss $\mathcal{L}_i^j$
 if $|\mathcal{L}_i^j - \mathcal{L}_i^{j-1}| < \varepsilon$ **then**
   Terminate the feature learning process of C_i
 else
   Send $\mathcal{L}_{ij}$ to client C_i
   end if
 end while
   Use the feature fusion sub-network to integrate clients' feature maps
   Use the landslide detection sub-network to predict landslides
   // *Client*
   Receive feature learning sub-network N_i from the server S
 while !$|\mathcal{L}_i^j - \mathcal{L}_I^{J-1}| < \varepsilon$ **do**
   Use local data X_i to train N_i by forward propagation
   Send feature map Z_{ij} to the sever S
   Receive loss $\mathcal{L}_i^j$ from the server
   Conduct backpropagation to update the parameters of N_i
end while
Fix the parameters of feature learning sub-network

3.3. The Server of FMLD

A semi-asynchronous server is proposed to manage the entire lifecycle of FMLD. Its core functions include implementing the task initialization protocol, constructing the split neural network, coordinating clients in training modality-specific feature extraction sub-networks, and performing landslide prediction based on multimodal feature fusion.

3.3.1. Task Initialization Protocol

The task initialization protocol is designed to address quality control, spatial alignment, label trustworthiness, and data relevance challenges inherent in privacy-preserving, vertically federated landslide detection.

1. Pre-Federation Collaboration Agreement. A shared geographic extent and spatial resolution grid are established to ensure data alignment across all data modalities. The task scope explicitly defines whether the focus is on shallow landslides, deep-seated landslides, or composite events. This ensures each modality used is appropriate for the landslide detection task. Additionally, the server requests each client to prepare the necessary software environment and hardware resources to ensure successful participation in the federated training process.
2. Data Quality Declaration and Verification. Each data owner provides metadata describing their data set, including: spatial resolution and coverage, acquisition time and revisit rate, and sensor type and calibration (e.g., orthorectification status, cloud masking procedures, and DEM interpolation methods).
3. Label Source Disclosure and Synchronization. The central server maintains documentation of label provenance, including sources such as public landslide inventories and human-labeled data sets from domain experts. The spatial footprint of each landslide label is cross-checked with available modalities to ensure at least partial visibility in each.
4. Data Overlap and Alignment Assurance. All clients are required to project their data into a common coordinate reference system and align to a shared tiling grid. The server verify that at least a defined percentage of labeled tiles have overlapping features across all required modalities.
5. Initialization Approval and Logging. All parties approve the final protocol checklist via digital signature. The federated learning process is initiated only after verification of data modality and label overlap, validation of label relevance. and confirmation of modality appropriateness. All steps are logged and reproducible to support review and audits.

3.3.2. Constructing Split Neural Network

Firstly, the server S constructs a split network (N_s, N_1, N_2, N_3), where the sub-networks N_1, N_2 and N_3 are allocated to client C_1, C_2 and C_3, respectively, for learning landslide detection features from multimodal data, such as optical imagery, DEM, and hillshade. Meanwhile, the sub-network N_s is maintained by the server to integrate the features from all the clients and to predict the polygon label of landslides.

3.3.3. Coordinating Clients to Train the Feature Learning Sub-Networks

After assigning sub-networks to individual clients, the server orchestrates their training through an asynchronous strategy. Specifically, the server waits to receive feature maps from the clients. Upon receiving a feature map Z_{ij} from client C_i, where j denotes the training iteration number, the server computes the pixel-wise cross-entropy (CE) loss $\mathcal{L}_{ij}$ using the landslide labels. This process enables classification at the pixel level. The loss function is presented in Equation 1, where the superscript l denotes the local loss associated with client. The computed loss $\mathcal{L}_{ij}$ is then return to client C_i. This approach enables clients to train their feature learning sub-networks in parallel and asynchronously, thereby improving overall training efficiency.

$$\mathcal{L}_{ij}^{(l)} = CE\left(\hat{Y}_{ij}^{(l)}, Y\right) \tag{1}$$

3.3.4. Predicting Landslides Using Multimodal Data

Once all the clients have finished the training of their sub-networks, the server uses its sub-network N_s to integrate the feature maps from all the clients and predict the landslide labels.

The network structure of N_s is shown in Figure 1. To enhance multimodal feature fusion, the model incorporates both channel and spatial attention mechanisms. The channel attention (Tang et al., 2021) layer assigns adaptive weights to features from different modalities (e.g., optical imagery, DEM, hillshade), enabling the model to emphasize the most relevant information based on landslide type, for example, prioritizing optical features for earthquake-triggered landslides, and DEM or Hillshade features for vegetation-covered ones. The spatial attention (D. Li, Tang, et al., 2023) layer further guides the model to focus on geospatial regions more likely to contain landslides, improving localization and reducing interference from irrelevant terrain or occlusion.

Feature maps extracted from optical imagery (1024 channels via SegFormer), DEM (720 channels via HRNet), and hillshade (1170 channels via HRFormer) are concatenated to form a unified tensor with 2914 channels. Prior to concatenation, DEM and hillshade features are interpolated to align with the spatial resolution of the optical features. Global average pooling is then applied to aggregate channel-wise information, serving as input for subsequent attention mechanisms. A Squeeze-and-Excitation module is used to refine the DEM and hillshade features (1890 channels in total) before fusion. The concatenated tensor subsequently undergoes global average pooling, followed by a sequence of convolutional layers that first compress the channel dimension to a maximum of 8, and then expand it back to 1024 and 1890 channels. ReLU and Sigmoid activations are applied sequentially to compute channel-wise attention weights, which are used to enhance critical feature representations. To further capture spatial dependencies, horizontal and vertical attention are applied using adaptive average pooling and convolutions layers with kernel sizes 7×1 and 1×7, followed by Sigmoid activation. Their element-wise product of these feature maps forms a spatial mask, which highlights landslide-prone regions. The resulting fused features are passed through a dropout layer with a rate of 0.1, a linear layer projecting to 1024 channels, and final 1 convolution layer to perform pixel-wise landslide classification. Model training of the server is guided by a cross-entropy loss function, defined in Equation 2, where the superscript g denotes the global loss of the server.

$$\mathcal{L}^{(g)} = CE\left(\hat{Y}^{(g)}, Y\right) \tag{2}$$

3.4. The Client of FMLD

After receiving feature learning sub-network N_i from the server, each client performs local training. Specifically, client C_i uses its local data X_i to train N_i. The training process consists of three stages, including forward propagation, backward propagation, and fix parameters. In the jth round, client C_i first uses a batch of its data to perform forward propagation on sub-network N_i. The output of N_i is a feature map Z_{ij}. Then, client C_i sends Z_{ij} to the server and waits for the server to return the loss $\mathcal{L}_{ij}$. When client C_i receives loss, it conducts back propagation. The training process continues for a maximum of N rounds, after which all parameters of sub-network N_i are fixed. Finally, while the server trains its sub-network N_s, all clients only perform forward propagation without updating their parameters.

The structure of the sub-networks N_1, N_2 and N_3 are illustrated in Figure 1. Different networks are adopted because each modality exhibits distinct spatial and spectral properties requiring specialized feature extraction. Specifically, SegFormer (Tang et al., 2022) is used for optical imagery because of its strong capability in capturing fine-scale spatial and semantic features from high-resolution RGB data. HRNet (J. Wang et al., 2020) is selected for DEMs since it maintains high-resolution representations, enabling effective modeling of continuous terrain structures. HRFormer (Yuan et al., 2021) is adopted for hillshade maps as it efficiently captures illumination-dependent topographic textures using transformer-based global attention. These architectures complement each other and allow each modality to contribute effectively within the unified FMLD framework. Furthermore, the integration of heterogeneous deep learning architectures within the FMLD framework demonstrates its flexibility and generality as a federated learning system capable of supporting diverse model types.

3.4.1. SegFormer for Extracting Optical Features of Landslides From Optical Imagery

The SegFormer network is employed to extract optical features of landslides, such as surface textures, edge patterns, and morphological disruptions, from high-resolution remote sensing images acquired by satellites or UAVs. As shown in Figure 1a, SegFormer is a lightweight and scalable vision transformer based on an encoder-decoder architecture. Landslides vary significantly in size, shape, and spatial distribution, often occurring in complex mountainous terrain. Capturing both fine local structure and broader contextual information is crucial for accurate detection. To this end, the encoder adopts a hierarchical structure composed of four transformer layers, which progressively downsample the input to generate multi-scale feature maps. This design allows the model to simultaneously represent slope-level and landscape-level features. Compared to conventional convolutional neural networks, vision transformers offer enhanced capability in modeling long-range dependencies. This is particularly advantageous in landslide detection, where spatially scattered features such as source zones, transport paths, and accumulation areas must be interpreted as part of a unified geomorphic process. The use of overlapping patch embedding and depthwise convolutions further preserves spatial detail and positional information relevant to landslide boundaries.

