




Article

Enhancing Landslide Detection with SBConv-Optimized U-Net Architecture Based on Multisource Remote Sensing Data

Yingxu Song ^{1,2,†} , Yujia Zou ², Yuan Li ^{3,4,†}, Yueshun He ^{2,*}, Weicheng Wu ³ , Ruiqing Niu ⁵ and Shuai Xu ^{6,*} 

¹ Engineering Research Center for Seismic Disaster Prevention and Engineering Geological Disaster Detection of Jiangxi Province (East China University of Technology), Nanchang 330013, China

² School of Information Engineering, East China University of Technology, Nanchang 330013, China; 2022110192@ecut.edu.cn

³ Key Lab of Digital Land and Resources, Faculty of Earth Sciences, East China University of Technology, Nanchang 330013, China; 2020110084@ecut.edu.cn (Y.L.); wuwch@ecut.edu.cn (W.W.)

⁴ School of Surveying and Geoinformation Engineering, East China University of Technology, Nanchang 330013, China

⁵ Institute of Geophysics and Geomatics, China University of Geosciences, Wuhan 430074, China; niuruiqing@cug.edu.cn

⁶ College of Vocational and Technical Education, South China Normal University, Shanwei 516600, China

* Correspondence: heys@ecut.edu.cn (Y.H.); xushuai@cug.edu.cn (S.X.)

† These authors contributed equally to this work.

Abstract: This study introduces a novel approach to landslide detection by incorporating the Spatial and Band Refinement Convolution (SBConv) module into the U-Net architecture, to extract features more efficiently. The original U-Net architecture employs convolutional layers for feature extraction, during which it may capture some redundant or less relevant features. Although this approach aids in building rich feature representations, it can also lead to an increased consumption of computational resources. To tackle this challenge, we propose the SBConv module, an efficient convolutional unit designed to reduce redundant computing and enhance representative feature learning. SBConv consists of two key components: the Spatial Refined Unit (SRU) and the Band Refined Unit (BRU). The SRU adopts a separate-and-reconstruct approach to mitigate spatial redundancy, while the BRU employs a split-transform-and-fuse strategy to decrease band redundancy. Empirical evaluation reveals that models equipped with SBConv not only show a reduction in redundant features but also achieve significant improvements in performance metrics. Notably, SBConv-embedded models demonstrate a marked increase in Recall and F1 Score, outperforming the standard U-Net model. For instance, the SBConvU-Net variant achieves a Recall of 75.74% and an F1 Score of 73.89%, while the SBConvResU-Net records a Recall of 70.98% and an F1 Score of 73.78%, compared to the standard U-Net's Recall of 60.59% and F1 Score of 70.91%, and the ResU-Net's Recall of 54.75% and F1 Score of 66.86%. These enhancements in detection accuracy underscore the efficacy of the SBConv module in refining the capabilities of U-Net architectures for landslide detection of multisource remote sensing data. This research contributes to the field of landslide detection based on remote sensing technology, providing a more effective and efficient solution. It highlights the potential of the improved U-Net architecture in environmental monitoring and also provides assistance in disaster prevention and mitigation efforts.

Keywords: deep learning; landslide detection; remote sensing; U-Net; SBConv



Citation: Song, Y.; Zou, Y.; Li, Y.; He, Y.; Wu, W.; Niu, R.; Xu, S. Enhancing Landslide Detection with SBConv-Optimized U-Net Architecture Based on Multisource Remote Sensing Data. *Land* **2024**, *13*, 835. <https://doi.org/10.3390/land13060835>

Academic Editors: Candide Lissak, Christopher Gomez and Vittoria Vandelli

Received: 24 March 2024

Revised: 23 May 2024

Accepted: 7 June 2024

Published: 12 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Landslides, as a prevalent and hazardous natural phenomenon, pose significant risks to human lives and infrastructure [1,2]. Landslides are recognized as a significant geo-environmental concern and a key geomorphological characteristic influenced by numerous surface processes. They represent a complex phenomenon, encompassing diverse geo-physical and hydro-meteorological factors, unlike any other natural hazard [3]. Therefore,

landslide detection and modeling plays a crucial role in their analysis [4]. These studies are categorized based on the triggering factors of landslides [5], primarily including those induced by rainfall [6–8] and earthquakes [9–11]. Key research topics in landslides include the prediction of displacement in single landslide [9,12], identification of landslides [13–15], susceptibility assessment [2,16], hazard evaluation [17], and risk assessment [18,19].

Traditional methods for landslide detection, often constrained by manual efforts and limited accuracy, necessitate advancements in automated detection technologies [13].

Remote sensing technology has a broad and profound impact on landslide identification or mapping [20–24]. Satellite and aerial images provide valuable data for landslide identification, analysis, and monitoring, with multi-temporal satellite images showing changes in landslides over time [21]. Synthetic Aperture Radar (SAR) technology can penetrate clouds and work in any weather condition, especially through Interferometric Synthetic Aperture Radar (InSAR) technology, which can measure ground displacement and provide important information for identifying landslide risk areas [25–27]. Optical and infrared images help identify potential landslide areas by analyzing the spectral characteristics of the images [28]. Data fusion technology combines multiple remote sensing data sources, such as optical, SAR, and LiDAR data, while machine learning algorithms can automate the analysis of these data, improving the accuracy of landslide detection [13,29]. LiDAR technology provides high-resolution topographic data for analyzing terrain and identifying landslide risk areas [30]. Research on real-time monitoring of landslides is also being explored, with the aim of achieving real-time/near-real-time early warning of landslides [31,32]. Finally, the integration of remote sensing data with other data sources such as geology, hydrology, and meteorology offers the possibility of a comprehensive understanding of landslide risk and the development of better prediction models. However, challenges such as data availability, resolution, and cost still exist, which may affect the effective application of remote sensing technology in landslide identification and monitoring.

Early landslide remote sensing identification mainly relied on visual interpretation, identifying potential landslide areas by analyzing surface features in remote sensing images [13], often constrained by manual efforts and limited accuracy, necessitate advancements in automated detection technologies. Change detection methods are also used for landslide detection, but these methods need 2 remote sensing images (before and after landslide events) at least, limited by the data availability.

With the development of computer technology, researchers began to use shallow machine learning techniques, such as support vector machines [33,34] and random forests [28], to automatically identify landslides. These methods usually require manual selection and extraction of features, and then use these features to train classifiers. In recent years, with the rise of deep learning technology, researchers have begun to use convolutional neural networks (CNNs) and other deep learning algorithms for landslide identification. Wang et al. published a study on remote sensing landslide identification based on convolutional neural networks [35]. Another study developed an integrated machine learning approach that combined multi-source data and pixel- and object-based processing to detect landslides, and also systematically studied the impact of training data size on detection performance [36]. Researchers have also used improved Transformers models to identify landslides using multi-feature remote sensing data [37,38]. Unlike shallow machine learning techniques, deep learning can automatically learn and extract features from data without manual intervention [39]. These studies provide an important research foundation for landslide identification based on remote sensing data and machine learning technology, and also show the current state and future development trends of this field [13].

The U-Net architecture [40], originally developed for biomedical image segmentation, has been particularly influential in advancing the field of semantic segmentation. Its novel architecture, characterized by a U-shaped design with a contracting path to capture context and an expansive path that enables precise localization, has been widely adopted and adapted for various image segmentation tasks. The effectiveness of U-Net is largely due to its use of skip connections, which allow for the combination of low-level detail with

high-level contextual information, making it particularly suitable for tasks that require the segmentation of fine-grained structures [10,41,42]. In the field of geospatial analysis, U-Net and its derivatives have been successfully applied to object detection and image segmentation tasks in remote sensing images. Similarly, they have also been extensively used for the identification and assessment of landslide risks [11,13,41].

However, including the U-Net architecture's use of convolutional layers for feature extraction, faces significant challenges in dealing with feature redundancy and optimizing channel feature processing [43]. To address these challenges, this paper integrates a plug-and-play convolution module—SBConv [43]—into the U-Net architecture. The SBConv module consists of two key components: the Spatial Refinement Unit (SRU) and the Band Refinement Unit (BRU), designed to reduce spatial and band redundancy, thereby enhancing the network's representation learning and improving landslide detection accuracy. Through this research, our goal is to leverage the advantages of spatial and band feature refinement for more precise and reliable landslide detection.

Through this research, our goal is to leverage the advantages of spatial and band feature refinement for more precise and reliable landslide detection. A key innovation is the integration of the SBConv module into the U-Net architecture, which enhance traditional convolutional layers. This enhancement significantly boosts the model's capabilities in aggregating features and representing complex patterns through the inclusion of the Spatial-Refined Unit (SRU) and the Band-Refined Unit (BRU). Experimental results confirm that the U-Net and ResU-Net architectures improved with SBConv outperform traditional models, showcasing marked improvements in landslide detection performance.

The remainder of this paper is structured as follows. The Materials and Methods section provides a detailed description of our improved U-Net approach and the data used in this study. The Results section presents the experiments and analysis, followed by a Discussion section. Finally, the Conclusion summarizes the findings and limitations of the study, offering perspectives for future work.

2. Materials and Methods

2.1. Materials

The Landslide4Sense dataset is designed as a multi-source benchmark for training deep learning (DL) models in landslide detection [44,45]. Given the challenges posed by small or homogeneous datasets, this benchmark incorporates data from four diverse geographic regions [44]. This approach ensures a broad representation of landslide characteristics. The specific geographical locations are marked on the map of Asia provided by ArcGIS 10.8.2 (see Figure 1). As illustrated in the figure, the four areas—the Iburi-Tobu Area of Hokkaido, the Kodagu District of Karnataka, the Rasuwa District of Bagmati, and Western Taitung County—are located in Japan, India, Nepal, and Taiwan, China, respectively. Among these, the landslides in Japan and Nepal are of the earthquake-induced type (the April 2015 Nepal earthquake and the August 2018 Hokkaido earthquake), while those in India and Taiwan, China, are triggered by heavy rainfall and Typhoon [44]. A detailed landslide distribution, typically existing landslide inventory maps could be seen in the article (Ghorbanzadeh et al., 2022) [44].

