

Article

Waterlogged Area Identification Models Based on Object-Oriented Image Analysis and Deep Learning Methods in Sloping Croplands of Northeast China

Peng Xie ^{1,2,3,4}, Shihang Wang ^{1,3,4}, Meiyan Wang ^{2,*} , Rui Ma ², Zhiyuan Tian ², Yin Liang ² and Xuezheng Shi ²

¹ School of Geomatics, Anhui University of Science and Technology, Huainan 232001, China; 201201626@aust.edu.cn (P.X.); 2011058@aust.edu.cn (S.W.)

² State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China; marui@issas.ac.cn (R.M.); tianzhiyuan@issas.ac.cn (Z.T.); yliang@issas.ac.cn (Y.L.); xzshi@issas.ac.cn (X.S.)

³ Key Laboratory of Aviation-Aerospace-Ground Cooperative Monitoring and Early Warning of Coal Mining-induced Disasters of Anhui Higher Education Institutes, Anhui University of Science and Technology, Huainan 232001, China

⁴ Coal Industry Engineering Research Center of Collaborative Monitoring of Mining Area's Environment and Disasters, Anhui University of Science and Technology, Huainan 232001, China

* Correspondence: mywang@issas.ac.cn

Abstract: Drainage difficulties in the waterlogged areas of sloping cropland not only impede crop development but also facilitate the formation of erosion gullies, resulting in significant soil and water loss. Investigating the distribution of these waterlogged areas is crucial for comprehending the erosion patterns of sloping cropland and preserving black soil resource. In this study, we built varied models based on two stages (one using only deep learning methods and the other combining object-based image analysis (OBIA) with deep learning methods) to identify waterlogged areas using high-resolution remote sensing data. The results showed that the five deep learning models using original remote sensing imagery achieved precision rates varying from 54.6% to 60.9%. Among these models, the DeepLabV3+-Xception model achieved the highest accuracy, as indicated by an *F1-score* of 53.4%. The identified imagery demonstrated a significant distinction in the two categories of waterlogged areas: sloping cropland erosion zones and erosion risk areas. The former had obvious borders and fewer misclassifications, exceeding the latter in terms of identification accuracy. Furthermore, the accuracy of the deep learning models was significantly improved when combined with object-oriented image analysis. The DeepLabV3+-MobileNetV2 model achieved the maximum accuracy, with an *F1-score* of 59%, which was 6% higher than that of the model using only original imagery. Moreover, this advancement mitigated issues related to boundary blurriness and image noise in the identification process. These results will provide scientific assistance in managing and reducing the impact in these places.

Keywords: high resolution; sloping cropland; waterlogged areas; object-oriented; deep learning



Citation: Xie, P.; Wang, S.; Wang, M.; Ma, R.; Tian, Z.; Liang, Y.; Shi, X. Waterlogged Area Identification Models Based on Object-Oriented Image Analysis and Deep Learning Methods in Sloping Croplands of Northeast China. *Sustainability* **2024**, *16*, 3917. <https://doi.org/10.3390/su16103917>

Academic Editor: Georgios Koubouris

Received: 23 March 2024

Revised: 30 April 2024

Accepted: 1 May 2024

Published: 8 May 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

Waterlogged areas in sloping cropland are sensitive regions for soil erosion. Research on the identification of these waterlogged areas holds significant practical importance for regional soil and water conservation as well as management. The undulating hilly regions in the northeast black soil area represent a typical terrain characterized by gentle and long slopes, with slope lengths generally ranging from 800 to 1500 m and gradients between three and seven degrees. This topography is prone to heavy slope runoff during the periods of intense rainfall, which mainly causes soil erosion in these areas, with sloping agriculture suffering the most severe erosion. Waterlogged zones form easily in low-lying areas on slopes due to the undulating topography and inadequate soil penetration. These areas not

only have an impact on crop growth but also act as the origins of erosion gullies. Significant soil and water loss occurs as fine gullies, shallow gullies, and rills arise as a result of the ongoing erosion in these places by water flow [1,2].

Waterlogged areas on slopes refer to the regions of a slope surface that are susceptible to water accumulation as a result of variations in the topography and challenges with soil infiltration. These waterlogged areas are small in area, scattered in distribution, and lack large-scale continuity, which is different from the traditional waterlogging areas. From a morphological perspective, these places can be categorized into two groups: those that already exhibit erosion gullies and those that do not. This distinction sets them apart from conventional slope erosion gullies. The boundaries of sloping cropland erosion areas are relatively clear and appear elongated, whereas erosion risk areas exhibit a planar spreading form in comparison. The present research on the identification of waterlogged areas predominantly centers on ecological investigations of vast waterlogged regions, with the objective of appraising and analyzing the effects of waterlogging events on crop development [3,4]. The presence of seasonal waterlogging disasters severely impacts regional food security [5,6]. The majority of research on erosion gullies focuses on gullies that have advanced in their development. In addition to conducting field surveys and analyzing topographical features, remote sensing visual interpretation and machine learning have emerged as the primary means of identification. However, challenges persist in terms of the efficiency and accuracy of identification [7–9]. In general, research in these fields primarily focuses on the watershed level, while less attention is given to identifying waterlogged areas on a smaller scale, such as sloping cropland.

