Earth Systems and Environment (2024) 8:1387–1398
https://doi.org/10.1007/s41748-024-00419-8

ORIGINAL ARTICLE

Evaluation of Various Deep Learning Algorithms for Landslide and Sinkhole Detection from UAV Imagery in a Semi-arid Environment

Narges Kariminejad[1] · Hejar Shahabi[2] · Omid Ghorbanzadeh[3] · Vahid Shafaie[4] · Mohsen Hosseinalizadeh[5] · Saied Homayouni[2] · Hamid Reza Pourghasemi[6]

Received: 4 March 2024 / Revised: 11 June 2024 / Accepted: 12 June 2024 / Published online: 27 June 2024

Abstract

Sinkholes and landslides occur due to soil collapse in different slope types, often triggered by heavy rainfall, presenting challenges in the semi-arid Golestan province, Iran. This study primarily focuses on the detection of these phenomena. Recent advancements in unmanned aerial vehicle (UAV) image acquisition and the incorporation of deep learning (DL) algorithms have enabled the creation of semi-automated methods for highly detailed soil landform detection across large areas. In this study, we explored the efficacy of six state-of-the-art deep learning segmentation algorithms—DeepLab-v3+, Link-Net, MA-Net, PSP-Net, ResU-Net, and SQ-Net—applied to UAV-derived datasets for mapping landslides and sinkholes. Our most promising outcomes demonstrated the successful mapping of landslides with an F1-Score of 0.95% and sinkholes with an F1-Score of 89% in a challenging environment. ResUNet exhibited an outstanding Precision of 0.97 and Recall of 0.92, culminating in the highest F1-Score of 0.95, indicating the best landslide detection model. MA-Net and SQ-Net resulted in the highest F1-Score for sinkhole detection. Our study underscores the significant impact of DL segmentation algorithm selection on the accuracy of landslide and sinkhole detection tasks. By leveraging DL segmentation algorithms, the accuracy of both landslide and sinkhole detection tasks can be significantly improved, promoting better hazard management and enhancing the safety of the affected areas.

Keywords Deep learning · Landslide detection · Sinkholes · UAV

1 Introduction

In Iran, rainfall-induced events such as landslides and sinkholes are frequent and have substantial impacts. Recent studies (Kariminejad et al. 2022) have revealed that landslides cause extensive damage to infrastructure and natural landscapes, influencing land-use and migration patterns. Sinkholes are prevalent in various global environments (Bernatek-Jakiel and Poesen 2018) and significantly affect landscape hydrology, contributing to landscape transformations and degradation (Jones and Crane 1984). This erosion type, linked to drainage densities (Bernatek 2015), also leads to channel expansion (Higgins et al. 1990). Sinkholes and landslides were chosen as the two subjects in this study because of their relevance and significance in the context of the semi-arid environment of the Golestan Province of Iran. Both sinkholes and landslides are geological hazards that can occur in various regions worldwide; however, their specific characteristics and causes differ. The analysis of these two landforms typically serves a dual purpose: to identify

✉ Omid Ghorbanzadeh
omid.ghorbanzadeh@boku.ac.at

1 Department of Natural Resources and Environmental Engineering, College of Agriculture, Shiraz University, Shiraz, Iran

2 Center Eau Terre Environnement, Institut National de la Recherche Scientifique (INRS), Quebec City, QC G1K 9A9, Canada

3 Institute of Geomatics, University of Natural Resources and Life Sciences (BOKU), Vienna, Peter-Jordan Strasse 82, 1190 Vienna, Austria

4 Department of Structural and Geotechnical Engineering, Széchenyi István University, Gyor 9026, Hungary

5 Department of Arid Zone Management, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran

6 Department of Soil Science, College of Agriculture, Shiraz University, Shiraz, Iran

occurrences and to create comprehensive inventories detailing affected areas (Naboureh et al. 2023a). These inventories encompass various details such as the spatial distribution of landslides and sinkholes, volume of slope failures, and specifics of each incident (Petschko et al. 2013). Successful landslide and sinkhole detection depends on the examination of existing hazard inventories, their geographic coverage, and their correlation, which helps identify other areas susceptible to these natural events (Agrawal et al. 2017; Chen et al. 2021). The consistency of the outcomes was inextricably connected to the accuracy of the input data.The development of improved remote sensing (RS) technology has greatly aided the monitoring of surface changes, notably landscape changes caused by extreme climatic events, human activities, and seismic activity (Naboureh et al. 2023b). RS is crucial because it enables large-scale mapping of slope failures (Fernández et al. 2016) and provides current data with reasonable spatial resolution (Moharrami et al. 2020). The UAVs have emerged as a great method for obtaining cost-effective, up-to-date, and extremely precise field datasets (Ghorbanzadeh et al. 2019; Zhang et al. 2019). The use of UAVs in circumstances involving natural catastrophes and emergency responses is gaining momentum (Watson and Kargel 2019; Windrim et al. 2019). Compared to traditional tools, UAVs have the potential to generate data with superior spatial resolution (Brovkina et al. 2018). The same hilly loess region was the subject of a recent study by Kariminejad et al. (2022) using UAV technology. They found that sinkholes were the most dangerous natural hazard, followed by landslides and headcuts.

While UAVs have been used in a number of studies to gather information for landslide inventories and piping/headcut monitoring (Yang et al. 2015; Fernández et al. 2016), this investigation has not yet fully explored the use of UAVs for the analysis and classification of remote sensing imagery to identify sinkholes and landslides. Pixel-based and object-based techniques are the categories into which the methods used for image analysis and classification fall. These techniques are then enhanced utilizing various machine- and deep-learning algorithms for the purpose of detecting natural hazards (Tavakkoli Piralilou et al. 2019).

Over the past decade, Convolutional Neural Networks (CNNs) have consistently demonstrated significant success in RS image classification tasks. The pioneering utilization of machine learning and CNNs in landslide detection surfaced in early 2019 through a groundbreaking study by Ghorbanzadeh et al. (2019). This pioneering study introduced two distinct CNNs meticulously designed to detect landslides using very high-resolution (VHR) imagery. Their innovative CNN-based approach yielded remarkable results, surpassing the performance of machine learning models, including the "random forest", "support vector machine", and "artificial neural network" models. It is crucial to emphasize that this study marked an inaugural endeavor to employ CNNs for landslide detection, establishing a significant milestone in this field. In the case of sinkholes, Lee et al. (2016) harnessed light CNNs to map sinkholes within urban settings using thermal far-infrared (FIR) imagery. Conversely, Hoai et al. (2019) introduced a sinkhole-tracking methodology that employed CNN transfer learning on FIR imagery. Nevertheless, it remains uncertain whether such techniques are applicable in highly complex semi-arid regions, where conventional RGB optical imagery is typically more accessible. Regarding landslides, the utilization of VHR-UAV imagery in combination with DL algorithms for automated mapping remains relatively scarce. For instance, Ghorbanzadeh et al. (2019) executed a noteworthy study to train diverse CNN architectures to map landslides within densely forested Himalayan terrains. Karantanellis et al. (2021) conducted a study that compared various machine learning models and object-based image analysis techniques to accurately outline the boundaries of two moving landslides using UAV images. It should be emphasized that this research focused exclusively on the delineation of two significant landslides.