It aims to improve model transferability to new regions by incorporating various landslide triggers and environmental conditions. The dataset also utilizes data from Sentinel-2 and ALOS PALSAR sensors. Key aspects include detailed annotations for landslide inventory and a focus on geographic diversity to enhance the robustness and generalizability of DL models trained with this dataset. The data is accessible on Future Development Leaderboard for future evaluation at <https://github.com/iarai/Landslide4Sense-2022> (accessed on 30 December 2023) [45].

The steps, procedures, and approaches employed in this study are outlined in the workflow depicted in Figure 2. The input data, comprising multi-spectral Sentinel-2 imagery and Digital Elevation Model (DEM) derivatives (slope and elevation), undergo a series of transformations through an encoder-decoder structure reminiscent of the U-Net architecture.

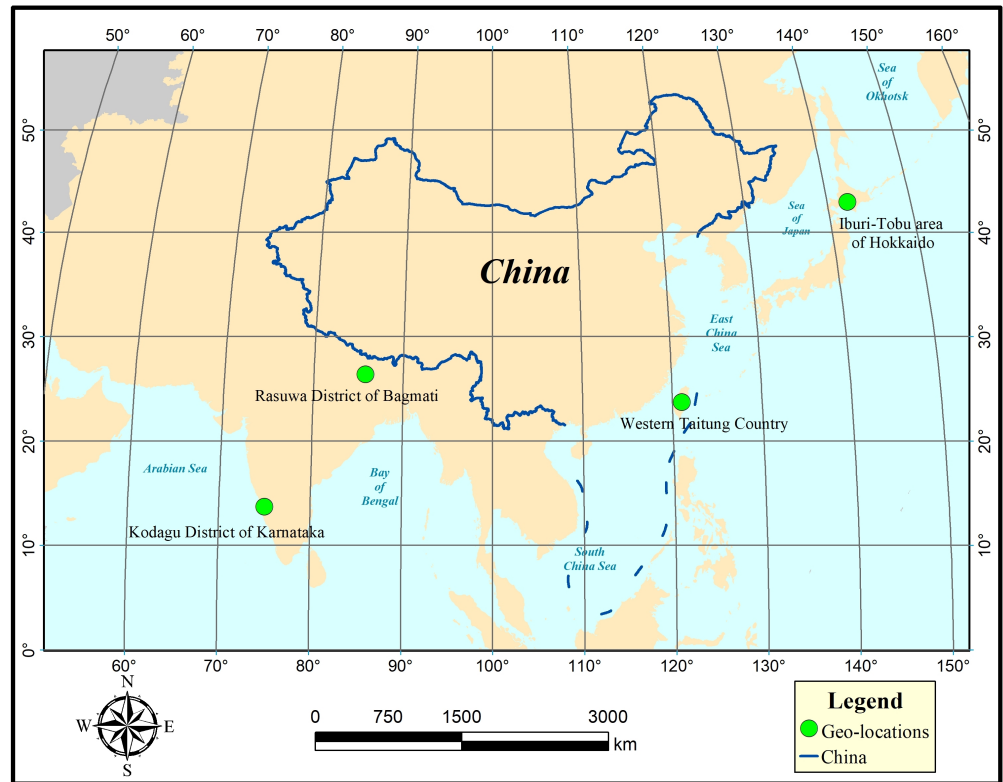


Figure 1. Geo-locations of the landslide samples source.

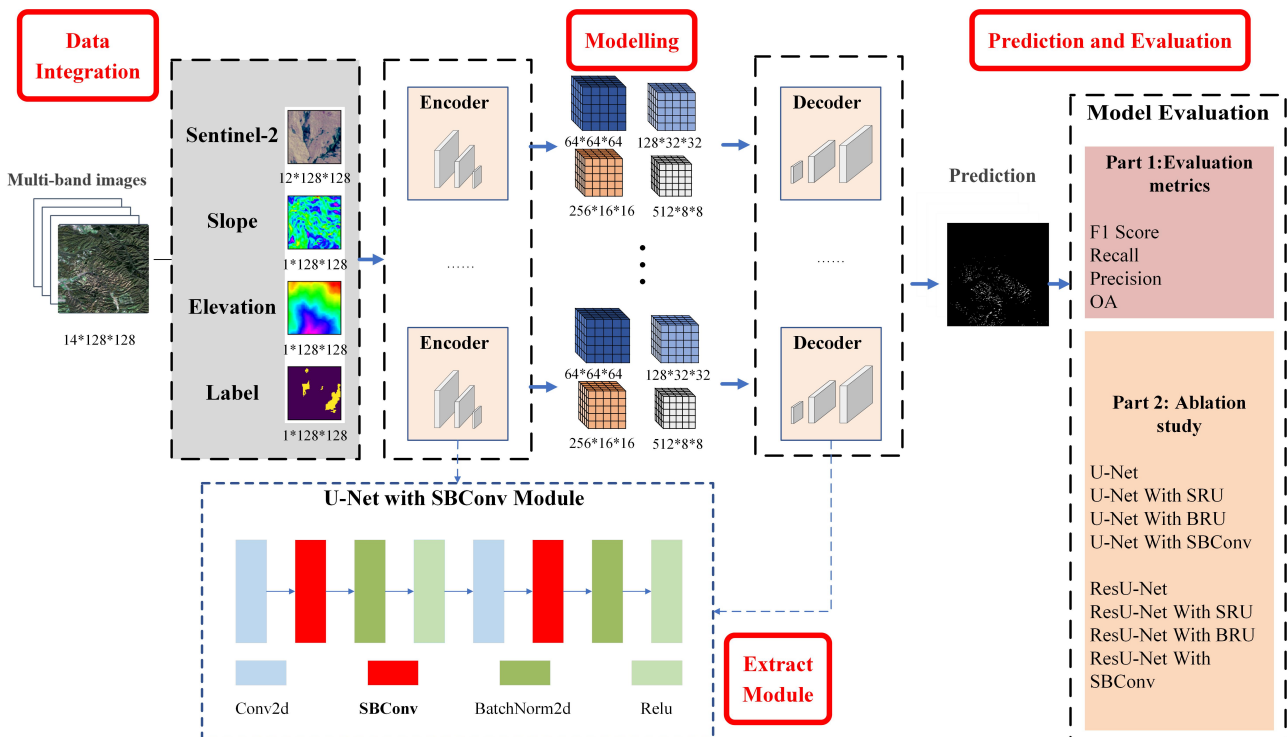


Figure 2. Research workflow.

The encoder segments sequentially extract the features of the input, employing a downsampling strategy to compress spatial information while enhancing feature specificity. This is visualized through the diminishing spatial dimensions in the encoder blocks, which reflect the hierarchy of learned representations.

In counterpoint, the decoder segments engage in an upsampling routine, reconstructing the spatial dimension of the original input from the encoded abstract feature maps. This is evident from the incrementing dimensions in the decoder blocks, signifying the expansion of data towards its original resolution.

The feature extract module, demarcated by a dashed blue outline, enumerates the pivotal components within the encoder-decoder interplay. These include Conv2d layers, which apply two-dimensional convolutions to extract spatial features; SBConv, a specialized convolution module; BatchNorm2d, which normalizes the features within a batch to improve stability and performance; and ReLU activations, which introduce non-linearity into the learning process.

The resultant output from this sophisticated computational apparatus is a binary prediction map, delineating the edge of landslide occurrences with crisp, unambiguous boundaries.

Finally, the model's inferential prowess is gauged against a suite of evaluation metrics—F1 Score, Recall, Precision, and Overall Accuracy (OA). These metrics collectively quantify the model's predictive accuracy and reliability, offering a multi-faceted assessment of its classification efficacy.

2.2. Methods

2.2.1. U-Net

U-Net is a convolutional neural network (CNN) architecture designed for semantic segmentation tasks, particularly in medical image analysis [40]. The name "U-Net" derives from its U-shaped architecture, comprising a contracting path and an expanding path. The contracting path is a traditional CNN consisting of convolutional and pooling layers that progressively reduce the spatial dimensions of the input image while increasing the number of feature channels to extract hierarchical features. The expanding path involves upsampling and concatenation operations to gradually restore the spatial resolution of the feature maps, with each upsampling step accompanied by a convolutional layer that reduces the number of feature channels. Skip connections are established between corresponding layers in the contracting and expanding paths to help the network retain detailed information from early layers, aiding in precise localization and segmentation. The final layer typically utilizes a convolutional layer with a Softmax activation function to generate pixel-wise predictions, where each pixel in the output corresponds to a class label indicating the likelihood of belonging to a specific semantic category. U-Net is often trained using pixel-wise cross-entropy loss or other suitable segmentation loss functions to minimize the discrepancy between predicted segmentation masks and ground truth masks during training. Widely adopted in medical imaging tasks such as cell segmentation, tumor detection, and organ segmentation, U-Net's effectiveness lies in its ability to efficiently capture both local and global context while preserving spatial information.