The decoder of SegFormer is implemented for high computational efficiency. It fuses the multi-scale features extracted by the encoder into a uniform-resolution representation using MLP layer and upsampling operations. A final 1×1 convolution performs pixel-wise classification, assigning each pixel to either landslide or non-landslide classes. By leveraging transformer-based modeling in the context of landslide detection, this approach improves the extraction of complex geomorphic features from optical imagery and enhances detection performance across varied landslide types and spatial scales.

3.4.2. HRNet for Extracting Geomorphological Features of Landslides From DEM

A DEM provides critical information about landslide's surface by representing terrain elevation. It provide complementary information for identifying geomorphological changes that are not always visible in optical imagery. Topographic features derived from DEM, such as slope, aspect, and curvature, are well-known indicators for detecting geomorphic anomalies associated with landslides. These include features such as scarps, ridge lines, tension cracks, and accumulation zones, which are often precursors or consequences of slope failure. To effectively extract these subtle topographic features, the High-Resolution Network (HRNet) is employed for extracting landslide-related features from DEM. Unlike traditional CNN-based approaches that repeatedly downsample feature maps, HRNet is particularly well-suited for DEM analysis because it maintains high-resolution representations throughout the network.

As illustrated in Figure 1b, HRNet takes the raw DEM as a single-channel grayscale input and processes it through parallel branches, each built upon four sequential connected convolutional layers. Convolutional operations provide a sufficient receptive field to capture landslide topographic features, while being more computationally efficient than Transformer-based architectures. By retaining high-resolution features throughout the network, HRNet preserves detailed topographic variations such as landslide scarps and sudden slope changes. This ability offers a approach for terrain-driven landslide detection, which is particularly suitable for scenarios where spectral information is limited or unavailable.

3.4.3. HRFormer for Extracting Surface Cues of Vegetation-Covered Landslides From Hillshade

Hillshade is a terrain visualization technique derived from DEM that simulates terrain illumination based on a hypothetical light source. By enhancing terrain morphology through shading and contrast, hillshade maps are particularly effective for identifying landslides in densely vegetated areas where optical imagery is limited. Compared to optical images, high-resolution hillshade maps contain less landscape information but preserve critical surface cues under forest cover, such as texture, shadow, and edge features. Meanwhile, LiDAR point clouds tend to be sparse in regions with dense vegetation, making it essential to retain high spatial resolution during feature extraction.

To address these challenges, HRFormer is employed for feature extraction from hillshade maps, which is shown in Figure 1b. While its overall structure resembles HRNet, HRFormer replaces convolutional layers with vision transformer blocks, enabling more effective modeling of long-term spatial relationships. Hillshade maps are treated as three-channel inputs, similar to optical images, facilitating integration into the network. The multi-resolution transformer branches in HRFormer allow for the simultaneous capture of fine-scale shading variations and broader terrain patterns. Through cross-scale feature fusion, HRFormer integrates detailed local morphological features with the global geomorphological context, enhancing the discrimination of landslide features. Moreover, its high-resolution representation throughout the network preserves pixel-level contrast, which is crucial for detecting small or subtle landslides, particularly in areas obscured by vegetation.

Notably, the FMLD framework provides a flexible and architecture-independent solution for landslide detection using heterogeneous geospatial data. Instead of enforcing a uniform model across all clients, FMLD allows each client to adopt the deep learning architecture best suited to its specific data modality, considering differences in spatial, spectral, and geometric properties. This is particularly important in practical scenarios where data such as optical imagery, DEMs, and hillshade maps vary widely in structure. Through a split neural network design, FMLD enables intermediate features from diverse architectures to be projected into a shared latent space. This design ensures compatibility across models and supports the integration of multiple data views, scales, and sources, which is crucial for accurate landslide detection in complex and vegetated environments. Demonstrated with models including SegFormer, HRNet, and HRFormer, FMLD is also extensible to Synthetic Aperture Radar (SAR) data using appropriate deep neural networks.

3.5. Incentive Mechanism

Evaluating the contribution of each data owner is crucial for understanding the utility of their landslide data. In vertical federated learning, different parties contribute distinct groups of features to the learning task. These features often have complex interactions. The joint mutual information (JMI) (Tang et al., 2018) is introduced to measure the dependency between a set of features and the landslide label. Equation 3 provide the definition of JMI,

$$I(X_1, X_2, \ldots, X_k; Y) = \sum_{X \in \mathcal{X}} I(S; Y) \tag{3}$$

where $I(S; Y)$ is the interaction information between Y and all the features in S. The contribution of a feature X_i is quantified by the marginal loss of its JMI, as shown in Equation 4.

$$C_i = I(X_1, X_2, \ldots, X_k; Y) - I(X_1, \ldots, X_{i-1}, \ldots, X_{i+1}, \ldots, X_k; Y) \tag{4}$$

However, accurately computing multivariate information measures can be challenging. The marginal loss of the trained neural network is introduced to estimate JMI. Algorithm 2 shows the process of using a neural network to calculate the marginal loss. The search strategy is known as Sequential Backward Search (SBS). Firstly, compute the overall F1 score using all types of data, denoted as I_0. Secondly, remove X_i from the input data sets and compute the F1 score using the remaining data sets. Finally, the contribution of the jth type of landslide data is given by $C_j = I_0 - I_j$, and the contribution ratio is $R_j = \frac{C(X_i)}{C_1 + C_2 + C_3}$.

Algorithm 2. An incentive algorithm for federated landslide detection.

Input: A trained neural network of FMLD, that is, $\{N_s, N_1, N_2, N_3\}$; the optical images (X_1), DEM (X_2) and hillshade maps (X_3); and the input landslide labels Y.
Output: The contributions of the participants $\{C(X_1), C(X_2), C(X_3)\}$.

```
Use the original trained network to compute the total F1 score of multimodal
landslide prediction accuracy, which is denoted as I_0
for j = 1 to 3 do
  Fix the feature map of N_j as 0
  Input the test data of D into sub-networks N_1, N_2, and N_3
  Feed the feature maps of N_1, N_2, and N_3 into N_s
  Compute the predict label Ŷ
  Use Y and Ŷ to compute F1 score I_j
  The contribution of X_j is C_j = I_0 − I_j
  The contribution ratio of X_j is R_j = C(X_j) / (C_1 + C_2 + C_3)
end for
```

4. Experiments

4.1. Data Sets

The proposed method was evaluated in three study areas (Figure 2). Detailed data preprocessing procedures are provided in Text S1 of Supporting Information S1.

The first study area is located in Jiuzhaigou County, China. Jiuzhai Valley National Park, a famous World Heritage site, is well-preserved with abundant vegetation. Situated in the transition zone between the Tibetan Plateau and the Sichuan Basin, the region is characterized by extremely complex geological conditions. Intense neotectonic movements and significant crustal uplift have resulted in diverse landforms and extensive tufa deposition due to karstification (Lei et al., 2018). On 8 August 2017, a Ms 7.0 earthquake occurred in Jiuzhaigou County (33.20°N, 103.82°E), triggering approximately 2,000 landslides (Fan et al., 2018). We developed a Jiuzhaigou landslide data set using multimodal data, including post-event airborne optical imagery, LiDAR-

10.1029/2025JH000857

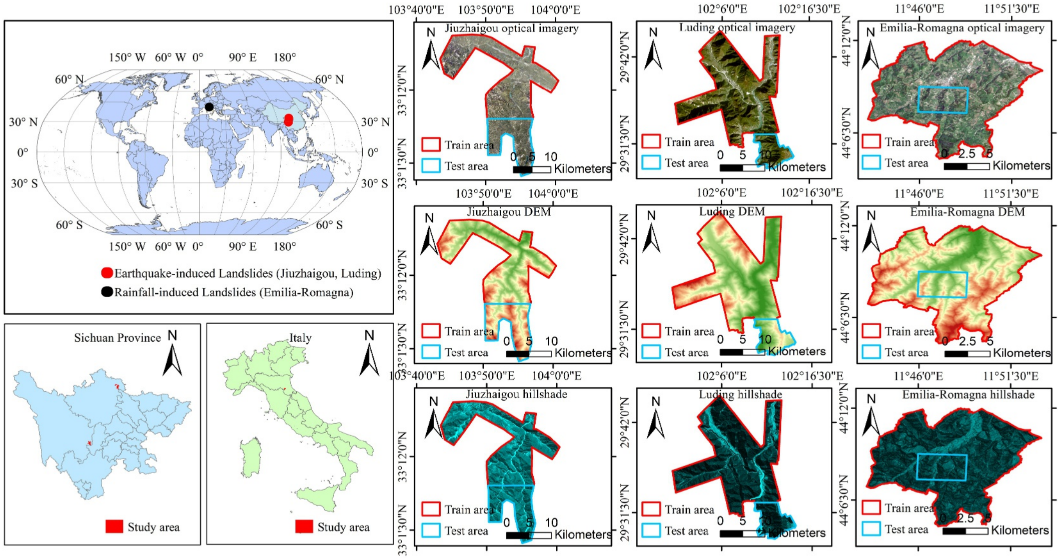

Figure 2. Study areas including Luding, Jiuzhaigou, and Emilia-Romagna, with separated training and testing regions and associated multimodal data (optical imagery, DEM, and hillshade map).

derived DEM and hillshade maps. The training set comprises 2,524 image patches and 2,524 landslides, while the test set contains 670 image patches and 727 landslides.