The literature indicates a limited number of studies dedicated to investigating the utilization of UAVs in conjunction with DL techniques for automated mapping of landslides and sinkholes. Remarkably, none of these studies has undertaken evaluations within the context of a semi-arid environment. The adoption of a DL framework represents an innovative and alternative approach to detect landslides and sinkholes, diverging from the conventional reliance solely on UAV technology. Overall, the importance of this research lies in its contribution to the field of hazard detection and management, particularly in semi-arid environments. By evaluating deep-learning algorithms and leveraging UAV imagery, this study offers insights and methodologies that can significantly enhance the accuracy and effectiveness of sinkhole and landslide detection, ultimately improving safety and resilience in vulnerable areas.

Additionally, it is worth noting that the application of such methodologies in semi-arid regions has been notably infrequent, largely due to challenges associated with extracting relevant data arising from subtle differences in spectral characteristics. By incorporating spectral information derived from UAV-generated imagery in combination with comprehensive slope datasets, we have effectively demonstrated the efficacy of DL-based automated methods. Our investigation also elucidated the advantages and limitations associated with the utilization of topographic datasets, particularly emphasizing the slope dataset, in the context of sinkhole and landslide detection.

The study area was initiated by a profound concern about the remarkable and multifaceted characteristics of sinkholes, along with the associated challenges they pose. These challenges include issues such as limited accessibility

to agricultural machinery, entrapment of cattle, and intricate hydrological responses within loess landscapes. It is noteworthy that some of these sinkholes, in conjunction with the subsequent development of gully heads, can reach depths of over 10 m, leading to unfortunate incidents involving cattle. Furthermore, the escalating numerical and dimensional expansion of sinkholes, coupled with the retreat and growth of gully heads, has resulted in substantial land degradation and prompted increased residential migration towards Kalaleh city and its neighboring towns. Rapid and severe land degradation caused by a combination of climatic and anthropogenic factors has led to the fragmentation of croplands into less fertile, highly erodible, and degraded plots. Notably, torrential rainfall in nearby regions has further exacerbated the situation, resulting in a significant surge in the number of sinkholes and gully heads. This underscores the critical importance of understanding landscape dynamics and their evolution. In the years 2020 and 2021, the study area continued to witness an increase in regions susceptible to various hazards, including sinkholes, landslides, and even compound events.

Furthermore, given that this form of soil erosion constitutes the primary source of significant soil degradation issues, such as the initiation and expansion of gullies, it is imperative to emphasize the significance of implementing advanced technologies, specifically UAVs, and deep learning methodologies. This recommendation carries substantial weight because it promises to yield invaluable insights into these phenomena at a more refined level of detail. In addition, prudent management of excessive grazing is critical as a key anthropogenic factor, particularly in regions characterized by loess-driven soils. This is of paramount concern because of the inherent fertility and vulnerability of these landscapes, which render them highly susceptible to a range of natural hazards including the formation of sinkholes and the occurrence of landslides.

2 Study Area and Materials

The study area is situated within the Golestan province in northeastern Iran, spanning from approximately 55° 39′ 30° E to 55° 41′ 30° E longitude and 37″ 36′ 30° N to 37″ 38′ 30° N latitude. Encompassing 515 ha, it primarily experiences a semiarid climate. The altitudes within this area range from 211 to 549 m above sea level. The prevailing soil texture is "silt loam," and the entire region is covered by loess deposits. In Fig. 1, we present the location and orthophoto image of the study area, along with visual examples of sinkholes and landslides. This region is characterized by loess-based soils and receives an average annual rainfall of 261 mm. The altitudes within this area range from 211 to 549 m above sea level. The prevailing soil texture is "silt loam," and the entire region is covered by loess deposits. In Fig. 1, we present the location and orthophoto image of the study area, along with visual examples of sinkholes and landslides. Additionally, it's

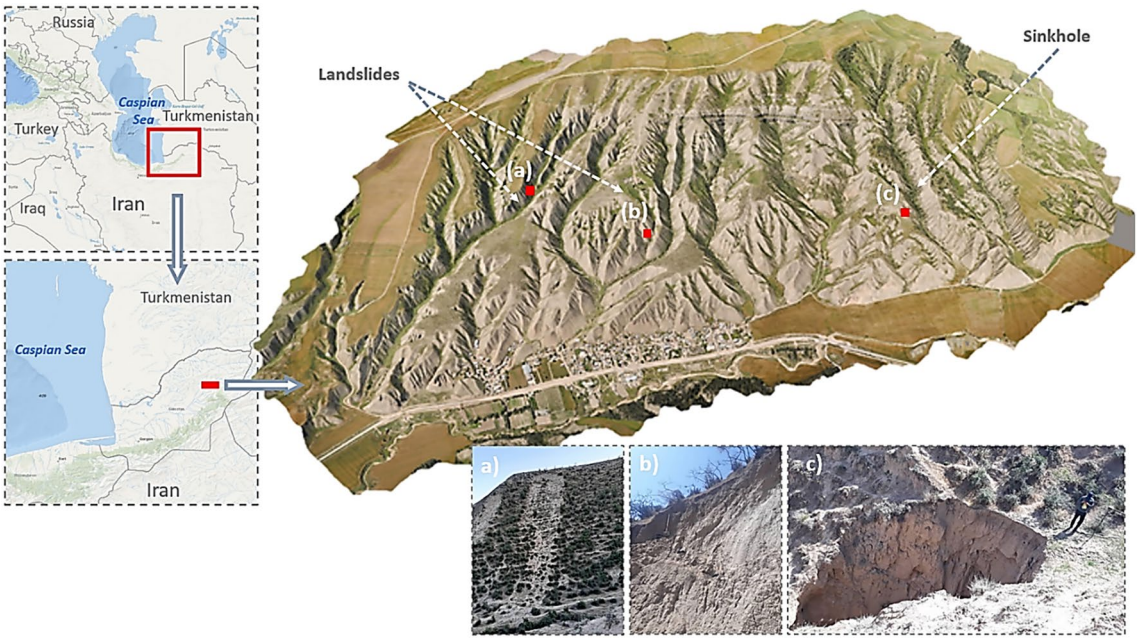

Fig. 1 The orthophoto image of the study area. Some examples of the studied sinkholes and landslides (the images are taken by Kariminejad et al. 2022)

worth noting that this region holds great significance for local ranchers and farmers due to the influence of soil piping erosion on grazing practices.