2.2.2. ResU-Net

ResU-Net [46] is a deep learning model specifically designed for medical image segmentation, and it has also been successfully applied to target detection in remote sensing images. This model is based on the classical U-Net architecture and incorporates residual learning to enhance performance. Structurally, ResU-Net consists of a contracting path and an expanding path, with residual connections added at each layer to effectively address the vanishing gradient problem, allowing direct signal propagation from lower to higher network layers. This design enables ResU-Net to handle deeper network structures, better capturing details, particularly suitable for dealing with images with complex or blurred boundaries. ResU-Net has been widely used in various image analysis tasks such as tumor detection, organ segmentation in medical images, and target detection in remote sensing images, demonstrating outstanding segmentation performance and generalization ability.

2.2.3. SBConv Module

To enhance the accuracy of landslide detection, we have optimized the U-Net architecture by integrating the SBConv module, aimed at reducing the redundancy in the U-Net structure during the information extraction process. U-Net is praised for its outstanding performance in medical image segmentation tasks, attributed to its symmetrical design and skip connections [40]. These features enable efficient extraction and fusion of features from both the encoder and decoder ends, thereby capturing detailed information more precisely. Although the architecture performs excellently, it reveals the need for further optimization when faced with the complex challenges of geographical spatial images characterized by multi-resolution and multi-spectral features.

As shown in Figure 3, the traditional U-Net architecture employs a DoubleConv module, which consists of two consecutive convolution layers for feature extraction. In our research, we adopted an innovative approach by integrating an SBConv module right after the standard Conv2d convolution layer within the DoubleConv module. The foundational Spatial and Channel Reconstruction Convolution (SCConv) operator was initially introduced in 2023 as a plug-in convolution module [43]. Building upon this innovation, ref. [47] expanded the SCConv operator, dubbing the enhanced module as ESBCConv. In our work, we adopted the ESBCConv module and designated it SBConv for our purposes.

This strategy was aimed at reducing information redundancy, leading to the development of the novel SBConvU-Net architecture. This enhanced DoubleConv module has been rebranded as the Improved DoubleConv Unit (IDCU), with its architecture depicted as part of the feature extraction or reconstruction module in Figure 2.

The enhancement of this architecture is rooted in the understanding that the feature extraction phase is crucial for the success of semantic segmentation tasks. By bolstering this stage, our aim is to increase the network’s sensitivity and specificity, thereby enhancing its ability to precisely identify potential landslide areas. The SBConv module stands as the cornerstone of our improved U-Net, utilizing sophisticated convolution operations to ensure the network efficiently learns the most significant features within complex geographic spatial datasets.

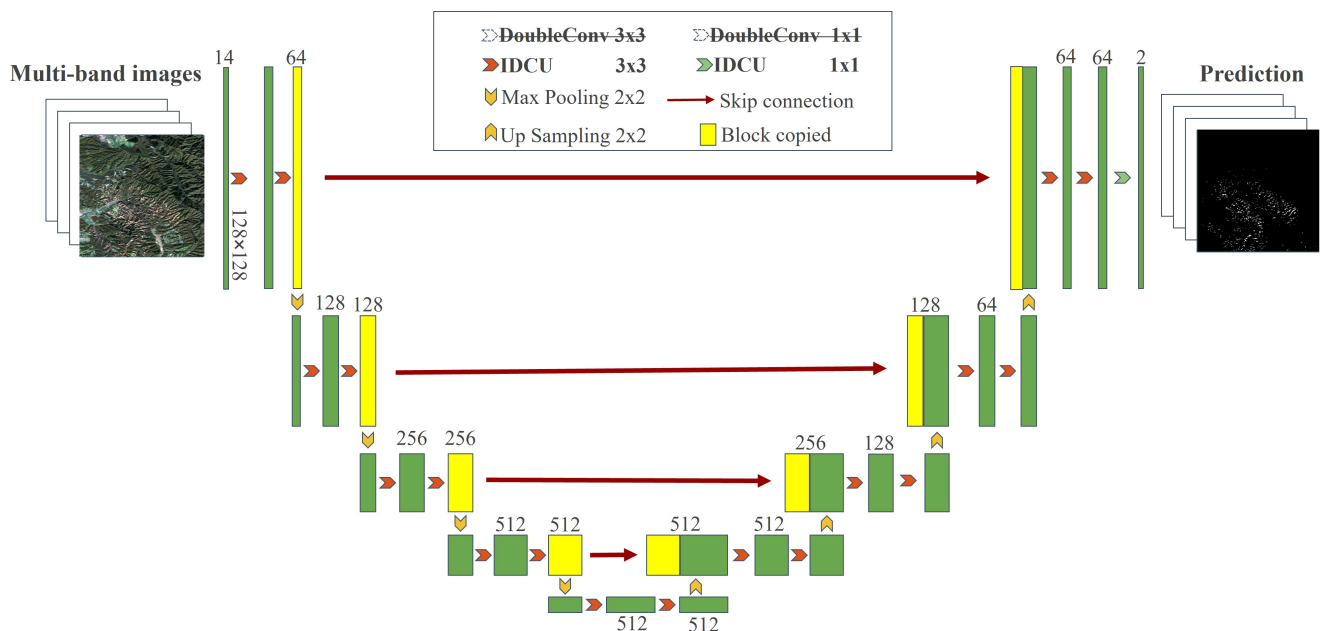


Figure 3. The structure of the improved U-Net architecture.

2.2.4. Spatial and Band Refined Convolution (SBConv)

The SBConv module is a novel component introduced in our architecture to augment the specificity of feature refinement, as shown in Figure 4. The SBConv module is designed

to process features through two sequential sub-modules: the Spatial Refined Unit (SRU) and the Band Refined Unit (BRU), as illustrated below.

- **Input ConvBlock:** The process begins with an input feature map that is first processed by a standard convolutional block (ConvBlock), preparing the features for subsequent refinement.
- **ResBlock + SBCConv:** The output from the initial ConvBlock is then fed into a residual block combined with the SBCConv. This combination allows for the incorporation of both residual learning and specialized convolution operations to enhance feature representation.
- **SRU:** Within the SBCConv, the Spatial Refined Unit (SRU) takes the input feature map X and applies a series of operations to refine the spatial characteristics of the features, yielding a spatially-refined feature map X' .
- **BRU:** Following spatial refinement, the Band Refined Unit (BRU) further processes X' to emphasize and recalibrate the spectral information, resulting in the band-refined feature map Y .
- **Output ConvBlock:** Finally, the refined feature map Y is passed through an output convolutional block (Output ConvBlock), producing the final output that is used in further layers or for constructing the final segmentation map.

The SBCConv module, encompassing the SRU and BRU, is encapsulated within our network to enhance the extraction and differentiation of features relevant to accurate image segmentation.

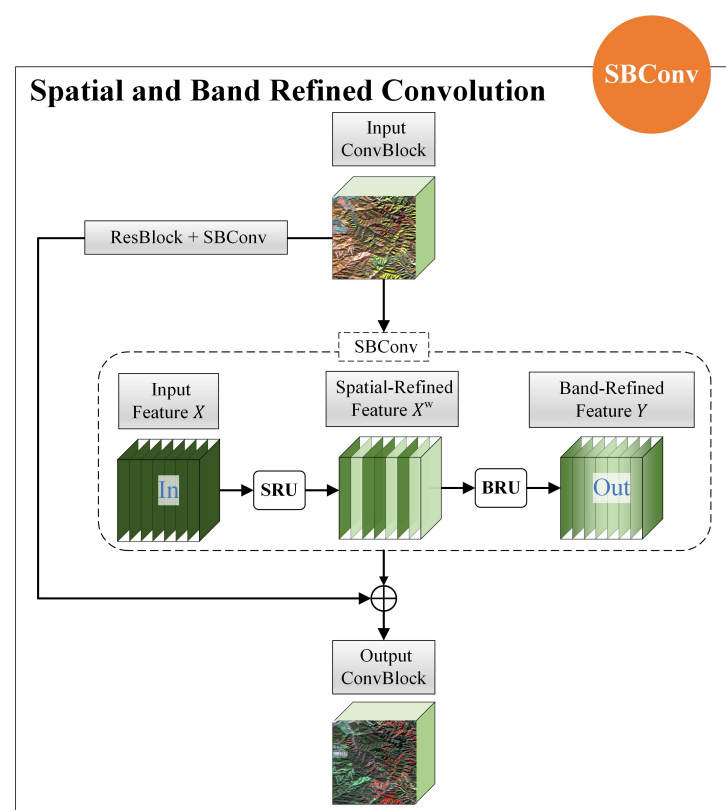


Figure 4. The structure of the IDCU module.

This module synergizes the Spatial Refined Unit (SRU) with the Band Refined Unit (BRU), aiming to significantly optimize spatial and spectral feature extraction capabilities tailored for landslide identification tasks. The SRU is engineered to amplify the delineation of spatial details, which is paramount in accurately segmenting landslide-prone regions from complex geospatial backgrounds. Meanwhile, the BRU is designed to recalibrate and refine the spectral features across the bands, ensuring that the model accentuates relevant

characteristics essential for identifying subtle differences in the multispectral data indicative of landslides. The architecture begins with a multi-band input, suggestive of a combination of spectral bands from remote sensing data, each with a resolution of 128×128 pixels. This input feeds into an encoder-decoder structure with skip connections, typical of a U-Net framework, but with the novel addition of SBConv modules at each stage. As we progress from the input to deeper layers, the feature map dimensions are systematically reduced using 2×2 max pooling operations, while the feature depth increases, peaking at 512 channels. This expansive feature extraction phase is balanced by the decoder path, where up-sampling operations incrementally restore the spatial resolution. Simultaneously, skip connections from corresponding encoder stages reintegrate contextually rich details by concatenating feature maps from the encoder with those of the decoder. The culmination of this intricate process is a prediction output that precisely maps the likelihood of landslide occurrences within the input imagery. It is the fusion of the SRU and BRU within the SBConv module that empowers our architecture with the discernment to effectively differentiate and segment landslides with a superior degree of accuracy than the conventional U-Net model. The improved U-Net architecture, bolstered by SBConv, providing a robust tool for the detection of landslides from multi-spectral remote sensing data.