The second study area is Luding County, China. In September 2022, a magnitude 6.8 earthquake occurred in Luding County (39.25°N, 102.08°E) (Huang et al., 2023), triggering over 5,000 landslides. The Luding earthquake struck the southeastern margin of the Qinghai-Tibet Plateau, an area known for its typical alpine canyon landforms. The Dadu River flows through this region from north to south, with significant elevation changes along its path. The predominant lithologies in this area include acidic plutonic rocks and mixed sedimentary rocks. Long-term intense tectonic activity and weathering have resulted in highly fragmented rock and soil masses, creating conditions conducive to geological hazards (Dai et al., 2023). We developed a Luding landslide data set using multimodal data, including post-event PlanetScope satellite optical imagery, LiDAR-derived DEMs, and hillshade maps. The training set contains 1,721 image patches and 1,721 landslides, while the test set includes 972 image patches and 417 landslides.

The third study area is located in the central region of Emilia-Romagna, Italy (44.50°N, 11.35°E). In May 2023, this region experienced multiple extreme rainfall events, triggering more than 80,000 landslides, primarily consisting of shallow soil slides and debris flows (Berti et al., 2025). These landslides were widely distributed across slopes that had previously been considered stable. A total of 6,962 landslides were mapped in the selected study area. Given the dominance of shallow landslide types in this region, we utilized multimodal remote sensing data, including post-event PlanetScope satellite optical imagery, DEM, and hillshade maps. The training set consists of 6,343 image patches and 6,343 landslides, while the test set contains 681 image patches and 619 landslides.

4.2. Experimental Results

This section presents the experimental results obtained from three study areas. The experiments are designed to evaluate the accuracy, efficiency, and generalizability of the proposed FMLD framework. Six widely used quantitative metrics are adopted to assess the landslide detection results: IoU, mIoU, Precision, Recall, Accuracy, and F1 score. The definitions and computational formulas of these evaluation metrics are provided in the Appendix. Detailed experimental configurations are presented in Text S2 of Supporting Information S1.

4.2.1. Overall Performance Evaluation

Table 1 compares the overall performance of landslide detection on the Jiuzhaigou data set using single-modal data and multimodal data. "Centralized" refers to the model trained when all multimodal data are stored

Table 1
The Overall Performance Comparison of Different Landslide Detection Methods on the Jiuzhaigou Data Set

Model	Data source	Precision	Recall	Accuracy	mIoU	F1
SegFormer	Optical	0.6583	0.6378	0.9501	0.5701	0.6472
HRNet	DEM	0.6444	0.6914	0.9408	0.5793	0.6640
HRFormer	Hillshade	0.7043	0.7628	0.9532	0.6353	0.7293
Centralized	Optical + DEM + Hillshade	0.7395	0.7595	0.9604	0.6549	0.7490
FMLD	Optical + DEM + Hillshade	0.7484	0.7454	0.9619	0.6533	0.7469

together on a single server, representing the ideal case of centralized learning without privacy or data-sharing constraints. In contrast, FMLD operates under a federated learning setting, where different data modalities are held by separate organizations and collaboratively trained without sharing raw data. For single-modal data, the SegFormer, HRNet and HRFormer models are used to detect landslides in optical images, DEM and hillshade, respectively. For multimodal data, both centralized and FMLD are used to detect landslides using all three data sources. The results show that the multimodal landslide detection methods (Centralized and FMLD) outperform single-modal landslide detection methods (SegFormer, HRNet and HRFormer). The performance of the centralized method and FMLD are quite close. Because FMLD is a split neural network, which divides a neural network into sub-networks. The forward and backward flows between the server and clients are transmitted via a reliable network connection. This setup ensures that all the transmitted data, including feature maps and gradients, are well protected. In contrast, the model aggregation algorithm in horizontal federated learning can lead to a reduction in prediction accuracy.

Table 2 compares the overall landslide detection performance on the Luding data set. The results indicate that multimodal methods, specifically FMLD and Centralized, outperform single-modal methods including "Optical + SegFormer," "DEM + HRNet," and "Hillshade + HRFormer." Additionally, the performance of FMLD and Centralized is very similar. These findings are consistent with the results obtained from the Jiuzhaigou data set. This suggests that FMLD is a robust, privacy-preserving landslide detection method with strong generalizability, as it performs well across different data sets and does not depend on specific landslide data set.

Table 3 presents the overall landslide detection performance on the Emilia-Romagna data set using both single-modal and multimodal approaches. For single-modal detection, the SegFormer model is used on optical imagery, HRNet on DEM, and HRFormer on hillshade data. The results indicate that among single-modal methods, the optical-based SegFormer model achieves the best performance, while DEM and hillshade modalities yield relatively lower scores across all metrics. When using multimodal data, both the centralized model and the proposed FMLD method demonstrate significantly better performance than single-modal models, with improvements particularly evident in mIoU and F1-score. The centralized and FMLD methods exhibit similar overall performance, with FMLD achieving slightly higher precision and accuracy. These results further support the effectiveness and generalizability of the FMLD framework in handling diverse landslide data sets while preserving data privacy.

In addition, FMLD achieves high prediction accuracy across three study areas located on different continents, including China and Italy, and involving two distinct triggering factors, namely earthquakes and rainfall. This demonstrates its generalizability across diverse geological settings. It also successfully integrates three different

Table 2
The Overall Performance Comparison of Different Landslide Detection Methods on the Luding Data Set

Model	Data source	Precision	Recall	Accuracy	mIoU	F1
SegFormer	Optical	0.7769	0.7988	0.9659	0.6930	0.7874
HRNet	DEM	0.5330	0.5214	0.9395	0.4909	0.5251
HRFormer	Hillshade	0.8096	0.7421	0.9685	0.6778	0.7713
Centralized	Optical + DEM + Hillshade	0.8234	0.8051	0.9721	0.7223	0.8140
FMLD	Optical + DEM + Hillshade	0.8135	0.8251	0.9716	0.7279	0.8192

Table 3
The Overall Performance Comparison of Different Landslide Detection Methods on the Emilia-Romagna Data Set

Model	Data source	Precision	Recall	Accuracy	mIoU	F1
SegFormer	Optical	0.7253	0.7495	0.9126	0.6303	0.7365
HRNet	DEM	0.5682	0.5601	0.8708	0.4892	0.5637
HRFormer	Hillshade	0.5847	0.5570	0.8855	0.4957	0.5663
Centralized	Optical + DEM + Hillshade	0.7660	0.7516	0.9261	0.6544	0.7585
FMLD	Optical + DEM + Hillshade	0.7723	0.7424	0.9275	0.6526	0.7563

types of deep neural networks, including the lightweight transformer-based SegFormer, the convolutional neural network-based HRNet, and the transformer-based HRFormer. This highlights its generalizability with respect to different of neural network architectures.

4.2.2. Individual Class Comparison

Table S1 in Supporting Information S1 compares the performance of the landslide and background classes on the Jiuzhaigou data set. For the landslide class, the multimodal methods FMLD and Centralized outperform the single-modal approaches, including "SegFormer + Optical", "HRNet + DEM" and "HRFormer + Hillshade." The complementary information provided by multimodal data enhances landslide detection performance, which is learned by the feature fusion sub-network of the multimodal landslide methods. FMLD achieves performance close to that of the centralized method, with a 0.46% and 0.5% difference in IoU and F1 for the landslide class, respectively. This small gap indicates that FMLD provides an effective implementation of the centralized approach while preserving data privacy. For the background class, performance across all methods is similar. This is because the number of pixels classified as background significantly exceeds the number of pixels classified as landslides, leading to less variation in performance for this class.