This region predominantly consists of rangeland with older sinkholes surrounded by woody vegetation, including species such as pomegranate, Ailanthus, Paliurus spina, and Lycium shawii. Consequently, sinkholes exhibit dual behavior, serving as sites for both land degradation and natural flowerpots, which provide suitable conditions for self-growing trees. The manifestation of these behaviors alternates between young and old sinkholes. Therefore, sinkhole biodiversity plays a crucial role and can be integrated into water/soil conservation practices. Therefore, the identification, monitoring, management, and bioengineering of these sinkholes are of paramount importance. Prior research conducted on the Iranian Loess Plateau (Kariminejad et al. 2022) highlights the concentration of gully heads and sinkholes in rangeland areas, with landowners and agricultural practices effectively managing and controlling them in croplands.

3 Methodology

The combination of computer vision and artificial intelligence has resulted in significant progress, especially in the areas of identifying sinkholes and landslides. Equipped with a carefully assembled dataset, we embarked on an analysis that employed six specific deep learning techniques. A comprehensive exploration of these methods is provided in subsequent sections.

3.1 Field Survay and Dataset Creation

Using UAV-based photogrammetric capture, the locations of sinkholes and landslides could be accurately determined. During the examination, sites exhibiting optical and morphological erosion and deposition were determined to have sinkholes and landslides (see Fig. 2). Sensefly eBee x, a fixed unmanned aerial vehicle, was equipped with a Sensefly Aeria X camera type with an 18.50 mm focal length. In order to guarantee an average ground sampling distance of 5.0 cm, acquisitions were planned accordingly. The orthomosaic's

Fig. 2 Enlarged examples of the studied sinkholes (**a**, **b**) and landslides from the study area and field survay perocedure (the field survay was done by Kariminejad et al. 2022)

final pixel size was 5.0 cm/pixel. The flight was organized by a skilled UAV operator using the SenseFly eMotion flight-control software. There was also the GPS RTK adjustment. The UAV trajectories were built with an 85% fore-and-aft overlap and a 70% lateral overlap. The average flying height is 220 m above the surface, and the aerial images collected during the research were analyzed using the Pix4Dmapper software.

We selected a large tile from the research area to create a composite image and corresponding mask for the dataset. The tile measured 512 × 512 pixels and contained three, four, or seven bands depending on the band arrangement. In the case of the landslide dataset, there were 2,064,288 pixels within landslides and 487,868,622 pixels in the background class, resulting in a ratio of 236. The tile and mask were divided into non-overlapping 512 × 512-pixel patches, totaling 6508 patches. For the sinkhole dataset, there were 2,096,241 pixels in the sinkholes class and 1,864,415,983 pixels in the background class, with a ratio of 899. Only images containing at least one target pixel (132 landslides and 203 sinkholes) were saved to balance class discrepancies and maintain a consistent quantity of target pixels. We allocated 70% of all patches for training purposes, reserving the remaining 30% for evaluating their performance. Before distributing the patches into the training, validation, and test sets, they were randomly shuffled. To prevent bias in the calibration process, the patches were sampled without overlapping. Subsequently, we standardized the pixel values within each patch to a range of 0 to 1. This study employed the RGBS band combination, which includes optical data and slope parameters, for detecting landslides and sinkholes. The decision to use this combination was based on the proven effectiveness of combining optical data with slope data for geohazard detection.

3.2 Models Architecture and Training Strategy

For this study, the leveraged DL models were derived from the "Segmentation models Pytorch" library (https://github.com/qubvel/segmentation_models.pytorch). To construct the dataset, 256 × 256 pixel image patches were extracted using GeoPatch, focusing on regions containing labeled data. During the data preparation phase, a dam slope layer was integrated into the images. Consequently, the subsequent raster layer encompasses three multispectral bands and a distinct slope layer, omitting the last layer that is devoid of substantive information.

In the training process, nearly 4,000 patches were harvested, designating 70% for model training and the remainder for evaluation. The models were configured with a batch size of 32 spanning 400 epochs. The augmentation strategies employed include both vertical and horizontal flips to ensure robustness. All input data were standardized based on the mean and standard deviation values.

In terms of model specifications, the ResNet50 architecture was chosen as the backbone encoder for all models. Configurational tweaks included adjusting the input channels to 4 and opting for the "softmax" as the activation function. Pretrained weights were not leveraged for the backbone, implying a comprehensive training of the entire architecture. The optimizer of choice was Adam, set at a learning rate of 0.001. The focal loss was adopted as the loss function, utilizing its default parameters. Training and computations were facilitated using an RTX 3060 8 GB GPU.

3.2.1 DeepLab-v3 +

The architecture of DeepLab-v3+, as articulated by Chen et al. (2018a, b), is a pioneering approach in semantic image segmentation (Chen et al. 2018a, b). It leverages a novel encoder-decoder mechanism complemented by atrous separable convolution, which is pivotal for delineating sharp object boundaries. Although the encoder is adept at compressing feature maps and assimilating high-level semantic details, the decoder excels in restoring and refining spatial information. This balance ensures a detailed representation of the objects. The incorporation of depth-wise separable convolution is a testament to the model's commitment to computational efficiency, without sacrificing the quality of segmentation.

3.2.2 LinkNet

LinkNet is a distinctive architecture formulated for real-time semantic segmentation, which is pivotal for prompt visual scene discernment (Chaurasia and Culurciello 2017). Central to LinkNet's design is an innovative methodology for directly transmitting spatial information from the encoder modules to their congruent decoder counterparts. This mechanism aims to bolster the accuracy of visual interpretations and concurrently reduce computational latency. The intrinsic ability of LinkNet to preserve the fidelity of object boundaries within images facilitated by this bypass strategy obviates the need for supplementary training parameters.