2.2.5. Spatial Refined Unit (SRU)

To deal with the spatial redundancy of features, we introduce the Spatial Refined Unit (SRU) proposed by [43], as depicted in Figure 5.

The SRU adopts a separate-restructure approach to mitigate the spatial redundancy of features, and its model structure is depicted in Figure 5. The purpose of the Separate operation is to separate the more informative feature maps, especially in terms of their spatial content. In Group Normalization (GN) layers, the scaling factors are used to evaluate the informative content of various feature maps.

Specifically, consider an feature map $X \in \mathbb{R}^{N \times C \times H \times W}$, in which N and C represent the batch and channel axis, and H and W represent the spatial height and width axes, respectively. To standardize X , subtract the mean μ and divide by the standard deviation σ . This process is achieved by the following formula:

$$X_{\text{out}} = \text{GN}(X) = \gamma \frac{X - \mu}{\sqrt{\sigma^2 + \varepsilon}} + \beta \quad (1)$$

A small positive constant ε is added to ensure numerical stability during division. The parameters γ and β are trainable factors associated with the affine transformation applied during normalization.

It is noteworthy that $\gamma \in \mathbb{R}^C$ is a trainable parameter in GN layers to measure the spatial pixels variance of batches and channels. A greater variance in spatial pixels, indicative of richer spatial information, corresponds to a larger value of γ . To signify the importance of feature maps, the normalized correlation weight $W_\gamma \in \mathbb{R}^C$ is derived using Equation (2).

$$W_\gamma = \{w_i\} = \frac{\gamma_i}{\sum_{j=1}^C \gamma_j}, i, j = 1, 2, \dots, C \quad (2)$$

Subsequently, W_γ are normalized to the range 0 to 1 using the sigmoid function and a predefined threshold. We assign a value of 1 to those weights exceeding the threshold, thereby defining the informative weights W_1 , while weights below the threshold are set to 0, resulting in the non-informative weights W_2 . In our experiments, this threshold is established at 0.5. W is succinctly represented by Equation (3):

$$W = \text{Gate}(\text{Sigmoid}(W_\gamma(\text{GN}(X)))) \quad (3)$$

In the final step, the input features X are multiplied by W_1 and W_2 respectively, resulting in two distinct sets of weighted features: X_1^w , which are informative and contain expressive spatial content, and X_2^w , which are less informative and considered to have minimal or no relevant information, hence deemed as redundant. This process effectively segregates

the input features into two categories: X_1^w representing the significant, information-rich features, and X_2^w encapsulating the less critical, potentially superfluous data.

To mitigate spatial redundancy, we further propose a Reconstruct operation. This operation involves summing features rich in information with those less informative, aiming to synthesize features that encapsulate richer information while conserving spatial resources. Rather than simply adding these two categories of features, we employ a cross reconstruct operation. This method is designed to effectively amalgamate the weighted, differing informative features, thereby enhancing the flow of information between them. Then, the features X^{w1} and X^{w2} are cross-reconstructed to yield the spatial-refined feature maps X^w . The entire Reconstruct operation can be articulated as follows:

$$\begin{cases} X_1^w = W_1 \otimes X \\ X_2^w = W_2 \otimes X \\ X_{11}^w \oplus X_{22}^w = X^{w1} \\ X_{21}^w \oplus X_{12}^w = X^{w2} \\ X^{w1} \cup X^{w2} = X^w \end{cases} \quad (4)$$

where \otimes , \oplus and \cup are element-wise multiplication, summation, and concatenation, respectively.

Symbols used in the figure and their corresponding operations include summation (\oplus), multiplication (\otimes), concatenation (C), group normalization (GN), weighting (w_i), thresholding (T), and sigmoid activation (S). The SRU's design is tailored to improve the network's ability to capture and emphasize spatial details, thereby facilitating more accurate segmentation in complex images, such as those used in landslide detection tasks.

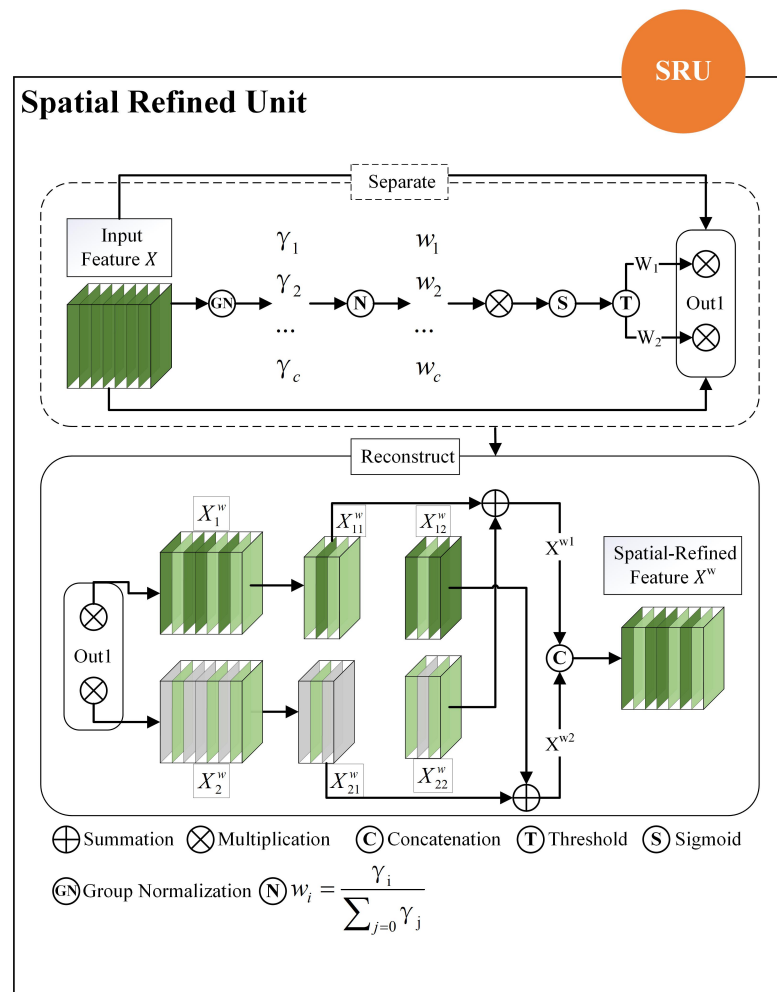


Figure 5. The structure of the SRU module.

2.2.6. Band Refined Unit (BRU)

The BRU employs a Split-Transform-Fuse strategy to address the issue of band redundancy in feature maps, depicted in Figure 6. Conventionally, feature extraction involves the use of repetitive standard $k \times k$ convolutions, often leading to the generation of somewhat redundant feature maps along the band dimension. Thus, we use $M^k \in \mathbb{R}^{c \times k \times k}$ instead of the $k \times k$ convolution kernel, and let X and Y denote the input and convolved output features, respectively, where $X, Y \in \mathbb{R}^{c \times h \times w}$.

A standard convolution, can be defined as $Y = M^k \cdot X$. In our approach, we substitute the standard convolution with the BRU, which is realized through three distinct operations: Split, Transform, and Fuse. The BRU operates through a multi-stage process detailed as follows:

Split: For a given set of spatial-refined features $X_w \in \mathbb{R}^{c \times h \times w}$, we initially divide its channels into two segments, consisting of αC and $(1 - \alpha)C$ channels, as illustrated in the splitting section of Figure 6. Here, α represents the split ratio, where $0 \leq \alpha \leq 1$. Subsequently, to enhance computational efficiency, we apply 1×1 convolutions to compress the channel dimensions of these feature maps. This compression introduces a ‘squeeze ratio’ r , which is utilized to regulate the feature channels, thereby balancing the computational load of the BRU (with $r = 2$). Following the split and squeeze operations, X_w are segregated into two components: X_{up} and X_{low} .

Transform: X_{up} is channeled into the upper transformation stage, functioning as a “Rich Feature Extractor”. To extract high-level representative information while simultaneously reducing computational demands, we adopt more efficient convolutional operations, namely Group-Wise Convolution (GWC) and Point-Wise Convolution (PWC) instead of the standard $k \times k$ convolutions. The sparse convolution connections in GWC lead to a reduction in the number of parameters and computational requirements, albeit at the cost of impeding the information flow between channel groups. PWC, in contrast, addresses this information loss and facilitates cross-channel information flow. Therefore, we apply both $k \times k$ GWC (setting the group size $g = 2$ in our experiments) and 1×1 PWC to X_{up} .

The outputs of these operations are then summed up to create a merged representative feature map Y_1 , as depicted in the Transform section of Figure 6. The process of the upper transformation stage can be expressed as follows:

$$Y_1 = M^G X_{up} + M^{P1} X_{up} \quad (5)$$

where $M^G \in \mathbb{R}^{\frac{ac}{gr} \times k \times k \times c}$ and $M^{P1} \in \mathbb{R}^{\frac{ac}{r} \times 1 \times 1 \times c}$ represent the learnable weight matrices of GWC and PWC, respectively. Additionally, $X_{up} \in \mathbb{R}^{\frac{ac}{r} \times h \times w}$ and $Y_1 \in \mathbb{R}^{c \times h \times w}$ are the input and output feature maps of the upper part, respectively. In essence, the upper transformation stage employs a synergistic combination of GWC and PWC on the same X_{up} . This approach is designed to extract rich representative features Y_1 , while maintaining a lower computational footprint.