Table S2 in Supporting Information S1 compares the performance of the individual classes on the Luding data set. Both the landslide and background classes perform better with multimodal methods compared to single-modal methods. The performance of the Centralized method and FMLD are similar. These individual class comparison results of the Luding data set are consistent with that of the Jiuzhaigou data set, showing that FMLD is a general framework for privacy-preserving landslide detection.

Table S3 in Supporting Information S1 presents the per-class performance on the Emilia-Romagna data set. For both the landslide and background classes, multimodal methods consistently outperform single-modal approaches. Notably, the performance of the Centralized method and FMLD is comparable. These results align with the findings from the Jiuzhaigou and Luding data sets, further demonstrating that FMLD serves as a robust and generalizable framework for privacy-preserving landslide detection.

4.2.3. Visualization Analysis

Figures 3–5 presents the visualization results of the proposed FMLD method on the Jiuzhaigou, Luding and Emilia-Romagna data sets. The results demonstrate that FMLD successfully identifies the majority of the landslides, confirming its effectiveness as a privacy-preserving landslide detection framework. Many landslides scattered along the river are accurately detected by FMLD, indicating its strong capability to capture geomorphic signatures associated with slope failures. Notably, FMLD also identifies landslides located in the shadow regions of the remote sensing image. The hillshade map, unaffected by shadows, provides complementary information to the optical remote-sensing images. FMLD successfully leverages multimodal landslide data to mitigate the impact of shadowing, enhancing detection accuracy.

However, several false positives are also observed, particularly in regions adjacent to rivers and water bodies. These areas often coincide with geomorphological features such as alluvial fans and river channel changes, which share similar spectral and topographic patterns with actual landslides. In addition, some agricultural areas, particularly cultivated farmland such as wheat fields, exhibit repetitive crop texture patterns that resemble the tonal and structural characteristics of landslides in optical imagery. These similarities in surface texture and brightness gradients may cause the model to misclassify such areas as potential landslides. This analysis

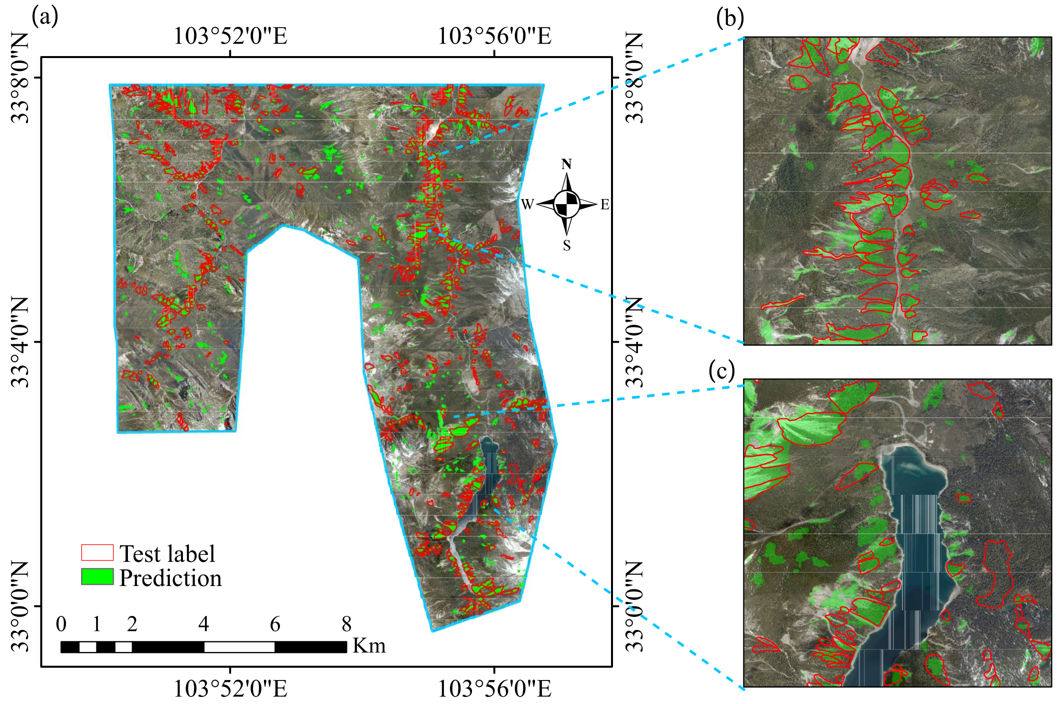

Figure 3. Visualizing the landslides identified by FMLD on the Jiuzhaigou data set. (a) Overall prediction results across the jiuzhaigou study area. (b) Enlarged view showing successful detection in shadowed regions. (c) Enlarged view highlighting accurate identification around water bodies and complex terrain.

highlights the importance of integrating additional contextual or temporal features in future work to further reduce false detections.

4.2.4. Ablation Study

To further evaluate the contribution of key components in the proposed FMLD framework, we conducted two ablation experiments: (a) removing the dual-attention module in the server-side decoder, and (b) removing the DEM modality while keeping the optical imagery and hillshade data. The results across the Jiuzhaigou, Luding, and Emilia-Romagna data sets are summarized in Table 4.

Removing the attention module leads to a substantial performance drop across all three regions, with decreases of 12.7%, 7.9%, and 9.0% in F1-score for Jiuzhaigou, Luding, and Emilia-Romagna, respectively. This confirms that the dual-attention mechanism plays a crucial role in enhancing multimodal feature fusion and improving spatial consistency. In contrast, removing the DEM modality results in only slight performance reductions compared with the baseline model. The F1-score decreases by 1.16% in Jiuzhaigou, 0.97% in Luding, and 0.75% in Emilia-Romagna, indicating that the DEM contributes relatively little to the overall prediction performance. This observation aligns with the contribution analysis, suggesting feature redundancy between DEM and hillshade data, which are derived from the same elevation source. Overall, the ablation results demonstrate that the dual-attention module is essential for effective multimodal fusion, while the DEM modality provides limited additional benefit under the current feature extraction setting.

4.2.5. Communication Overhead and Training Efficiency

To further evaluate the system performance of the proposed FMLD framework, we analyzed the communication overhead and total training time across three representative study areas, including Jiuzhaigou, Luding, and Emilia–Romagna. In Jiuzhaigou, the total transmitted feature size per training round was 1119.50 MB, and the total training time was approximately 12.9 hr. In Luding, a similar transmission size of 1119.50 MB was

10.1029/2025JH000857

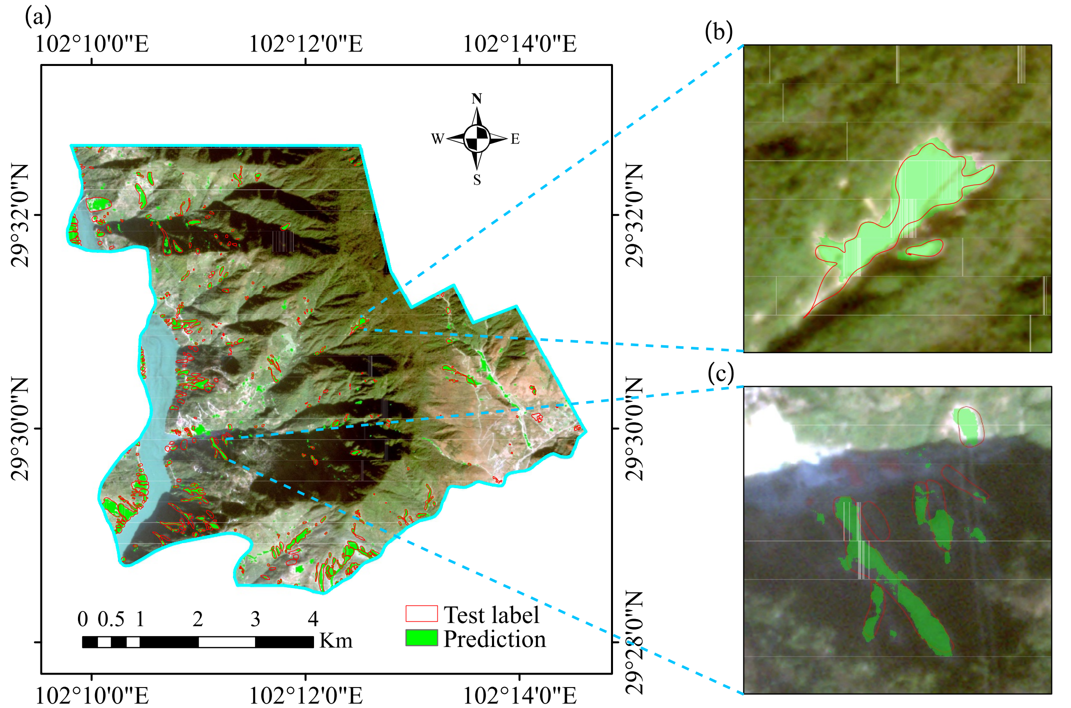

Figure 4. Visualizing the landslides identified by FMLD on the Luding data set. (a) Overall prediction results across the Luding study area. (b) Enlarged view showing accurate detection of small-scale landslides in steep terrain. (c) Enlarged view illustrating successful identification of shadowed and vegetation-covered landslides.