3.2.3 MANet

A Multi-scale Attention Net (MANet) represents a sophisticated neural network architecture tailored specifically for applications in remote sensing. In the context of remote sensing, where the analysis of satellite or aerial imagery plays a fundamental role in environmental monitoring, MANet distinguishes itself through its adept multiscale analysis capabilities (Tao et al. 2020). These capabilities enable MANet to concurrently assess imagery at diverse

Table 1 Landslide mapping

Model	DeepLab-v3+	LinkNet	MANet	PSPNet	ResU-Net	SQNet
Fully convolutional	Yes	Yes	Yes	Yes	Yes	No
Deep architecture	Yes	No	Yes	Yes	Yes	Yes
Attention mechanism	Yes	No	Yes	Yes	Yes	Yes
Multi-scale features	Yes	No	Yes	Yes	Yes	Yes
Complexity	High	Low	High	High	Moderate	High

resolutions, enabling it to capture a wide spectrum of geographic characteristics and subtle nuances within imagery (Ying et al. 2019). This proficiency is underpinned by the implementation of attention mechanisms, which highlight areas of particular interest, such as alterations in terrain, variations in vegetation, or topographic irregularities indicative of environmental transformations, including the potential occurrence of landslides (Yu et al. 2022). In addition to its multiscale prowess, MANet excels in semantic segmentation, ensuring precise pixel-level labeling, which is a crucial capability for tasks such as land cover classification and the identification of regions susceptible to landslides. Furthermore, MANets possess the ability to seamlessly integrate diverse data sources, including topographic and spectral information, thereby enhancing their precision and utility in remote sensing applications. Its architectural design strikes an essential balance between computational efficiency and accuracy, an imperative consideration for the thorough analysis of extensive geographic areas and the real-time monitoring of dynamic environmental changes in the domain of remote sensing.

3.2.4 PSPNet

Developed as a response to the intricate demands of scene parsing, PSPNet stands out because it adeptly navigates the challenges of diverse scenes and broad vocabularies (Zhao et al. 2016). By employing its unique pyramid pooling module, the network aggregates and synthesizes features across multiple scales, thereby enriching its context. This multi-scale aggregation, which encompasses four distinct pyramid levels, not only enhances the network's feature extraction capabilities, but also solidifies its prowess in handling a myriad of scene types.

3.2.5 ResU-Net

Embodying the synergistic integration of Residual Networks (ResNets) with the foundational U-Net architecture, ResU-Net is an innovative stride in semantic segmentation (He et al. 2016). This enhanced model, rather than adhering to the conventional U-Net framework, introduces structures grounded in residual learning blocks, marking a departure from the plain convolution layers (Zhang et al. 2018). At the core of these blocks are minimalist skip connections, which are designed to transmit only the output of the previous layer, eliminating the introduction of redundant parameters. Such an architectural choice not only amplifies the model's learning efficiency, but also serves as a countermeasure against the pitfalls of gradient vanishing. A deeper exploration reveals that the essence of ResU-Net's architecture, as detailed by Zhang et al. (2018) is the reformulated layers that emphasize residual functions, ultimately streamlining the optimization of U-Net and bolstering the overall model performance. This design innovation, combined with its inherent capabilities, has positioned ResU-Net as a sought-after tool in the remote sensing domain, especially for discerning landslides from satellite imagery (Ghorbanzadeh et al. 2021).

3.2.6 SQNet

In the realm of autonomous driving, where efficiency is paramount, SQNet has emerged as a streamlined architecture tailored for semantic segmentation tasks (Treml et al. 2016). At the heart of SQNet lies an amalgamation of key components: the integration of ELU activation functions, an encoder reminiscent of SqueezeNet, and the innovative use of parallel dilated convolutions paired with a decoder enriched by SharpMask-like refinement modules. Of note is the adept utilization of the parallel dilated convolution layer of the network, which harmoniously fuses feature maps from various receptive fields. Through the application of four uniquely dilated convolutions, SQNet achieves multi-scale feature extraction while maintaining a commendable efficiency in parameter usage.

For the application of these six models in landslide and sinkhole detection, parameter efficiency becomes crucial. In this context, parameter efficiency refers to how well the models can accurately delineate landslide and sinkhole areas while minimizing the number of parameters used. Models with higher parameter efficiency can achieve accurate segmentation results with fewer parameters, making them more suitable for practical applications where computational resources may be limited. This is particularly important for tasks like landslide and sinkhole detection, where efficient use of computational resources can lead to faster processing times and better scalability for large-scale analysis. Some

other characteristics of our six algorithms are summerized in Table 1.

3.2.7 Model Training

The training of the models in this study was a meticulously structured process aimed at achieving accurate landslide and sinkhole detection. These models were trained using a dataset of nearly 4000 image patches, where 70% of the patches were designated for model training and the remaining 30% were reserved for performance evaluation. The training process adhered to specific parameters, including a batch size of 32 and 400 training epochs and data augmentation techniques such as flips to enhance model robustness. Data standardization based on the mean and standard deviation values was consistently applied for uniformity. The optimization was carried out applying the Adam optimizer with a learning rate of 0.001, and the chosen loss function was the focal loss. Importantly, the models were constructed using the ResNet50 architecture as their backbone and were exclusively trained from scratch, omitting the pretrained weights. This rigorous training methodology ensured that the models were well-equipped to accurately detect landslides and sinkholes in challenging environments.

3.3 Evaluation Metrics

To validate our findings, we conducted a comparative analysis of the output maps generated through automated approaches and manually compiled inventories of sinkholes and landslides derived from various sources. Subsequently, the identified instances of landslides and sinkholes were rigorously validated using established computer vision validation techniques.

To rigorously assess the performance of the six deep learning models (DeepLab-v3+, LinkNet, MANet, PSP-Net, ResU-Net, and SQNet) in detecting landslides and sinkholes, several established accuracy assessment metrics were employed. These metrics are fundamental in the realm of remote sensing and deep learning studies, ensuring a systematic and consistent evaluation across all models. True Positives (TP) represent the correctly identified instances of the target class (either landslides or sinkholes). True Negatives (TN) signify instances that were correctly identified as not being part of the target class. False Positives (FP) refer to pixels or instances that were incorrectly identified as part of the target class, while False Negatives (FN) represent instances of the target class that were missed by the models.

From these core classifications, the following evaluation metrics were derived:

1. **Precision:** A measure of the accuracy of positive predictions.

$$Precision = \frac{TP}{TP + FP} + \text{FP} \tag{1}$$

2. **Recall (or Sensitivity):** This quantifies the ability of the model to identify all the relevant instances.