X_{low} is directed into the lower transformation stage. At this stage, we employ cost-effective 1×1 PWC to generate feature maps, which could be used to reveal shallow hidden details and serve as a complement to the “Rich Feature Extractor” of the upper transformation stage. Additionally, we repurpose the features in X_{low} to derive more feature maps, thereby augmenting our feature set without incurring additional computational costs. To form the output Y_2 of the lower stage, both the newly generated and reused features are concatenated in the final step of this stage, as follows:

$$Y_2 = M^{P2} X_{low} \cup X_{low} \quad (6)$$

In this equation, $M^{P2} \in \mathbb{R}^{\frac{(1-\alpha)c}{r} \times 1 \times 1 \times (1-\frac{\alpha}{r})c}$ is weight matrix that could be learned for PWC. The operation \cup signifies concatenation. The input feature map $X_{low} \in \mathbb{R}^{\frac{(1-\alpha)c}{r} \times h \times w}$ undergoes a transformation via PWC, and then the resulting feature map is concatenated with the original input X_{low} . This process effectively combines the transformed and original

features, resulting in the final output feature map $Y_2 \in \mathbb{R}^{c \times h \times w}$, which encompasses supplementary detailed information extracted from the lower stage.

Fuse: Following the transformation stages, a simplified version of the SKNet method [48] is employed to merge the output features Y_1 and Y_2 from the upper and lower transformation stages, as illustrated in the Fuse section of Figure 6. Initially, we apply global average pooling (referred to as Pooling) to aggregate global spatial information. This process generates channel-wise statistics $S_m \in \mathbb{R}^{c \times 1 \times 1}$, which are computed as follows:

$$S_m = \text{Pooling}(Y_m) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W Y_c(i, j), m = 1, 2 \tag{7}$$

Next, we combine the global channel-wise descriptors S_1 and S_2 from the upper and lower stages, respectively, and proceed to apply a channel-wise soft attention operation. This operation is designed to obtain the feature importance vectors $\beta_1, \beta_2 \in \mathbb{R}^c$, which are computed as follows:

$$\beta_1 = \frac{e^{s_1}}{e^{s_1} + e^{s_2}}, \beta_2 = \frac{e^{s_2}}{e^{s_1} + e^{s_2}}, \beta_1 + \beta_2 = 1 \tag{8}$$

Finally, guided by the feature importance vectors β_1 and β_2 , the channel-refined features Y are derived by channel-wise merging of Y_1 and Y_2 . This merging process can be expressed as follows:

$$Y = \beta_1 Y_1 + \beta_2 Y_2 \tag{9}$$

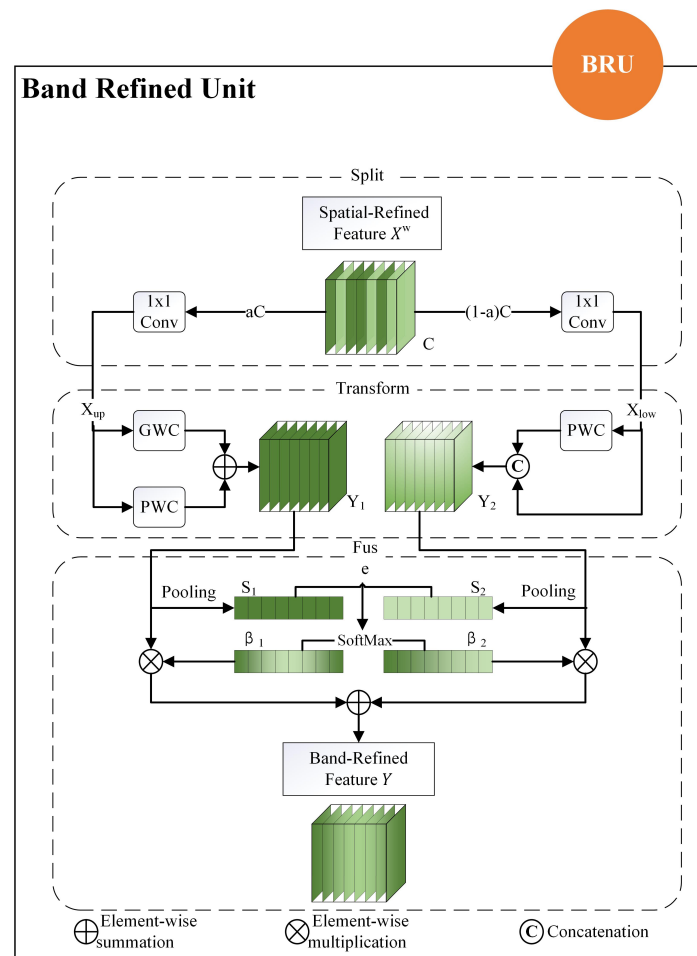


Figure 6. The structure of the BRU module.

In summary, the BRU is utilized, following a Split-Transform-and-Fuse strategy, to reduce the redundancy of X^w . The BRU effectively extracts representative features via lightweight convolutional operations and handles redundant features through cost-effective operations and feature reuse strategies. Significantly, the BRU can function independently or be synergistically combined with the SRU operation. By sequentially integrating the SRU and BRU, we establish the proposed SBConv, a highly efficient architecture that serves as a viable alternative to standard convolution operations.

Symbols in the figure denote element-wise summation (\oplus), element-wise multiplication (\otimes), and concatenation (C), indicating the operations performed at each stage of the BRU. This unit is integral to our network's ability to discern and enhance spectral features, further contributing to the robustness of the landslide detection task.

2.3. Model Evaluation

Precision, recall, overall accuracy (OA) and F1 score (F1) are computed to evaluate the performance of the proposed model. These 5 indexes are defined as:

$$\text{precision} = \frac{TP}{TP + FP} \quad (10)$$

$$\text{recall} = \frac{TP}{TP + FN} \quad (11)$$

$$F1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (12)$$

$$OA = \frac{TP + TN}{FP + FN + TP + TN} \quad (13)$$

In this context, TP represents the count of true positives, which are the pixels accurately identified as landslides. TN denotes the true negatives, referring to the pixels correctly classified as non-landslide areas. FP is the number of false positives, indicating pixels that are actually non-landslide (background) but erroneously classified as landslides. Conversely, FN stands for the false negatives, which are the pixels that should be classified as landslides according to ground truth but are mislabeled as non-landslide areas. Precision is a measure that quantifies the accuracy of the landslide predictions, signifying the proportion of pixels correctly classified as landslides out of all pixels predicted as landslides. Recall, on the other hand, represents the ability of the model to correctly identify landslide pixels, calculated as the proportion of pixels correctly classified as landslides out of the total ground truth landslide pixels. F1 score is considered the harmonic mean of precision and recall. This value provides a comprehensive measure of overall performance in accurately classifying landslides by balancing precision and recall [13].

3. Results

3.1. Experimental Settings

In the course of this study, a total of 3799 image patches were employed for the training of various models. Each experimental model underwent training across 5000 epochs to ensure thorough learning. A comparative analysis was conducted using several architectures: standard U-Net, SBConv-Unet, BruU-Net, SruU-Net, ResU-Net, SBConvResU-Net, BruResU-Net, and SruResU-Net, all applied to the identical dataset. Notably, the training of these models was executed without the aid of any pre-trained models, ensuring an unbiased learning process. The implementation of these experiments was carried out using the PyTorch framework, a popular choice for its flexibility and efficiency in deep learning tasks. The hardware setup for these experiments included a personal computer equipped with an Intel 12400F CPU and an NVIDIA GeForce RTX 4060Ti GPU, supported by 16 GB of memory.

3.2. Experimental Results

As detailed in Table 1, our study highlights the superior performance of the proposed network in landslide detection, as evidenced by quantitative comparisons. Specifically, the SBConvU-Net, our proposed model, demonstrated noteworthy performance in terms of recall and F1 scores, achieving 75.74% and 70.89%, respectively. In the task of landslide detection or landslide susceptibility mapping, the model's ability to identify landslides (the positive class) is of paramount importance [16]. Therefore, a higher recall and a comparatively high F1 score significantly indicate the practicality and efficacy of the model. This is particularly significant when contrasted with other models used for comparison; for instance, the ResU-Net produced the lowest recall and F1 scores among all models evaluated.

Table 1. Comparison results of different models for landsilide detection.

Model	Precision	Recall	F1	Mean F1	OA
U-Net	85.47	60.59	70.91	85.16	98.85
SruU-Net	81.57	64.89	72.28	85.85	98.85
BruU-Net	73.75	68.35	70.95	85.14	98.7
SBConvU-Net	72.13	75.74	73.89	86.63	98.76
ResU-Net	85.85	54.75	66.86	83.11	98.74
SruResU-Net	83.75	58.89	69.16	84.27	98.78
BruResU-Net	80.69	59.24	68.32	83.84	98.73
SBConvResU-Net	76.82	70.98	73.78	86.59	98.83

Notably, while the SBConvU-Net secured the highest F1 score, it did record a relatively modest precision score of 72.13%. This is acceptable in the context of landslide extraction, where sample imbalance is a common challenge. A higher recall and F1 score are desirable attributes in models for this application, as they indicate a stronger capability to correctly identify a greater number of landslide pixels compared to non-landslide pixels. Further analysis of the performance across different network architectures—U-Net, SruU-Net, BruU-Net, and SBConvU-Net—reveals that the integration of the BRU and SRU modules can enhance the U-Net model's effectiveness. Of particular note is the Mixed module, which exhibited the most robust capability for landslide extraction. This finding emphasizes the value of incorporating these modules into the U-Net framework to achieve improved performance in landslide detection tasks.