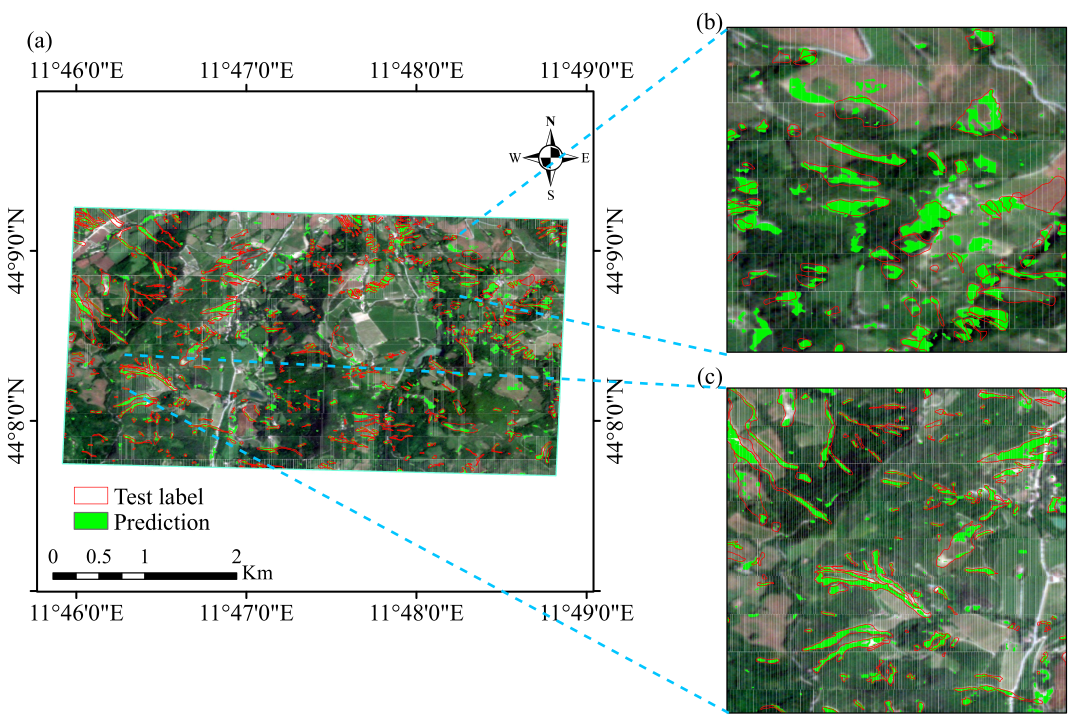

Figure 5. Visualizing the landslides identified by FMLD on the Italy data set. (a) Overall prediction results across the Emilia-Romagna study area. (b, c) Enlarged views showing scattered small-scale landslides accurately detected in hilly terrain.

Table 4
Ablation Study Results of the FMLD Framework Across Three Study Areas

Component	Study area	mIoU	Precision	Recall	Accuracy	F1
Baseline	Jiuzhaigou	0.6533	0.7484	0.7454	0.9619	0.7469
	Luding	0.7279	0.8135	0.8251	0.9716	0.8192
	Emilia-Romagna	0.6526	0.7723	0.7424	0.9275	0.7563
Without Attention Module	Jiuzhaigou	0.5430	0.5987	0.6607	0.9242	0.6201
	Luding	0.6449	0.7120	0.7787	0.9535	0.7401
	Emilia-Romagna	0.5696	0.6872	0.6501	0.9052	0.6658
Without DEM Data	Jiuzhaigou	0.6415	0.7185	0.7554	0.9566	0.7353
	Luding	0.7168	0.7983	0.8215	0.9694	0.8095
	Emilia-Romagna	0.6447	0.7591	0.7395	0.9239	0.7488

observed, with a total training time of about 5.0 hr. In Emilia-Romagna, the transmitted feature size was 91.06 MB, and the total training time was around 3.1 hr. These results demonstrate that the proposed FMLD framework maintains efficient communication and scalable training performance across diverse regional conditions, effectively balancing model accuracy and system efficiency in federated multimodal landslide detection.

5. Discussion

High-quality, large-scale landslide data sets are critical for advancing machine learning-based landslide detection models. Rapid global landslide detection demands cross-institutional collaboration and data sharing. However, concerns regarding data privacy, proprietary rights, and regulatory compliance often hinder such efforts. To address these challenges, this study proposes FMLD, a federated machine learning framework for landslide detection based on vertical federated learning. FMLD enables collaborative model training across institutions while protecting sensitive data, facilitating secure knowledge integration without raw data exchange. This section discusses the advantages of vertical over horizontal federated learning, strategies for decentralized data and label quality control, and insights from experimental results and multimodal data analysis.

5.1. Comparing Vertical With Horizontal Federated Learning for Landslide Detection

FMLD offers two primary advantages. First, geohazard research institutions typically hold diverse complementary data types, such as satellite-based optical imagery, airborne LiDAR point clouds, and ground-based deformation measurements. Traditional centralized approaches require pooling such heterogeneous data, which is often impractical due to privacy regulations or institutional policies. FMLD provides a privacy-preserving alternative, enabling multi-source, multimodal data to contribute to model training without the need for raw data exchange. Second, FMLD maintains performance comparable to centralized models. Unlike HFL approaches (Y. Yang et al., 2023; Tang et al., 2024), which often suffer from accuracy degradation due to non-IID data distributions and model aggregation issues, FMLD partitions the neural network into modality-specific subnetworks. By protecting both forward and backward propagation flows between clients and server, FMLD minimizes information leakage while maintaining model accuracy.

The distinction between VFL and HFL lies in their underlying data distribution assumptions and use cases. HFL is suited for scenarios where clients hold homogeneous data features from different geographic regions, for instance, optical satellite imagery across different administrative zones. In such cases, each client trains a local model on indetical input features, enabling collaborative learning across regions without centralizing data. In contrast, VFL is designed for scenarios in which institutions hold different but complementary modalities for the same region. This reflects real-world landslide detection constraints where institutions hold non-overlapping features (e.g., optical imagery, DEMs, or LiDAR-derived hillshade maps) for the same geographic region. VFL allows such features to be integrated securely without sharing raw data. For example, in forested mountainous terrains where canopy cover obscures landslides in optical imagery, hillshade maps can uncover hidden scarps and terrain changes (D. Li, Tang, et al., 2023). VFL enables feature-level fusion across modalities while safeguarding data privacy and ownership.

Furthermore, VFL supports modality-specific modeling, which is crucial given that different landslide types manifest uniquely across data sources. Rapid, shallow landslides are often visible in optical imagery, whereas deep-seated or slow-moving landslides require elevation-based analysis using DEMs or LiDAR. FMLD employs tailored sub-networks optimized for each modality (e.g., SegFormer for optical imagery, HRNet for DEM, and HRFormer for hillshade map) and followed by server-side feature fusion. Such modular and flexible architecture is not feasible under HFL, which requires consistent input structures and feature spaces across clients.

FMLD demonstrates how federated learning can be applied to balance data privacy and prediction accuracy in landslide detection. Its prediction accuracy is comparable to that of the corresponding centralized model, indicating that FMLD can preserve data privacy without sacrificing landslide prediction performance. Compared to HFL (Tang et al., 2024), the VFL paradigm adopted by FMLD enables a more effective trade-off between privacy protection and model accuracy. More broadly, federated learning holds significant promise for privacy-preserving modeling across a range of natural hazards, including earthquakes, debris flows, wildfires, floods, and tsunamis. These fields often involve sensitive data such as seismographs, rainfall records, remote sensing imagery, and water level measurements across research institutions, commercial entities, and government agencies. By enabling secure collaboration without raw data exchange, federated learning can alleviate privacy concerns and promote inclusive, large-scale geoscientific research. Future work could further explore the use of federated learning to integrate additional geoscience data types, such as displacement time series, GNSS measurements, and seismic waveforms, to enhance robust, privacy-aware natural hazard modeling.