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

3. **F1-Score:** Harmonizing Precision and Recall, the F1-Score balances the two metrics in a single figure.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{3}$$

These metrics were selected to provide a comprehensive and robust framework for understanding and comparing the performance of the models across the complexities inherent in UAV imagery. The results section presents the outcomes of these evaluations for each model.

4 Results

Here, we explained the results derived from six deep-learning models: DeepLab-v3+, LinkNet, MANet, PSPNet, ResU-Net, and SQNet. These models were systematically assessed for their effectiveness in detecting landslides and sinkholes using UAV imagery. The resulting curves for the aforementioned models for the 400 training epochs are shown in Fig. 3. The outcomes of their segmentation capabilities are detailed, with the metrics for landslides provided in Table 1 and for sinkholes in Table 2. To optimize performance, the hyperparameters were fine-tuned for each model.

4.1 Landslides

In assessing various DL segmentation algorithms for landslide detection, noteworthy performance metrics were observed (Table 2). Models like ResUNet exhibit high precision and recall, indicating fewer FP and TP respectively, while models like PSP-Net show relatively lower recall, suggesting a higher number of FN. DeepLab-v3+ demonstrated a precision of 0.91, signifying a 91% accuracy in predicting landslides, but with a recall of 0.78, suggesting some potential misses. The F1-Score of 0.84 reflects an overall balanced performance. Link-Net excelled with a high Precision of 0.94 and Recall of 0.87, leading to a remarkable F1-Score of 0.9, indicating effective landslide identification and few false positives. MA-Net exhibited a precision of 0.94, along with a recall of 0.91, resulting in a commendable F1-Score of 0.93, highlighting its robust performance. The PSPNet exhibited the lowest detection accuracy. ResUNet exhibited an outstanding Precision of 0.97 and Recall of

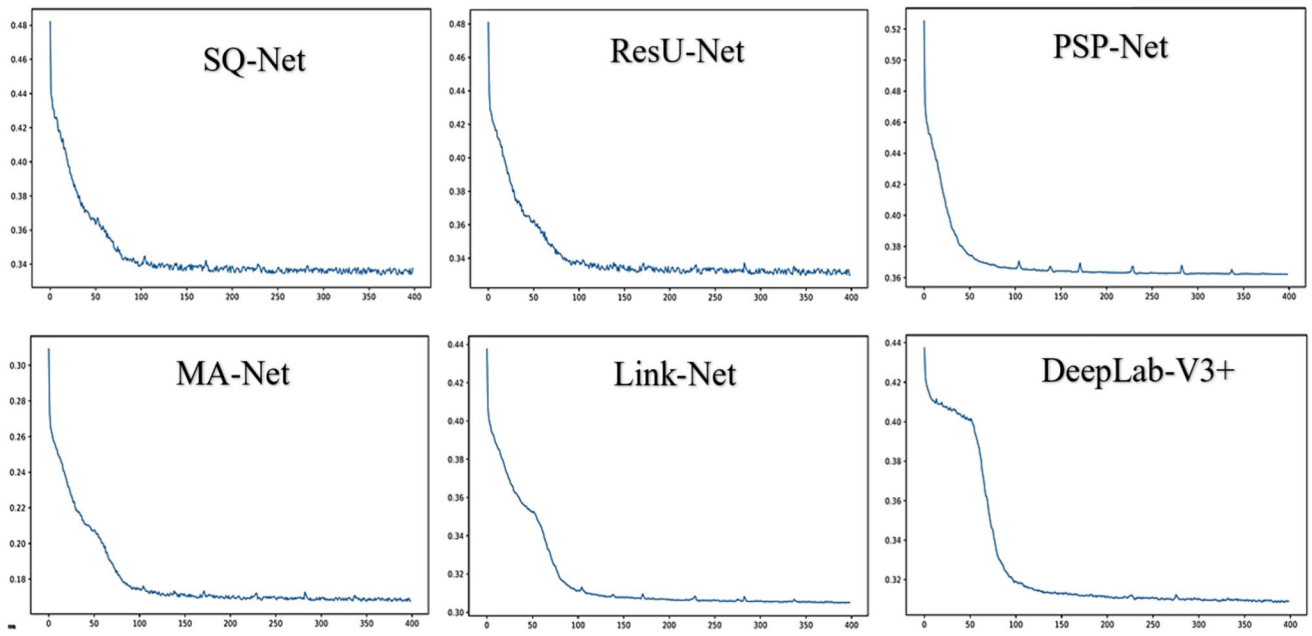

Fig. 3 The resulting curves for each model for 400 training epochs

Table 2 Landslide mapping: the evaluation metrics for the six DL algorithems on the unseen test data

Model	Precision	Recall	F1-Score
DeepLab-v3+	0.91	0.78	0.84
Link-Net	0.94	0.87	0.9
MA-Net	0.94	0.91	0.93
PSP-Net	0.90	0.55	0.68
ResUNet	0.97	0.92	**0.95**
SQ-Net	0.93	0.89	0.91

Highest F1-score is in bold

0.92, culminating in the highest F1-Score of 0.95, indicating the best landslide detection model. SQ-Net achieved a balanced Precision of 0.93 and Recall of 0.89, resulting in a commendable F1-Score of 0.91, effectively identifying landslides while maintaining a good precision-recall equilibrium.

Figure 4 demonstrates the predictive accuracy of DL segmentation algorithms. It is evident that U-Net can incorrectly predict landslides that do not actually exist. On the other hand, ADSMS U-Net addresses this issue and provides a more precise outline of the target landslide polygons.

4.2 Sinkholes

In our pursuit of accurate sinkhole detection from UAV imagery, we ventured into diverse DL segmentation algorithms, each revealing its unique capabilities and intricacies. These algorithms hold the promise of unveiling hidden mysteries beneath Earth's surface, and our findings shed light on their performance (see Table 3). Among the models, MA-Net demonstrated the lowest FP and FN rates, indicating strong accuracy in identifying sinkhole pixels. PSP-Net exhibited higher FP and lower TP, suggesting potential misclassification of non- sinkhole pixels. ResUNet and SQ-Net both achieved low FP and FN rates, with ResUNet slightly edging out in TP classification. Overall, MA-Net and ResUNet show promise for accurate sinkhole pixel identification. DeepLab-v3+ displayed promise with a Precision of 0.91 and the lowest recall of 0.76, culminating in an F1-Score of 0.83. Notably, Link-Net emerged as a standout performer, boasting a remarkable Precision of 0.95 and a recall of 0.82, resulting in a robust F1-Score of 0.88, positioning it as a leading contender in sinkhole detection. It showed a keen aptitude for identifying sinkholes while maintaining a low rate of false positives. MA-Net also demonstrated its prowess, with a commendable Precision of 0.94 and a recall of 0.85, resulting in a competitive F1-Score of 0.89, highlighting its reliability in sinkhole detection. ResUNet further affirmed its capabilities with a high Precision of 0.94 and Recall of 0.83, and F1-Score of 0.88. In contrast, PSP-Net maintained consistent Precision and Recall at 0.77, resulting in an F1-Score of 0.77, indicating a performance plateau. SQ-Net achieved an F1-Score of 0.89, precision of 0.89, and recall of 0.90, solidifying its dependability. MA-Net and SQ-Net resulted in the highest F1-Score for sinkhole detection.