3.3. Prediction with Different Models

To further evaluate the effectiveness of the proposed method, this study randomly selected 5 image samples and conducted ablation experiments. We calculated the results of using U-Net (ResU-Net) alone, U-Net with the added SRU module (SruU-Net or SruResU-Net), U-Net with the added BRU module (BruU-Net or BruResU-Net), and U-Net with both the SRU and BRU modules added (SBConvU-Net or SBConvResU-Net). Figure 7 intuitively illustrates the landslide detection results achieved by each method. In the figure, the first column displays the true-color remote sensing images, the second column shows the Ground Truth provided by the dataset used in the experiment, and the third to tenth columns show the prediction results of different methods, with blue representing the Ground Truth and red indicating the landslide areas predicted by the model. By observing the first row as an illustrative example, we can notice significant differences in the detection capabilities of various models. Traditional models like U-Net and ResU-Net rely on more conventional convolutional neural network architectures, thus can only detect relatively small landslide areas. This limitation can be attributed to their standard feature extraction mechanism, which may not capture the subtle differences often present in complex geographical terrains.

In contrast, models incorporating Spatial Refinement Units (SRU) or Band Refinement Units (BRU), whether integrated into the U-Net or ResU-Net framework, demonstrate significantly enhanced capabilities, capable of identifying larger and more subtle landslide

areas affected by spatial influences. This improvement indicates the advanced feature aggregation and analysis capabilities endowed by SRU and BRU modules, enabling the models to discern finer details and variations in the landscape, thus more accurately identifying landslides. Among all evaluated models, SBConvU-Net and SBConvResU-Net stand out for their outstanding performance. These models equipped with SRU and BRU enhancement functionalities showcase the ability to detect the widest and most detailed landslide areas.

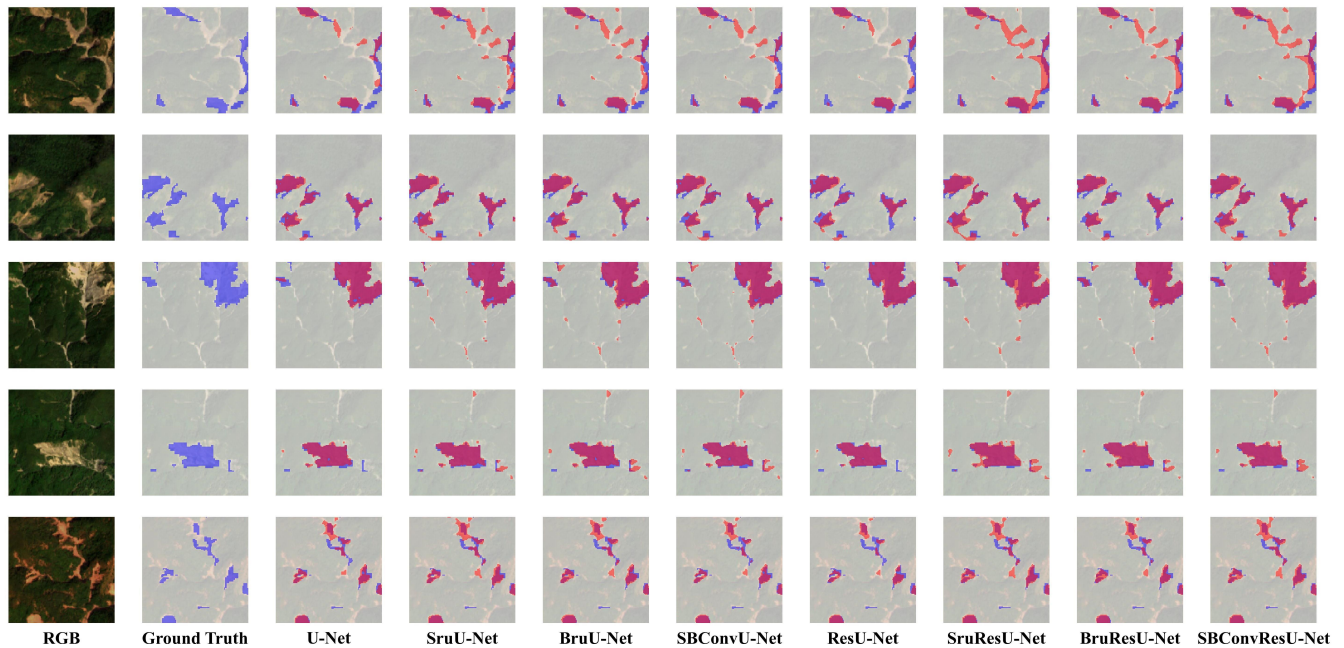


Figure 7. Visualization comparison of prediction results of various models. Red represents the predicted landslide area, and blue represents the label data (ground truth) provided by the dataset.

4. Discussion

The rapid development of deep learning technologies has spurred growth in the field of computer vision. Architectures such as CNN, RNN, LSTM, U-Net, and ResNet have been widely applied to various tasks including image classification, recognition, and semantic segmentation. Remote sensing images, characterized by multiple bands, spatiotemporal resolutions, and sensors, have also benefited from advancements in computer vision and deep learning.

U-Net, a convolutional neural network (CNN) based fully convolutional network, is particularly effective for image segmentation tasks due to its encoder-decoder structure and skip connections. It excels in extracting features and identifying targets in remote sensing images [40]. ResNet addresses the vanishing gradient problem in deep networks by introducing residual connections, allowing for the effective training of deeper networks. This makes it highly effective for classifying various land cover types in remote sensing images, such as urban areas, forests, and agricultural fields [49]. CNNs, as the foundational architecture of deep learning, are widely applied in remote sensing image processing tasks, showing exceptional performance in image classification, object detection, and scene recognition due to their layered convolution and pooling design [50]. RNNs and LSTMs, known for their prowess in handling sequential data, are crucial in spatiotemporal data analysis in remote sensing. They are applied in tasks such as environmental monitoring, climate prediction, and disaster assessment [51].

In conclusion, deep learning technologies have significantly advanced the processing of remote sensing images. The successful application of various neural network architectures in image classification, segmentation, and semantic recognition has enhanced

the capabilities for processing and analyzing remote sensing data, providing robust technical support for fields such as environmental monitoring, resource management, and disaster assessment.

This study presents the integration of the Spatial and Band Refinement Convolution (SBConv) module into the U-Net architecture, enhancing landslide detection capabilities using multisource remote sensing data. Compared to existing methods, our SBConv-embedded U-Net models demonstrate significant improvements in performance metrics such as recall and F1 score, which are critical for accurate and reliable landslide detection.

4.1. Comparison with Existing Approaches

Recent advancements in landslide detection have primarily leveraged conventional convolutional architectures like standard U-Net and its derivatives, which, while effective, often fail to address the complexity and heterogeneity of multispectral remote sensing data used in detecting landslides [14,35]. Our approach differs by incorporating the SBConv module, which systematically reduces redundancy in spatial and spectral feature processing. For instance, our SBConvU-Net model achieves a recall of 75.74% and an F1 Score of 73.89%, outperforming the base U-Net's recall of 60.59% and F1 Score of 70.91%. This indicates a substantial enhancement in detecting true positives, a crucial aspect often overlooked in previous studies.

To further compare the model proposed in this study with results obtained by other researchers using the same dataset, we present a comparison between our results and those found in Reference [44] in Table 2.

Table 2. Quantitative Results of Different Deep Neural Networks for the Landslide Detection (%).

Model	Recall	Precision	F1
PSPNet	52.03	61.55	56.39
ContextNet	49.29	70.77	58.11
DeepLab-v2	63.68	60.8	62.21
DeepLab-v3+	62.11	69.91	65.78
FCN-8s	63.05	68.66	65.73
LinkNet	67.02	66.76	66.89
FRRN-A	64.4	76.57	69.96
FRRN-B	76.16	64.93	70.1
SQNet	66.69	74.2	70.24
SBConvU-Net	75.74	72.13	73.89
SBConvResU-Net	70.98	76.82	73.78

The table includes models such as PSPNet, ContextNet, various DeepLab versions, FCN-8s, LinkNet, and FRRN types, which were employed in [44], alongside the SBConvU-Net and SBConvResU-Net proposed in this study. These latter models demonstrate a marked improvement in performance metrics. The SBConvU-Net and SBConvResU-Net exhibit superior performance with F1 Scores of 73.89% and 73.78%, respectively. These models surpass other approaches, illustrating their capability to effectively balance precision and recall within this specific application context. The enhanced performance of these models emphasizes the potential of specialized architectures to significantly enhance both the accuracy and reliability of landslide detection systems.

4.2. Broader Implications and Future Work

The implications of our findings extend beyond academic research into practical applications, including enhanced real-time monitoring and early warning systems for landslide-prone areas. By improving the accuracy of landslide detection, our methodology supports more informed decision-making in disaster management and mitigation strategies.

However, despite these advancements, our study has limitations that warrant further research. The current SBConv module focuses predominantly on spatial and spectral data without integrating other environmental factors like soil moisture content and terrain

stability, which can also influence landslide occurrences. Future studies could explore incorporating these variables to provide a more comprehensive assessment of landslide risks.

Additionally, ongoing advancements in deep learning architectures may offer further opportunities to enhance the SBConv module. Exploring the integration of newer neural network architectures or advanced regularization techniques could potentially lead to even higher performance gains.