5.2. Decentralized Data and Label Quality Control

A core challenge in federated learning is ensuring high data and label quality when the data are distributed and not directly accessible. To address this, FMLD incorporates a Task Initialization Protocol that coordinates the server and clients before training begins. This protocol covers several key aspects: declaration and verification of data quality, disclosure and synchronization of label sources, checks for data overlap and alignment, and formal approval and logging of initialization records. These mechanisms help ensure consistency and traceability across clients, reducing the risk of degraded model performance due to noisy or misaligned data.

In cases where a client temporarily disconnects or withdraws during training, the system maintains training stability by substituting the missing modality with placeholder feature embeddings and recording the dropout event. This ensures that model updates remain synchronized across participating clients, although a minor decrease in accuracy may occur due to incomplete multimodal information. Improving the robustness of FMLD remains an open challenge. Future work will explore adaptive aggregation strategies and dynamic reinitialization mechanisms that allow the model to automatically adjust to client dropout events, thereby enhancing the resilience and scalability of the framework under real-world federated learning environments.

5.3. Experimental Insights and Multimodal Contribution Analysis

Experimental results across three geographically and causally diverse regions, including two earthquake-triggered and one rainfall-induced landslide cases, demonstrate that FMLD achieves performance comparable to centralized models in both IoU and F1 score. These findings highlight the generalizability of the FMLD framework across different landslide types and terrain conditions while preserving the privacy of distributed data owners. In addition, the multimodal fusion strategy in FMLD significantly enhances detection accuracy. Models trained with combined optical, DEM, and hillshade inputs outperform those trained on single modalities. The modality contribution analysis shows that hillshade maps contribute the most, followed by optical imagery, while DEMs contribute the least. Quantitatively, the contributions of optical imagery, DEM, and hillshade are 17.6%, 3.2%, and 79.2% in Jiuzhaigou; 17.5%, 0.1%, and 82.4% in Luding; and 18.9%, 2.4%, and 78.7% in Emilia-Romagna.

The higher contribution of hillshade can be attributed to its ability to enhance topographic contrast and highlight geomorphological structures such as scarps, gullies, and slope breaks, which are key features for identifying landslides. Unlike optical imagery, which is affected by vegetation, lighting, and atmospheric conditions, hillshade maps emphasize intrinsic surface morphology and effectively reveal landslide traces even in shadowed or vegetated areas. Consequently, hillshade provides clearer and more stable indicators of landslide activity in mountainous terrain. In contrast, the relatively low contribution of DEMs, which is below five percent across all data sets, indicates feature redundancy between DEM and hillshade data because both represent elevation-derived

10.1029/2025JH000857

terrain morphology. To verify this, an ablation experiment was conducted by removing the DEM modality. The results showed only minor decreases in F1 score: 1.16% in Jiuzhaigou, 0.97% in Luding, and 0.75% in Emilia-Romagna. These results confirm that the DEM provides limited additional information under the current feature extraction scheme. Nevertheless, DEM data remain valuable for providing precise elevation information and can be utilized more effectively through advanced feature extraction. Future research should explore incorporating DEM-derived topographic factors such as slope, curvature, and terrain roughness to better exploit the geomorphological relevance of DEM data. These improvements could enhance both interpretability and computational efficiency in federated multimodal landslide detection.

6. Conclusions

This study proposed a new multimodal deep learning model called FMLD for landslide detection, which leveraged vertical federated learning to protect data privacy while enabling effective collaborative modeling. FMLD achieves comparable detection accuracy across diverse geographic and causal scenarios, without requiring raw data exchange between participating institutions. In this approach, the centralized landslide detection model is divided into two sub-networks: a client-side data processing sub-network, and a server-side feature fusion network. These sub-networks are organized under a Client/Server structure. Each client utilizes a specific type of landslide data to train a data processing sub-network (e.g., optical imagery, DEMs, hillshade maps). The server then uses the feature fusion network to aggregate the feature maps of all the clients. Communication between the server and clients is limited to the exchange of neural network parameters and gradients, ensuring that the original landslide data of each data owner are never transmitted to others. Consequently, the privacy of all data owners is effectively protected. FMLD was evaluated in three distinct study areas, namely Jiuzhaigou and Luding County in China, and Emilia-Romagna in Italy. Experimental results demonstrate that FMLD achieves performance comparable to centralized models and consistently outperforms single-modality baselines. The results further confirm that FMLD successfully exploits the complementary information among different modalities, particularly between optical and topographic data, leading to improved detection accuracy in complex terrain conditions. Visualization analysis further confirms that FMLD effectively detects landslide occurrences across varied terrains and data types. Future work will focus on extending this framework to incorporate additional hazard-related modalities, enhancing robustness under real-world deployment conditions, and reducing communication overhead to further improve system efficiency. In addition, future studies will explore the integration of multi-temporal optical imagery and ancillary geospatial data sets, such as land use and vegetation maps, to better distinguish between actual landslides and visually similar non-landslide features, as well as improving DEM utilization through derived topographic factors and optimized feature extraction.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The source code and data supporting this research are available at (Tang et al., 2026). Due to restrictions associated with required NDAs and licensing agreements, the data are provided at reduced resolution.

Acknowledgments

This work was supported by the National Science Fund for Distinguished Young Scholars (No. 42125702); the Sichuan Science and Technology Program (No. 2026NSFSC0439, 2024ZYD0140, 2024JDHJ0038, 2024ZDZX0020); the Opening fund of State Key Laboratory of Geohazard Prevention and Geoenvironment Protection (Chengdu University of Technology) (No. SKLGP2025K025); the Academic Degree and Postgraduate Education Reform Project of Sichuan Province (No. YJGXM24-C036); and the National Natural Science Foundation of China (Grant 52579103); The co-authors from UNIPD were funded by "The Geosciences for Sustainable Developmen" project (Budget Ministero dell'Universitá e della Ricerca-Dipartimenti di Eccellenza 2023–2027 [CUP C93C23002690001]) of the Department of Geosciences, University of Padova.

References

Atarashi, K., & Ishihata, M. (2021). Vertical federated learning for higher-order factorization machines. In *Pacific-Asia Conference on knowledge discovery and data mining* (pp. 346–357). https://doi.org/10.1007/978-3-030-75765-6_28

Berti, M., Pizziolo, M., Scaroni, M., Generali, M., Critelli, V., Mulas, M., et al. (2025). RER2023: The landslide inventory dataset of the May 2023 Emilia-Romagna meteorological event. *Earth System Science Data*, *17*(3), 1055–1074. https://doi.org/10.5194/essd-17-1055-2025

Cai, J., Zhang, L., Dong, J., Guo, J., Wang, Y., & Liao, M. (2023a). Automatic identification of active landslides over wide areas from time-series InSAR measurements using faster RCNN. *International Journal of Applied Earth Observation and Geoinformation*, *124*, 103516. https://doi.org/10.1016/j.jag.2023.103516

Cai, W., Gao, M., Ding, Y., Ning, X., Bai, X., & Qian, P. (2023b). Stereo attention cross-decoupling fusion-guided federated neural learning for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, *61*, 1–16. https://doi.org/10.1109/TGRS.2023.3320044

Casagli, N., Intrieri, E., Tofani, V., Gigli, G., & Raspini, F. (2023). Landslide detection, monitoring and prediction with remote-sensing techniques. *Nature Reviews Earth & Environment*, *4*(1), 51–64. https://doi.org/10.1038/s43017-022-00373-x

Chen, W., Li, X., Wang, Y., Chen, G., & Liu, S. (2014). Forested landslide detection using LiDAR data and the random forest algorithm: A case study of the three gorges, China. *Remote Sensing of Environment*, *152*, 291–301. https://doi.org/10.1016/j.rse.2014.07.004

Coca, M., Neagoe, I., & Datcu, M. (2020). Physically meaningful dictionaries for eo crowdsourcing: A ML for blockchain architecture. In *IEEE international geoscience and remote sensing symposium* (pp. 3688–3691). https://doi.org/10.1109/IGARSS39084.2020.9324361

Dai, L., Fan, X., Wang, X., Fang, C., Zou, C., Tang, X., et al. (2023). Coseismic landslides triggered by the 2022 Luding Ms 6.8 earthquake, China. *Landslides*, *20*(6), 1277–1292. https://doi.org/10.1007/s10346-023-02061-3

Dong, A., Dou, J., Li, C., Chen, Z., Ji, J., Xing, K., et al. (2024). Accelerating cross-scene co-seismic landslide detection through progressive transfer learning and lightweight deep learning strategies. *IEEE Transactions on Geoscience and Remote Sensing*, *62*, 4410213. https://doi.org/10.1109/TGRS.2024.3424680