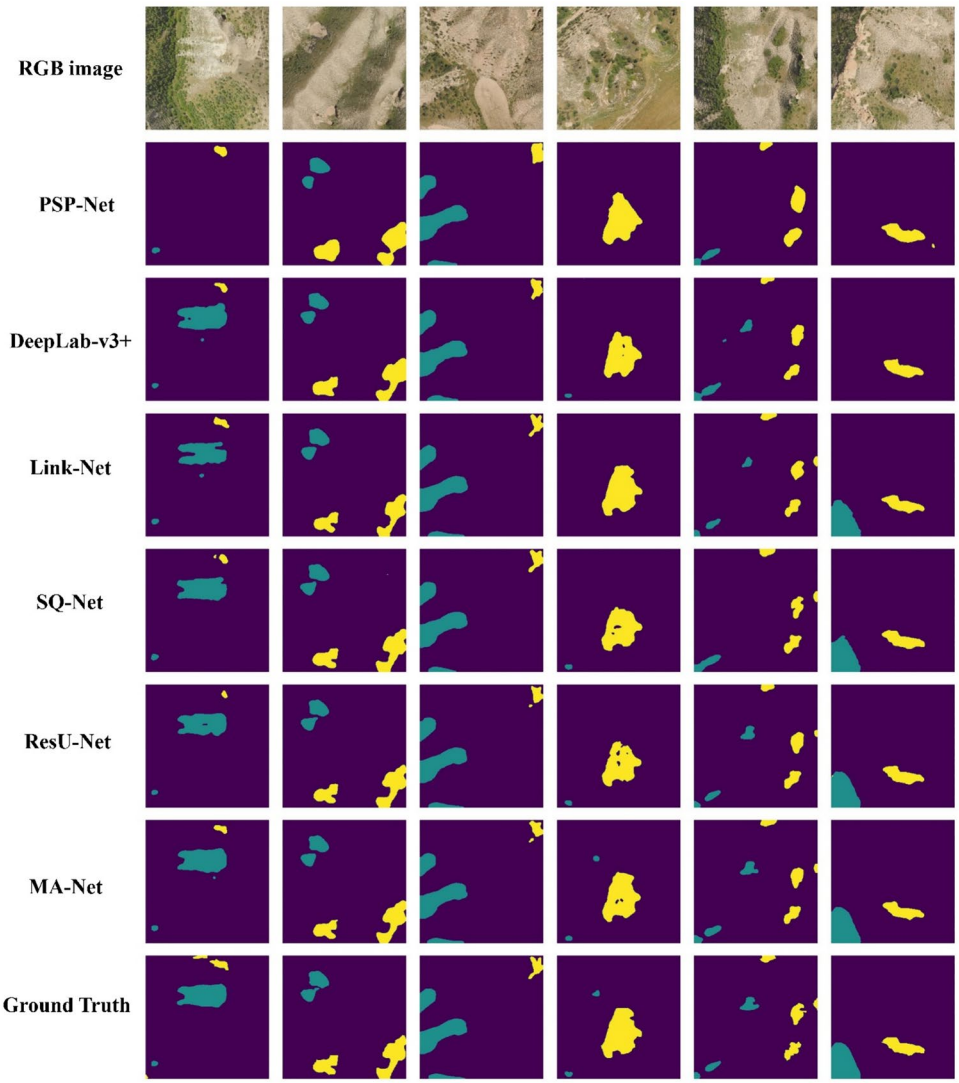

Fig. 4 Landslide (blue spots) and sinkhole (yellow spots) predictions based on the six DL models used in this research on the unseen test set patches on dataset

Table 3 Sinkhole mapping: the evaluation metrics for six DL algorithems on the unseen test data

Model	Precision	Recall	F1-Score
DeepLab-v3+	0.91	0.76	0.83
Link-Net	0.95	0.82	0.88
MA-Net	0.94	0.85	**0.89**
PSP-Net	0.77	0.77	0.77
ResUNet	0.94	0.83	0.88
SQ-Net	0.89	0.90	**0.89**

Highest F1-scores are in bold

5 Discussion

The results of our study, involving the application of various DL segmentation algorithms to both landslide and sinkhole detection tasks, revealed the substantial impact of algorithm choice on detection accuracy. Notably, algorithms such as LinkNet, MANet, and ResUNet have demonstrated exceptional performance in terms of the F1-Score, which effectively combines the metrics of Precision and Recall. LinkNet's performance reflects a remarkable equilibrium between Precision and Recall, establishing it as a robust choice for landslide and sinkhole detection tasks. Similarly, MANet has emerged as a powerful contender with high F1-Scores, implying its efficiency in identifying both landslides and sinkholes while maintaining a low rate of false positives. By showing remarkable Precision and Recall, ResUNet proves its adaptability to situations where detection accuracy is paramount, whether for landslides or sinkholes.

In contrast, the relatively high precision of PSPNet is offset by a considerably lower recall, suggesting a conservative approach for labeling both landslides and sinkholes. While

PSPNet maintains consistency, it risks missing a notable number of actual events, and it performs at a plateau, resulting in an F1-Score indicating a limitation in its detection performance. On the other hand, SQ-Net strikes a balance between Precision and Recall, providing an F1-Score that, while slightly lower than LinkNet and MANet, denotes its steadfast dependability in detecting both landslides and sinkholes.