The method proposed contributes significantly to the field of remote sensing and landslide detection by introducing a refined convolutional approach that optimizes both spatial and spectral feature processing. By setting a new benchmark in the performance metrics of landslide detection models, this research paves the way for future investigations into more sophisticated and comprehensive approaches for environmental monitoring.

5. Conclusions

Our research presents a transformative approach to landslide detection through the innovative application of the SBConv module within the U-Net architecture. The empirical evaluation of our models demonstrates not only a reduction in feature redundancy but also a substantial improvement in detecting landslide events, as evidenced by the performance metrics.

The strategic enhancement of spatial and spectral feature processing capabilities has proven to be a decisive factor in improving the accuracy and reliability of landslide detection. Our improved U-Net architecture, augmented with the SBConv module, offers a more effective tool for environmental monitoring.

This study contributes to the body of knowledge in remote sensing technology for landslide identification and monitoring. The proposed model's performance indicates a promising direction for future research in applying advanced CNN architectures to environmental surveillance and analysis. The success of the SBConv-U-Net model in this study encourages further exploration into the integration of specialized convolutional modules for enhancing feature recognition in various remote sensing applications.

While our SBConv integrated U-Net model demonstrates significant advancements in landslide detection, there are several aspects that merit further investigation: One limitation of the current study lies in the data processing approach. Our model primarily focused on spatial and spectral features but did not fully incorporate other critical topographic factors. Future research should include additional variables such as slope aspect, terrain curvature, and surface roughness. These factors could provide a more comprehensive understanding of landslide dynamics and improve detection accuracy. Moreover, the controlling and triggering factors of landslides require deeper exploration. While our model efficiently identifies landslide occurrences, understanding the underlying causes and mechanisms is essential for predictive analysis and risk management. Integrating geological, hydrological, and climatic data could offer valuable insights into the factors that contribute to landslide susceptibility and occurrence. Looking ahead, there is significant scope for enhancing the model by integrating a broader range of environmental variables and exploring more sophisticated deep learning techniques. The ongoing advancement of convolutional neural network (CNN) architectures offers a promising avenue for crafting more sophisticated and resilient models tailored for geospatial analysis. Such progress holds considerable potential to enhance the effectiveness of landslide monitoring efforts.

In conclusion, our study lays the groundwork for future advancements in automated landslide detection using deep learning approaches. By addressing the identified limitations and exploring the outlined future directions, we can move towards more accurate, reliable, and comprehensive landslide monitoring systems.

Author Contributions: Conceptualization, Y.S. and Y.H.; methodology, Y.S. and S.X.; software, Y.Z.; validation, Y.Z. and Y.L.; formal analysis, W.W. and R.N.; investigation, Y.L.; resources, Y.S.; data curation, Y.L.; writing—original draft preparation, Y.L.; writing—review and editing, Y.S., Y.L. and S.X.; visualization, Y.L.; supervision, Y.L.; project administration, Y.H.; funding acquisition, Y.S. All authors have read and agreed to the published version of the manuscript.

Funding: The study was jointly funded by Open Fund from Engineering Research Center for Seismic Disaster Prevention and Engineering Geological Disaster Detection of Jiangxi Province (No. SDGD202203), Open Fund from Key Laboratory for Digital Land and Resources of Jiangxi Province, East China University of Technology (No. DLLJ202204), and Research on Key Technologies of Robot Vision Perception and Dexterous Operations for 3C Manufacturing (No. 20232ABC03A09).

Data Availability Statement: The multi-source landslide benchmark data (Landslide4Sense) are publicly available at <https://github.com/iarai/Landslide4Sense-2022> (accessed on 31 December 2023). The ground truth of Iburi study area can be downloaded at the website of <https://zenodo.org/records/2577300#.YzKiSXZByUI> (accessed on 31 December 2023).

Acknowledgments: We would like to express our sincere gratitude to the funding agencies that supported this research. The code of this article refers to <https://github.com/cheng-haha/ScConv/tree/main> (accessed on 31 December 2023) and <https://github.com/iarai/Landslide4Sense-2022> (accessed on 31 December 2023). We would also thank the Institute of Advanced Research in Artificial Intelligence (IARAI) for sharing the landslide and remote sensing data.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

BRU	Band Refined Unit
CNN	Convolutional Neural Network
DEM	Digital Elevation Model
DL	Deep Learning
IDCU	Improved Double Convolution Unit
InSAR	Interferometric Synthetic Aperture Radar
LiDAR	Light Detection And Ranging
ML	Machine Learning
SAR	Synthetic Aperture Radar
SBCConv	Spatial and Band Refined Convolution
SRU	Spatial Refined Unit

References

- Arrogante-Funes, P.; Bruzón, A.G.; Álvarez Ripado, A.; Arrogante-Funes, F.; Martín-González, F.; Novillo, C.J. Assessment of the Regeneration of Landslides Areas Using Unsupervised and Supervised Methods and Explainable Machine Learning Models. *Landslides* **2024**, *21*, 275–290. [\[CrossRef\]](#)
- Hong, H.; Liu, J.; Bui, D.T.; Pradhan, B.; Acharya, T.D.; Pham, B.T.; Zhu, A.X.; Chen, W.; Ahmad, B.B. Landslide Susceptibility Mapping Using J48 Decision Tree with AdaBoost, Bagging and Rotation Forest Ensembles in the Guangchang Area (China). *Catena* **2018**, *163*, 399–413. [\[CrossRef\]](#)
- Lukić, T.; Bjelajac, D.; Fitzsimmons, K.E.; Marković, S.B.; Basarin, B.; Mlađan, D.; Micić, T.; Schaetzl, R.J.; Gavrilov, M.B.; Milanović, M.; et al. Factors Triggering Landslide Occurrence on the Zemun Loess Plateau, Belgrade Area, Serbia. *Environ. Earth Sci.* **2018**, *77*, 519. [\[CrossRef\]](#)
- Li, X.; Wang, L.; Hong, B.; Li, L.; Liu, J.; Lei, H. Erosion Characteristics of Loess Tunnels on the Loess Plateau: A Field Investigation and Experimental Study. *Earth Surf. Process. Landforms* **2020**, *45*, 1945–1958. [\[CrossRef\]](#)
- Meng, Z.J.; Ma, P.H.; Peng, J.B. Characteristics of Loess Landslides Triggered by Different Factors in the Chinese Loess Plateau. *J. Mt. Sci.* **2021**, *18*, 3218–3229. [\[CrossRef\]](#)
- Wang, H.B.; Sassa, K. Rainfall-Induced Landslide Hazard Assessment Using Artificial Neural Networks. *Earth Surf. Process. Landforms* **2006**, *31*, 235–247. [\[CrossRef\]](#)
- Wu, T.; Xie, X.; Wu, H.; Zeng, H.; Zhu, X. A Quantitative Analysis Method of Regional Rainfall-Induced Landslide Deformation Response Variation Based on a Time-Domain Correlation Model. *Land* **2022**, *11*, 703. [\[CrossRef\]](#)
- Liu, G.; Zhou, Z.; Xu, S.; Cheng, Y. Post Evaluation of Slope Cutting on Loess Slopes under Long-Term Rainfall Based on a Model Test. *Sustainability* **2022**, *14*, 15838. [\[CrossRef\]](#)
- Jing, J.; Wu, Z.; Yan, W.; Ma, W.; Liang, C.; Lu, Y.; Chen, D. Experimental Study on Progressive Deformation and Failure Mode of Loess Fill Slopes under Freeze-Thaw Cycles and Earthquakes. *Eng. Geol.* **2022**, *310*, 106896. [\[CrossRef\]](#)
- Chen, Y.; Wei, Y.; Wang, Q.; Chen, F.; Lu, C.; Lei, S. Mapping Post-Earthquake Landslide Susceptibility: A U-net like Approach. *Remote. Sens.* **2020**, *12*, 2767. [\[CrossRef\]](#)