Elmoulat, M., Debauche, O., Mahmoudi, S., Mahmoudi, S. A., Guttadauria, A., Manneback, P., & Lebeau, F. (2021). Towards landslides early warning system with fog-edge computing and artificial intelligence. *International Journal of Ubiquitous Systems and Pervasive Networks*, *15*(2), 11–17. https://doi.org/10.5383/JUSPN.15.02.002

Fadlullah, Z. M., & Kato, N. (2021). On smart IoT remote sensing over integrated terrestrial-aerial-space networks: An asynchronous federated learning approach. *IEEE Network*, *35*(5), 129–135. https://doi.org/10.1109/MNET.101.2100125

Fan, X., Scaringi, G., Domènech, G., Yang, F., Guo, X., Dai, L., et al. (2019). Two multi-temporal datasets that track the enhanced landsliding after the 2008 Wenchuan earthquake. *Earth System Science Data*, *11*(1), 35–55. https://doi.org/10.5194/essd-11-35-2019

Fan, X., Scaringi, G., Xu, Q., Zhan, W., Dai, L., Li, Y., et al. (2018). Coseismic landslides triggered by the 8th August 2017 Ms 7.0 Jiuzhaigou earthquake (Sichuan, China): Factors controlling their spatial distribution and implications for the seismogenic blind fault identification. *Landslides*, *15*(5), 967–983. https://doi.org/10.1007/s10346-018-0960-x

Fang, C., Fan, X., Wang, X., Nava, L., Zhong, H., Dong, X., et al. (2024). A globally distributed dataset of coseismic landslide mapping via multi-source high-resolution remote sensing images. *Earth System Science Data Discussions, 2024*, 1–42. https://doi.org/10.5194/essd-2024-239

Fang, C., Fan, X., Zhong, H., Lombardo, L., Tanyas, H., & Wang, X. (2022). A novel historical landslide detection approach based on LiDAR and lightweight attention U-Net. *Remote Sensing*, *14*(17), 4357. https://doi.org/10.3390/rs14174357

Farmakis, I., DiFrancesco, P.-M., Hutchinson, D. J., & Vlachopoulos, N. (2022). Rockfall detection using LiDAR and deep learning. *Engineering Geology*, *309*, 106836. https://doi.org/10.1016/j.enggeo.2022.106836

Fu, C., Zhang, X., Ji, S., Chen, J., Wu, J., Guo, S., et al. (2022). Label inference attacks against vertical federated learning. In *USENIX security symposium* (pp. 1397–1414). Retrieved from https://www.usenix.org/conference/usenixsecurity22/presentation/fu-chong

Gao, Y., Liu, L., Hu, B., Lei, T., & Ma, H. (2020). Federated region-learning for environment sensing in edge computing system. *IEEE Transactions on Network Science and Engineering*, *7*(4), 2192–2204. https://doi.org/10.1109/TNSE.2020.3016035

Ghorbanzadeh, O., Xu, Y., Ghamisi, P., Kopp, M., & Kreil, D. (2022a). Landslide4Sense: Reference benchmark data and deep learning models for landslide detection. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–17. https://doi.org/10.1109/TGRS.2022.3215209

Ghorbanzadeh, O., Xu, Y., Zhao, H., Wang, J., Zhong, Y., Zhao, D., et al. (2022b). The outcome of the 2022 Landslide4Sense competition: Advanced landslide detection from multisource satellite imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, 9927–9942. https://doi.org/10.1109/JSTARS.2022.3220845

Hardy, S., Henecka, W., Ivey-Law, H., Nock, R., Patrini, G., Smith, G., & Thorne, B. (2017). Private federated learning on vertically partitioned data via entity resolution and additively homomorphic encryption. *arXiv preprint arXiv:1711.10677*. https://doi.org/10.48550/arXiv.1711.10677

He, Y., Kang, Y., Zhao, X., Luo, J., Fan, L., Han, Y., & Yang, Q. (2024). A hybrid self-supervised learning framework for vertical federated learning. *IEEE Transactions on Big Data*, *11*(5), 2210–2223. https://doi.org/10.1109/TBDATA.2024.3403386

Huang, Y., Xie, C., Li, T., Xu, C., He, X., Shao, X., et al. (2023). An open-accessed inventory of landslides triggered by the MS 6.8 Luding earthquake, China on September 5, 2022. *Earthquake Research Advances*, *3*(1), 100181. https://doi.org/10.1016/j.eqrea.2022.100181

Joshi, A., Agarwal, S., Kanungo, D. P., & Panigrahi, R. K. (2023). Integration of Edge–AI into IoT–Cloud architecture for landslide monitoring and prediction. *IEEE Transactions on Industrial Informatics*, *20*(3), 4246–4258. https://doi.org/10.1109/TII.2023.3319671

Kang, Y., He, Y., Luo, J., Fan, T., Liu, Y., & Yang, Q. (2022). Privacy-preserving federated adversarial domain adaptation over feature groups for interpretability. *IEEE Transactions on Big Data*, *10*(6), 879–890. https://doi.org/10.1109/TBDATA.2022.3188292

Lee, W. (2022). Federated reinforcement learning-based UAV swarm system for aerial remote sensing. *Wireless Communications and Mobile Computing*, *2022*(1), 4327380. https://doi.org/10.1155/2022/4327380

Lei, H., Wang, X., Hou, H., Su, L., Yu, D., & Wang, H. (2018). The earthquake in Jiuzhaigou county of Northern Sichuan, China on August 8, 2017. *Natural Hazards*, *90*(2), 1021–1030. https://doi.org/10.1007/s11069-017-3064-3

Li, D., Tang, X., Tu, Z., Fang, C., & Ju, Y. (2023a). Automatic detection of forested landslides: A case study in Jiuzhaigou county, China. *Remote Sensing*, *15*(15), 3850. https://doi.org/10.3390/rs15153850

Li, D., Xie, W., Li, Y., & Fang, L. (2023b). Fedfusion: Manifold driven federated learning for multi-satellite and multi-modality fusion. *IEEE Transactions on Geoscience and Remote Sensing*, *62*, 5500813. https://doi.org/10.1109/TGRS.2023.3339522

Li, W., Fu, Y., Fan, S., Xin, M., & Bai, H. (2023c). DCI-PGCN: Dual channel interaction portable graph convolutional network for landslide detection. *IEEE Transactions on Geoscience and Remote Sensing*, *61*, 1–16. https://doi.org/10.1109/TGRS.2023.3273623

Liu, X., Peng, Y., Lu, Z., Li, W., Yu, J., Ge, D., & Xiang, W. (2023). Feature-fusion segmentation network for landslide detection using high-resolution remote sensing images and digital elevation model data. *IEEE Transactions on Geoscience and Remote Sensing*, *61*, 1–14. https://doi.org/10.1109/TGRS.2022.3233637

Liu, Y., Kang, Y., Zou, T., Pu, Y., He, Y., Ye, X., et al. (2024). Vertical federated learning: Concepts, advances, and challenges. *IEEE Transactions on Knowledge and Data Engineering*, *36*(7), 3615–3634. https://doi.org/10.1109/TKDE.2024.3352628

Liu, Y., Nie, J., Li, X., Ahmed, S. H., Lim, W. Y. B., & Miao, C. (2020). Federated learning in the sky: Aerial-ground air quality sensing framework with UAV swarms. *IEEE Internet of Things Journal*, *8*(12), 9827–9837. https://doi.org/10.1109/JIOT.2020.3021006

Lu, J., Pan, B., Seid, A. M., Li, B., Hu, G., & Wan, S. (2022). Truthful incentive mechanism design via internalizing externalities and LP relaxation for vertical federated learning. *IEEE Transactions on Computational Social Systems*, *10*(6), 2909–2923. https://doi.org/10.1109/TCSS.2022.3227270

Lu, Z., Peng, Y., Li, W., Yu, J., Ge, D., Han, L., & Xiang, W. (2023). An iterative classification and semantic segmentation network for old landslide detection using high-resolution remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, *61*, 1–13. https://doi.org/10.1109/TGRS.2023.3313586

Luo, X., Wu, Y., Xiao, X., & Ooi, B. C. (2021). Feature inference attack on model predictions in vertical federated learning. In *IEEE International Conference on data engineering* (pp. 181–192). https://doi.org/10.1109/ICDE51399.2021.00023

Lv, P., Ma, L., Li, Q., & Du, F. (2023). ShapeFormer: A shape-enhanced vision transformer model for optical remote sensing image landslide detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *16*, 2681–2689. https://doi.org/10.1109/JSTARS.2023.3253769