Our findings underscore the importance of carefully selecting the most suitable DL segmentation algorithm according to specific objectives, and the need to weigh the trade-offs between Precision and Recall. LinkNet and MANet have demonstrated strength as prime contenders for accurate landslide and sinkhole detection, effectively balancing Precision and Recall. ResUNet is preferable when maximizing precision is the primary concern. However, the choice of the algorithm should consider task-specific requirements and the relative importance of false positives and false negatives. Therefore, the influence of DL segmentation algorithms on both landslide and sinkhole detection accuracy highlights the need for a tailored approach for algorithm selection based on the specific objectives of the application. DL methods for automated sinkhole mapping remain relatively underrepresented in existing literature (Hoai et al. 2019). While some research papers have utilized harnessed UAV images along with DL algorithms to map landslides, such occurrences are scarce. For example, Ghorbanzadeh et al. (2019) utilized various CNN structures to chart landslides in a densely forested Himalayan area, achieving an F1-score of 0.85. However, our research location, marked by a semi-arid climate, presents a further obstacle due to the intricacies involved in differentiating between the predominant bare ground and particular objectives, which include both sinkholes and landslides.

Karantanellis et al. (2021) evaluated a range of ML models alongside object-based image analysis to accurately define the boundaries of two rotational slides in Greece, achieving an F1-score of 0.85. However, this scenario involved only two landslides within the test set, raising questions about the feasibility of scaling such predictions to a much larger geographical extent. The adoption of DL algorithms holds tremendous promise for significantly improving the accuracy of sinkhole and landslide detection, thereby contributing to the advancement of hazard management and overall safety of vulnerable areas. Notably, conventional approaches for generating soil landform maps often rely on the visual interpretation of aerial and satellite images and are occasionally complemented by field surveys (Kariminejad et al. 2022). This means that by combining UAV technology, DL algorithms, and field surveys, the detection of sinkholes and landslides can become more efficient, accurate, and comprehensive. This integrated approach enhances hazard management, supports decision-making processes, and contributes to the safety and well-being of vulnerable regions. In other words, using UAV, DL, and field surveys to detect landslides and sinkholes offers several benefits, such as high-resolution data acquisition, improved accessibility, rapid data collection, real-time monitoring, detailed 3D mapping, and ground truth validation (Mittal et al. 2020; Osco et al. 2021). Furthermore, this integrated approach allows for more precise detection and mapping of hazards, such as landslides and sinkholes, land subsidence, or even gullies. By leveraging high-resolution data from UAVs, advanced analysis with DL algorithms, and ground truth validation through field surveys, decision makers can efficiently assess risks, implement timely mitigation measures, and ultimately enhance the safety and resilience of communities in the face of natural hazards.

Moreover, the research findings from this study have several potential future applications, including hazard mapping and risk assessment, early warning systems, urban planning and infrastructure design, emergency response and disaster management, and R&D. Overall, the research findings have significant implications for improving hazard management, enhancing safety, and promoting resilience in semiarid environments. By leveraging deep learning algorithms and UAV imagery, this study provides valuable tools and methodologies for a range of applications, from urban planning and infrastructure design to emergency response and long-term risk reduction strategies.

6 Conclusion

This research examined the quick identification of sinkholes and landslides in a difficult semi-arid location, closely analyzing different DL segmentation algorithms. Additionally, we delved into the utilization of UAV-derived data for sinkhole detection by employing two cutting-edge models. This strategic choice of a demanding mapping location allowed us to develop a versatile and robust approach tailored for rigorous detection settings while retaining scalability for application in less challenging conditions, including densely vegetated areas. This research represents a pioneering endeavor in the application of DL segmentation algorithms for mapping both sinkholes and landslides in a demanding semi-arid context, potentially setting a new standard in this field. To the best of our knowledge, this is the first instance of employing DL segmentation algorithms for simultaneous detection of two distinct features in a semi-arid region. The outcomes of our study demonstrated highly promising results, with successful mapping of landslides achieving an impressive F1-Score of 0.95% and sinkholes achieving an F1-Score of 89% in a challenging semi-arid environment. ResUNet demonstrated exceptional performance, achieving an impressive Precision of 0.97 and Recall of 0.92. These

remarkable results culminated in the highest F1-Score of 0.95, establishing ResUNet as the top-performing model for landslide detection. On the other hand, MA-Net and SQ-Net emerged as the leading models, yielding the highest F1-Score for sinkhole detection. Our results highlight the value of combining DL algorithms with UAV data for reliable landslide mapping in these areas. However, it is important to recognize that topographical derivatives from UAV-generated DSMs may not be the best approach for mapping sinkholes. Future research will focus on applying this approach to monitor the evolution of these natural hazards over time, shedding light on their dynamic behavior. Notably, LinkNet and MANet have emerged as strong performers, demonstrating their effectiveness in identifying both landslides and sinkholes with a minimal rate of false positives. In contrast, the conservative approach of PSPNet, which balances high precision with lower recall, suggests limitations in its detection performance. Overall, our findings emphasize the importance of meticulous algorithm selection, considering specific objectives, and the trade-offs between Precision and Recall. In practical applications, LinkNet and MANet proved to be reliable choices for accurate detection, underscoring the need for a tailored approach for algorithm selection aligned with the objectives of the application.

Acknowledgements Open access funding was provided by University of Natural Resources and Life Sciences, Vienna (BOKU). It also thanks to the support of the Iran National Science Foundation (Grant No. 99011991).

Funding Open access funding provided by University of Natural Resources and Life Sciences Vienna (BOKU). The authors have not disclosed any funding.

Data Availability Data will be made available on reasonable request to corresponding author.

Declarations

Conflict of interest The authors have not disclosed any competing interests.

References

Agrawal K, Baweja Y, Dwivedi D, Saha R, Prasad P, Agrawal S, Kapoor S, Chaturvedi P, Mali N, Kala VU, Dutt V (2017) A comparison of class imbalance techniques for real-world landslide predictions. In: 2017 international conference on machine learning and data science (MLDS). IEEE, pp 1–8

Bernatek A (2015) The influence of piping on mid-mountain relief: a case study from the polish bieszczady Mts. (Eastern Carpathians). Carpathian J Earth Environ Sci 10(1):107–120

Bernatek-Jakiel A, Poesen J (2018) Subsurface erosion by soil piping: significance and research needs. Earth-Sci Rev 185:1107–1128

Brovkina O, Cienciala E, Surový P, Janata P (2018) Unmanned aerial vehicles (uav) for assessment of qualitative classification of norway spruce in temperate forest stands. Geo Spat Inf Sci 21:12–20

Chaurasia A, Culurciello E (2017) LinkNet: exploiting encoder representations for efficient semantic segmentation. In: 2017 IEEE Visual Communications and Image Processing (VCIP), pp 1–4. https://doi.org/10.1109/VCIP.2017.8305148