11. Shafapourtehrany, M.; Rezaie, F.; Jun, C.; Heggy, E.; Bateni, S.M.; Panahi, M.; Özener, H.; Shabani, F.; Moeini, H. Mapping Post-Earthquake Landslide Susceptibility Using U-net, VGG-16, VGG-19, and Metaheuristic Algorithms. *Remote Sens.* **2023**, *15*, 4501. [[CrossRef](#)]
12. Wang, J.; Liu, Y.; Zhang, G.; Hu, X.; Xing, B.; Wang, D. Reservoir Landslide Displacement Prediction under Rainfall Based on the ILF-FFT Method. *Bull. Eng. Geol. Environ.* **2023**, *82*, 179. [[CrossRef](#)]
13. Lu, W.; Hu, Y.; Zhang, Z.; Cao, W. A Dual-Encoder U-net for Landslide Detection Using Sentinel-2 and DEM Data. *Landslides* **2023**, *20*, 1975–1987. [[CrossRef](#)]
14. Zhou, Y.; Wang, H.; Yang, R.; Yao, G.; Xu, Q.; Zhang, X. A Novel Weakly Supervised Remote Sensing Landslide Semantic Segmentation Method: Combining CAM and cycleGAN Algorithms. *Remote Sens.* **2022**, *14*, 3650. [[CrossRef](#)]
15. Ge, X.; Zhao, Q.; Wang, B.; Chen, M. Lightweight Landslide Detection Network for Emergency Scenarios. *Remote Sens.* **2023**, *15*, 1085. [[CrossRef](#)]
16. Xu, S.; Song, Y.; Hao, X. A Comparative Study of Shallow Machine Learning Models and Deep Learning Models for Landslide Susceptibility Assessment Based on Imbalanced Data. *Forests* **2022**, *13*, 1908. [[CrossRef](#)]
17. Cabral, V.; Reis, F.; Veloso, V.; Ogura, A.; Zarfl, C. A Multi-Step Hazard Assessment for Debris-Flow Prone Areas Influenced by Hydroclimatic Events. *Eng. Geol.* **2023**, *313*, 106961. [[CrossRef](#)]
18. Zhang, Y.; Ayyub, B.M.; Gong, W.; Tang, H. Risk Assessment of Roadway Networks Exposed to Landslides in Mountainous Regions—A Case Study in Fengjie County, China. *Landslides* **2023**, *20*, 1419–1431. [[CrossRef](#)]
19. Zhang, X.; Li, Y.; Liu, Y.; Huang, Y.; Wang, Y.; Lu, Z. Characteristics and Prevention Mechanisms of Artificial Slope Instability in the Chinese Loess Plateau. *Catena* **2021**, *207*, 105621. [[CrossRef](#)]
20. Adnan, M.S.G.; Rahman, M.S.; Ahmed, N.; Ahmed, B.; Rabbi, M.F.; Rahman, R.M. Improving Spatial Agreement in Machine Learning-Based Landslide Susceptibility Mapping. *Remote Sens.* **2020**, *12*, 3347. [[CrossRef](#)]
21. An, B.; Wang, C.; Liu, C.; Li, P. A Multi-Source Remote Sensing Satellite View of the February 22nd Xinjing Landslide in the Mining Area of Alxa Left Banner, China. *Landslides* **2023**, *20*, 2517–2523. [[CrossRef](#)]
22. Abbas, F.; Zhang, F.; Abbas, F.; Ismail, M.; Iqbal, J.; Hussain, D.; Khan, G.; Alrefaei, A.F.; Albeshr, M.F. Landslide Susceptibility Mapping: Analysis of Different Feature Selection Techniques with Artificial Neural Network Tuned by Bayesian and Metaheuristic Algorithms. *Remote Sens.* **2023**, *15*, 4330. [[CrossRef](#)]
23. Abraham, M.T.; Vaddapally, M.; Satyam, N.; Pradhan, B. Spatio-Temporal Landslide Forecasting Using Process-Based and Data-Driven Approaches: A Case Study from Western Ghats, India. *Catena* **2023**, *223*, 106948. [[CrossRef](#)]
24. Arabameri, A.; Karimi-Sangchini, E.; Pal, S.C.; Saha, A.; Chowdhuri, I.; Lee, S.; Tien Bui, D. Novel Credal Decision Tree-Based Ensemble Approaches for Predicting the Landslide Susceptibility. *Remote Sens.* **2020**, *12*, 3389. [[CrossRef](#)]
25. Cai, J.; Zhang, L.; Dong, J.; Guo, J.; Wang, Y.; Liao, M. Automatic Identification of Active Landslides over Wide Areas from Time-Series InSAR Measurements Using Faster RCNN. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *124*, 103516. [[CrossRef](#)]
26. Chen, H.; Zhao, C.; Li, B.; Gao, Y.; Chen, L.; Liu, D. Monitoring Spatiotemporal Evolution of Kaiyang Landslides Induced by Phosphate Mining Using Distributed Scatterers InSAR Technique. *Landslides* **2023**, *20*, 695–706. [[CrossRef](#)]
27. Ciampalini, A.; Raspini, F.; Lagomarsino, D.; Catani, F.; Casagli, N. Landslide Susceptibility Map Refinement Using PSInSAR Data. *Remote Sens. Environ.* **2016**, *184*, 302–315. [[CrossRef](#)]
28. Stumpf, A.; Kerle, N. Combining Random Forests and Object-Oriented Analysis for Landslide Mapping from Very High Resolution Imagery. *Procedia Environ. Sci.* **2011**, *3*, 123–129. [[CrossRef](#)]
29. Mladenova, I.E.; Bolten, J.D.; Crow, W.T.; Anderson, M.C.; Hain, C.R.; Johnson, D.M.; Mueller, R. Intercomparison of Soil Moisture, Evaporative Stress, and Vegetation Indices for Estimating Corn and Soybean Yields Over the U.S. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 1328–1343. [[CrossRef](#)]
30. Jaboyedoff, M.; Oppikofer, T.; Abellán, A.; Derron, M.H.; Loye, A.; Metzger, R.; Pedrazzini, A. Use of LIDAR in Landslide Investigations: A Review. *Nat. Hazards* **2012**, *61*, 5–28. [[CrossRef](#)]
31. Rossi, M.; Kirschbaum, D.; Luciani, S.; Mondini, A.C.; Guzzetti, F. TRMM Satellite Rainfall Estimates for Landslide Early Warning in Italy: Preliminary Results. In Proceedings of the Remote Sensing of the Atmosphere, Clouds, and Precipitation IV, Kyoto, Japan, 29 October–1 November 2012; Hayasaka, T., Nakamura, K., Im, E., Eds.; SPIE: Bellingham, WA, USA, 2012. [[CrossRef](#)]
32. Lagomarsino, D.; Segoni, S.; Fanti, R.; Catani, F. Updating and Tuning a Regional-Scale Landslide Early Warning System. *Landslides* **2013**, *10*, 91–97. [[CrossRef](#)]
33. Marjanović, M.; Kovačević, M.; Bajat, B.; Voženílek, V. Landslide Susceptibility Assessment Using SVM Machine Learning Algorithm. *Eng. Geol.* **2011**, *123*, 225–234. [[CrossRef](#)]
34. Zeng, F.; Nait Amar, M.; Mohammed, A.S.; Motahari, M.R.; Hasanipanah, M. Improving the Performance of LSSVM Model in Predicting the Safety Factor for Circular Failure Slope through Optimization Algorithms. *Eng. Comput.* **2022**, *38*, 1755–1766. [[CrossRef](#)]
35. Wang, J.; Chen, G.; Jaboyedoff, M.; Derron, M.H.; Fei, L.; Li, H.; Luo, X. Loess Landslides Detection via a Partially Supervised Learning and Improved Mask-RCNN with Multi-Source Remote Sensing Data. *Catena* **2023**, *231*, 107371. [[CrossRef](#)]
36. Zhang, W.; Li, H.; Han, L.; Chen, L.; Wang, L. Slope Stability Prediction Using Ensemble Learning Techniques: A Case Study in Yunyang County, Chongqing, China. *J. Rock Mech. Geotech. Eng.* **2022**, *14*, 1089–1099. [[CrossRef](#)]
37. Huang, P.C. Establishing a Shallow-Landslide Prediction Method by Using Machine-Learning Techniques Based on the Physics-Based Calculation of Soil Slope Stability. *Landslides* **2023**, *20*, 2741–2756. [[CrossRef](#)]

38. Huang, R.; Chen, T. Landslide Recognition from Multi-Feature Remote Sensing Data Based on Improved Transformers. *Remote Sens.* **2023**, *15*, 3340. [[CrossRef](#)]
39. Chen, W.; Chen, Y.; Tsangaratos, P.; Ilia, I.; Wang, X. Combining Evolutionary Algorithms and Machine Learning Models in Landslide Susceptibility Assessments. *Remote Sens.* **2020**, *12*, 3854. [[CrossRef](#)]
40. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*; Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Volume 9351, pp. 234–241. [[CrossRef](#)]
41. Chen, H.; He, Y.; Zhang, L.; Yao, S.; Yang, W.; Fang, Y.; Liu, Y.; Gao, B. A Landslide Extraction Method of Channel Attention Mechanism U-net Network Based on Sentinel-2A Remote Sensing Images. *Int. J. Digit. Earth* **2023**, *16*, 552–577. [[CrossRef](#)]
42. Li, G.; Zhou, X.; Chen, C.; Xu, L.; Zhou, F.; Shi, F.; Tang, J. Multitype Geomagnetic Noise Removal via an Improved U-Net Deep Learning Network. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5916512. [[CrossRef](#)]
43. Li, J.; Wen, Y.; He, L. SCConv: Spatial and Channel Reconstruction Convolution for Feature Redundancy. In Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada, 17–24 June 2023; pp. 6153–6162. [[CrossRef](#)]
44. Ghorbanzadeh, O.; Xu, Y.; Ghamisi, P.; Kopp, M.; Kreil, D. Landslide4Sense: Reference Benchmark Data and Deep Learning Models for Landslide Detection. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5633017. [[CrossRef](#)]
45. Ghorbanzadeh, O.; Xu, Y.; Zhao, H.; Wang, J.; Zhong, Y.; Zhao, D.; Zang, Q.; Wang, S.; Zhang, F.; Shi, Y.; et al. The Outcome of the 2022 Landslide4Sense Competition: Advanced Landslide Detection From Multisource Satellite Imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **2022**, *15*, 9927–9942. [[CrossRef](#)]
46. Zhang, Z.; Liu, Q.; Wang, Y. Road Extraction by Deep Residual U-Net. *IEEE Geosci. Remote. Sens. Lett.* **2018**, *15*, 749–753. [[CrossRef](#)]
47. Chen, M.; Dong, W.; Yu, H.; Woodhouse, I.; Ryan, C.M.; Liu, H.; Georgiou, S.; Mitchard, E.T.A. Multimodal Deep Learning for Mapping Forest Dominant Height by Fusing GEDI with Earth Observation Data. *arXiv* **2023**, arXiv:2311.11777. [[CrossRef](#)]
48. Li, X.; Wang, W.; Hu, X.; Yang, J. Selective Kernel Networks. *arXiv* **2019**, arXiv:1903.06586. [[CrossRef](#)]
49. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. *arXiv* **2015**, arXiv:1512.03385. [[CrossRef](#)]
50. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems*; Pereira, F., Burges, C., Bottou, L., Weinberger, K., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2012; Volume 25.
51. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.