McMahan, B., Moore, E., Ramage, D., Hampson, S., & yArcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics* (pp. 1273–1282). Retrieved from https://proceedings.mlr.press/v54/mcmahan17a.html

Meena, S. R., Nava, L., Bhuyan, K., Puliero, S., Soares, L. P., Dias, H. C., et al. (2023). HR-GLDD: A globally distributed dataset using generalized deep learning (DL) for rapid landslide mapping on high-resolution (HR) satellite imagery. *Earth System Science Data*, *15*(7), 3283–3298. https://doi.org/10.5194/essd-15-3283-2023

Mohan, A., Singh, A. K., Kumar, B., & Dwivedi, R. (2021). Review on remote sensing methods for landslide detection using machine and deep learning. *Transactions on Emerging Telecommunications Technologies*, *32*(7), e3998. https://doi.org/10.1002/ett.3998

Nava, L., Bhuyan, K., Meena, S. R., Monserrat, O., & Catani, F. (2022). Rapid mapping of landslides on SAR data by attention U-Net. *Remote Sensing*, *14*(6), 1449. https://doi.org/10.3390/rs14061449

Nava, L., Mondini, A. C., Bhuyan, K., Fang, C., Monserrat, O., Novellino, A., & Catani, F. (2024). Sentinel-1 SAR-based globally distributed landslide detection by deep neural networks. *EarthArXiv Preprint*. https://doi.org/10.31223/X59D6M

Qi, P., Chiaro, D., & Piccialli, F. (2023). FL-FD: Federated learning-based fall detection with multimodal data fusion. *Information Fusion*, *99*, 101890. https://doi.org/10.1016/j.inffus.2023.101890

Saadati, Y., Imteaj, A., & Amini, M. H. (2024). Vertical federated learning: Principles, applications, and future frontiers. In *Distributed machine learning and computing: Theory and applications* (pp. 111–127). Springer. https://doi.org/10.1007/978-3-031-57567-9_5

Shi, Y., Xia, S., Zhou, Y., Mao, Y., Jiang, C., & Tao, M. (2023). Vertical federated learning over cloud-ran: Convergence analysis and system optimization. *IEEE Transactions on Wireless Communications*, *23*(2), 1327–1342. https://doi.org/10.1109/TWC.2023.3288122

Tam, P., Math, S., Nam, C., & Kim, S. (2021). Adaptive resource optimized edge federated learning in real-time image sensing classifications. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 10929–10940. https://doi.org/10.1109/JSTARS.2021.3120724

Tang, X., Dai, Y., Sun, P., & Meng, S. (2018). Interaction-based feature selection using Factorial design. *Neurocomputing*, *281*, 47–54. https://doi.org/10.1016/j.neucom.2017.11.058

Tang, X., Liu, M., Zhong, H., Ju, Y., Li, W., & Xu, Q. (2021). MILL: Channel attention–based deep multiple instance learning for landslide recognition. ACM transactions on multimedia computing. *Communications, and Applications*, *17*(2s), 1–11. https://doi.org/10.1145/3454009

Tang, X., Lu, Z., Fan, X., Li, D., Yuan, X., Li, H., et al. (2026). Release v1.0.0: FMLD: Vertical federated learning for multi-modal landslide detection. *Github*. https://doi.org/10.5281/zenodo.18345539

Tang, X., Lu, Z., Fan, X., Yan, X., Yuan, X., Li, D., et al. (2025). Mamba for landslide detection: A lightweight model for mapping landslides with very high-resolution images. *IEEE Transactions on Geoscience and Remote Sensing*, *63*, 5637117. https://doi.org/10.1109/TGRS.2025.3598446

Tang, X., Tu, Z., Wang, Y., Liu, M., Li, D., & Fan, X. (2022). Automatic detection of coseismic landslides using a new transformer method. *Remote Sensing*, *14*(12), 2884. https://doi.org/10.3390/rs14122884

Tang, X., Yan, X., Catani, F., Liu, X., Lu, Z., Wang, Y., et al. (2024). FedLD: Federated learning for privacy-preserving collaborative landslide detection. *IEEE Geoscience and Remote Sensing Letters*, *21*, 8003105. https://doi.org/10.1109/LGRS.2024.3437743

Tehrani, F. S., Calvello, M., Liu, Z., Zhang, L., & Lacasse, S. (2022). Machine learning and landslide studies: Recent advances and applications. *Natural Hazards*, *114*(2), 1197–1245. https://doi.org/10.1007/s11069-022-05423-7

Tehseen, R., Farooq, M. S., & Abid, A. (2021). A framework for the prediction of earthquake using federated learning. *PeerJ Computer Science*, *7*, e540. https://doi.org/10.7717/peerj-cs.540

Teimoori, Z., Yassine, A., & Hossain, M. S. (2022). A secure cloudlet-based charging station recommendation for electric vehicles empowered by federated learning. *IEEE Transactions on Industrial Informatics*, *18*(9), 6464–6473. https://doi.org/10.1109/TII.2022.3148997

Thapa, C., Arachchige, P. C. M., Camtepe, S., & Sun, L. (2022). SplitFeD: When federated learning meets split learning. In *Proceedings of the AAAI Conference on artificial intelligence* (Vol. 36, pp. 8485–8493). https://doi.org/10.1609/aaai.v36i8.20825

Wang, G., Gu, B., Zhang, Q., Li, X., Wang, B., & Ling, C. X. (2024). A unified solution for privacy and communication efficiency in vertical federated learning. Retrieved from https://openreview.net/forum?id=AYiRHZirD2

Wang, J., Sun, K., Cheng, T., Jiang, B., Deng, C., Zhao, Y., et al. (2020). Deep high-resolution representation learning for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *43*(10), 3349–3364. https://doi.org/10.1109/TPAMI.2020.2983686

Wei, R., Ye, C., Ge, Y., Li, Y., & Li, J. (2023). Dynamic graph attention networks for point cloud landslide segmentation. *International Journal of Applied Earth Observation and Geoinformation*, *124*, 103542. https://doi.org/10.1016/j.jag.2023.103542

Xu, Y., Bai, T., Yu, W., Chang, S., Atkinson, P. M., & Ghamisi, P. (2023). AI security for geoscience and remote sensing: Challenges and future trends. *IEEE Geoscience and Remote Sensing Magazine*, *11*(2), 60–85. https://doi.org/10.1109/MGRS.2023.3272825

Xu, Y., Ouyang, C., Xu, Q., Wang, D., Zhao, B., & Luo, Y. (2024). CAS landslide dataset: A large-scale and multisensor dataset for deep learning-based landslide detection. *Scientific Data*, *11*(1), 12. https://doi.org/10.1038/s41597-023-02847-z

Yan, Y., Wang, H., Huang, Y., He, N., Zhu, L., Xu, Y., et al. (2024). Cross-modal vertical federated learning for MRI reconstruction. *IEEE Journal of Biomedical and Health Informatics*, *28*(11), 6384–6394. https://doi.org/10.1109/JBHI.2024.3360720

Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, *10*(2), 1–19. https://doi.org/10.1145/3298981

Yang, Y., Lu, Y., & Mei, G. (2023). A federated learning based approach for predicting landslide displacement considering data security. *Future Generation Computer Systems*, *149*, 184–199. https://doi.org/10.1016/j.future.2023.07.021

Yin, X., Zhu, Y., & Hu, J. (2021). A comprehensive survey of privacy-preserving federated learning: A taxonomy, review, and future directions. *ACM Computing Surveys*, *54*(6), 1–36. https://doi.org/10.1145/346042

Yuan, Y., Fu, R., Huang, L., Lin, W., Zhang, C., Chen, X., & Wang, J. (2021). HRFormer: High-resolution transformer for dense prediction. In *International Conference on neural information processing systems* (pp. 7281–7293). https://doi.org/10.5555/3540261.3540818

Zhang, X., Zhang, B., Yu, W., & Kang, X. (2023a). Federated deep learning with prototype matching for object extraction from very-high-resolution remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, *61*, 1–16. https://doi.org/10.1109/TGRS.2023.3244136

Zhang, Z., Ma, X., & Ma, J. (2023b). Local differential privacy based membership-privacy-preserving federated learning for deep-learning-driven remote sensing. *Remote Sensing*, *15*(20), 5050. https://doi.org/10.3390/rs15205050

Zhu, J., Wu, J., Bashir, A. K., Pan, Q., & Yang, W. (2023). Privacy-preserving federated learning of remote sensing image classification with dishonest majority. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *16*, 4685–4698. https://doi.org/10.1109/JSTARS.2023.3276781