Chen W, Shahabi H, Zhang S, Khosravi K, Shirzadi A, Chapi K, Pham BT, Zhang T, Zhang L, Chai H et al (2018a) Landslide susceptibility modeling based on gis and novel bagging-based kernel logistic regression. Appl Sci 8:2540

Chen L, Zhu Y, Papandreou G, Schroff F, Adam H (2018b) Encoder–decoder with atrous separable convolution for semantic image segmentation. In: European Conference on Computer Vision (ECCV 2018). pp 833–851. https://doi.org/10.1007/978-3-030-01234-2_49

Chen Y, Chen W, Janizadeh S, Bhunia GS, Bera A, Pham QB, Linh NTT, Balogun AL, Wang X (2021) Deep learning and boosting framework for piping erosion susceptibility modeling: spatial evaluation of agricultural areas in the semi-arid region. Geocarto Int 37:4628–4654

Fernández T, Pérez JL, Cardenal J, Gómez JM, Colomo C, Delgado J (2016) Analysis of landslide evolution affecting olive groves using uav and photogrammetric techniques. Remote Sens 8:837

Ghorbanzadeh O, Meena SR, Blaschke T, Aryal J (2019) UAV-based landslide detection using deep-learning convolutional neural networks. Remote Sens 11(17):2046

Ghorbanzadeh O, Crivellari A, Ghamisi P, Shahabi H, Blaschke T (2021) A comprehensive transferability evaluation of U-Net and ResU-Net for landslide detection from Sentinel-2 data (case study areas from Taiwan, China, and Japan). Sci Rep 11(1):14629. https://doi.org/10.1038/s41598-021-94190-9

He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp 770–778. https://doi.org/10.1109/CVPR.2016.90.

Higgins CG, Coates DR et al (1990) Groundwater geomorphology: the role of subsurface water in Earth-surface processes and landforms, vol 252. Geological Society of America

Hoai NV, Dung NM, Ro S (2019) Sinkhole detection by deep learning and data association. In: 2019 Eleventh International Conference on Ubiquitous and Future Networks (ICUFN). IEEE, pp 211–213

Jones JAA, Crane FG (1984) Pipeflow and pipe erosion in the Maesnant experimental catchment. In: International Geographical Union Commission on Field Experiments in Geomorphology. Meeting, pp 55–72

Karantanellis E, Marinos V, Vassilakis E, Hölbling D (2021) Evaluation of machine learning algorithms for object-based mapping of landslide zones using UAV data. Geosciences 11(8):305

Kariminejad N, Pourghasemi HR, Hosseinalizadeh M (2022) Analytical techniques for mapping multi-hazard with geo-environmental modeling approaches and UAV images. Sci Rep 12:14946

Mittal P, Singh R, Sharma A (2020) Deep learning-based object detection in low-altitude UAV datasets: a survey. Image vis Comput 104:104046

Moharrami M, Naboureh A, Gudiyangada Nachappa T, Ghorbanzadeh O, Guan X, Blaschke T (2020) National-scale landslide susceptibility mapping in Austria using fuzzy best-worst multi-criteria decision-making. ISPRS Int J Geo Inf 9(6):393

Naboureh A, Li A, Bian J, Lei G, Nan X (2023a) Land cover dataset of the China Central-asia West-asia Economic Corridor from 1993 to 2018. Sci Data 10(1):728

Naboureh A, Li A, Bian J, Lei G (2023b) National scale land cover classification using the semiautomatic high-quality reference sample generation (HRSG) method and an adaptive supervised classification scheme. IEEE J Sel Top Appl Earth Obs Remote Sens 16:1858–1870

Osco LP, Junior JM, Ramos APM, de Castro Jorge LA, Fatholahi SN, de Andrade Silva J, Matsubara ET, Pistori H, Gonçalves WN, Li J (2021) A review on deep learning in UAV remote sensing. Int J Appl Earth Obs Geoinf 102:102456

Petschko H, Bell R, Leopold P, Heiss G, Glade T (2013) Landslide inventories for reliable susceptibility maps in Lower Austria. In: Landslide science and practice: volume 1: landslide inventory and susceptibility and hazard zoning, pp 281–286

Tao A, Sapra K, Catanzaro B (2020) Hierarchical multi-scale attention for semantic segmentation. arXiv preprint arXiv:2005.10821

Tavakkoli Piralilou S, Shahabi H, Jarihani B, Ghorbanzadeh O, Blaschke T, Gholamnia K, Meena S et al (2019) Landslide detection using multi-scale image segmentation and different machine learning models in the higher himalayas. Remote Sens 11(21):2575. https://doi.org/10.3390/rs11212575

Treml M, Arjona-Medina JA, Unterthiner T, Durgesh R, Friedmann F, Schuberth P, Mayr A, Heusel M, Hofmarcher M, Widrich M, Nessler B, Hochreiter S (2016) Speeding up semantic segmentation for autonomous driving. In: 29th Conference on Neural Information Processing Systems (NIPS 2016). Barcelona, Spain

Watson CS, Kargel JS, Tiruwa B (2019) UAV-derived himalayan topography: hazard assessments and comparison with global dem products. Drones 3:18

Windrim L, Bryson M, McLean M, Randle J, Stone C (2019) Automated mapping of woody debris over harvested forest plantations using UAVs, high-resolution imagery, and machine learning. Remote Sens 11:733

Yang Z-H, Lan H-X, Gao X, Li L-P, Meng Y-S, Wu Y-M (2015) Urgent landslide susceptibility assessment in the 2013 lushan earthquake-impacted area, Sichuan province, China. Nat Hazards 75:2467–2487

Ying X, Wang Q, Li X, Yu M, Jiang H, Gao J, Yu R (2019) Multi-attention object detection model in remote sensing images based on multi-scale. IEEE Access 7:94508–94519

Yu B, Xu C, Chen F, Wang N, Wang L (2022) HADeenNet: a hierarchical-attention multi-scale deconvolution network for landslide detection. Int J Appl Earth Obs Geoinf 111:102853

Zhang Z, Liu Q, Wang Y (2018) Road extraction by deep residual U-Net. IEEE Geosci Remote Sens Lett 15:749–753. https://doi.org/10.1109/LGRS.2018.2802944

Zhang Y, Yue P, Zhang G, Guan T, Lv M, Zhong D (2019) Augmented reality mapping of rock mass discontinuities and rockfall susceptibility based on unmanned aerial vehicle photogrammetry. Remote Sens 11:1311

Zhao H, Shi J, Qi X, Wang X, Jia J (2016) Pyramid scene parsing network. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA, 2017, pp 6230–6239. https://doi.org/10.1109/CVPR.2017.660