The semantic segmentation method based on deep learning is currently used for feature recognition [10]. It can use the spectral information of the image elements for classification, in addition to conventional visual interpretation and inversion techniques. In recent years, models such as convolutional neural networks (CNN) [11], fully convolutional networks (FCN8s) [12], and densely connected convolutional networks (DenseNet) [13] have been widely applied in fields like medical imaging and environmental monitoring. The application of these technologies makes it possible to rapidly extract waterlogged areas. With the gradual improvement in the resolution of existing images, the classification accuracy of tiny features in a small area has further improved. However, the accurate identification of features, especially erosion gullies, is challenging due to issues such as ambiguity in boundary delineation [14], the complexity of features at high resolution, and significant seasonal variations [15]. The object-oriented method views a target object as a unified entity and utilizes its spectral, geometric, and textural features. Its segmentation process groups similar image elements together to define the boundaries of an object's features. This method has been widely used [16,17]. The construction of image objects is foundational to object-based image analysis (OBIA), a process that typically involves a variety of segmentation algorithms, such as those based on multi-scale, grayscale, texture, knowledge, and watershed segmentation methods [18]. A key advancement in OBIA is the integration of deep learning technologies, which includes incorporating deep features into the geographic object-based image analysis (GEOBIA) paradigm for remote sensing image classification, or combining deep learning techniques for semantic segmentation of high-resolution images. The integration of OBIA and deep learning can enhance the precision of semantic segmentation to a certain extent. This technique has been successfully applied in various applications, such as the extraction of vegetation [19], medical research [20], and the detection of buildings [21]. However, its application in identifying sloping cropland within a waterlogged area has not been reported. Remote sensing optical data and Synthetic Aperture Radar (SAR) data have been widely used to identify waterlogged areas. The former assesses the magnitude of waterlogging's influence on vegetation using measures such as the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI) to monitor waterlogged regions [22]. In contrast, the SAR data method employs an object-oriented analysis methodology to determine areas

by assessing the influence of the water content in waterlogged places on the reflectance of radar waves [23].

This study focused on the Qianjin subwatershed, which is a tributary of the Hailun River in Hailun city, located in Heilongjiang Province, China. This study utilized high-resolution remote sensing imagery as a data source to recognize regions of waterlogging in sloping croplands. To facilitate clear communication and comprehension, this study classified waterlogged areas on slopes into two different types: places that have already experienced the formation of erosion gullies, denoted as slope erosion zones, and areas that have not yet developed erosion gullies, referred to as erosion risk zones. We employed an object-oriented multi-resolution segmentation (MRS) algorithm to segment the study area images and generated a feature dataset comprising remote sensing images and segmentation results. By incorporating deep learning techniques and comparing model accuracies before and after integrating the OBIA, we aim to automate the extraction of waterlogged areas. This approach offers technical assistance for the swift and precise detection of waterlogged areas in sloping croplands.

2. Materials and Methods

2.1. Overview of the Study Area

The Qianjin subwatershed ($47^{\circ}20'–47^{\circ}24' \text{ N}$, $126^{\circ}48'–126^{\circ}52' \text{ E}$) is located in the central part of Hailun city, Heilongjiang Province, China (Figure 1), which has a mid-temperate continental monsoon climate with a frost-free period of approximately 120 days, an average annual temperature of 1.5°C , and an average annual precipitation of 530 mm. The precipitation exhibits a concentrated pattern during the period from May to September and is characterized by a brief duration, intense magnitudes, and uneven allocation. This watershed covers an area of approximately 23.52 km^2 , with elevations ranging from 192 to 258 m. The regional topography is characterized by rolling hills, with long and gentle slopes (mostly exceeding 200 m and below 5°). The primary soil type in the study area is typical black soil, the thickness ranges from 50 to 60 cm. This soil has a pH value of 5.8 and exhibits high-fertility characteristics, including an organic matter content of 54.71 g/kg, a total nitrogen content of 2.58 g/kg, and a total phosphorus content of 1.00 g/kg. The soil texture is mostly classified as clay loam, indicating that the soil is fine, which contributes to its high fertility and capacity to retain moisture and nutrients, making it suitable for the cultivation of crops such as soybeans and corn, which are predominant in this region.

2.2. Technical Flowchart

This study established enhanced target recognition models by integrating remote sensing data and object-oriented auxiliary data into the dataset used for the model. Models constructed using two different datasets—one based only on original remote sensing imagery and another combining remote sensing imagery with OBIA—were assessed for their accuracy in recognizing objects. The purpose of this study was to evaluate and compare the precision and efficiency of multiple models in identifying waterlogged areas of slopping cropland within a specific study area. Initially, a dataset was established using original remote sensing imagery, which was then further processed using the MRS method to obtain another data set marking as OBIA feature data set. Deep learning models were separately trained on these two types of data set, and the accuracy of each model was evaluated with the ultimate goal of achieving precise identification of waterlogged areas. Figure 2 illustrates the flowchart of this study.

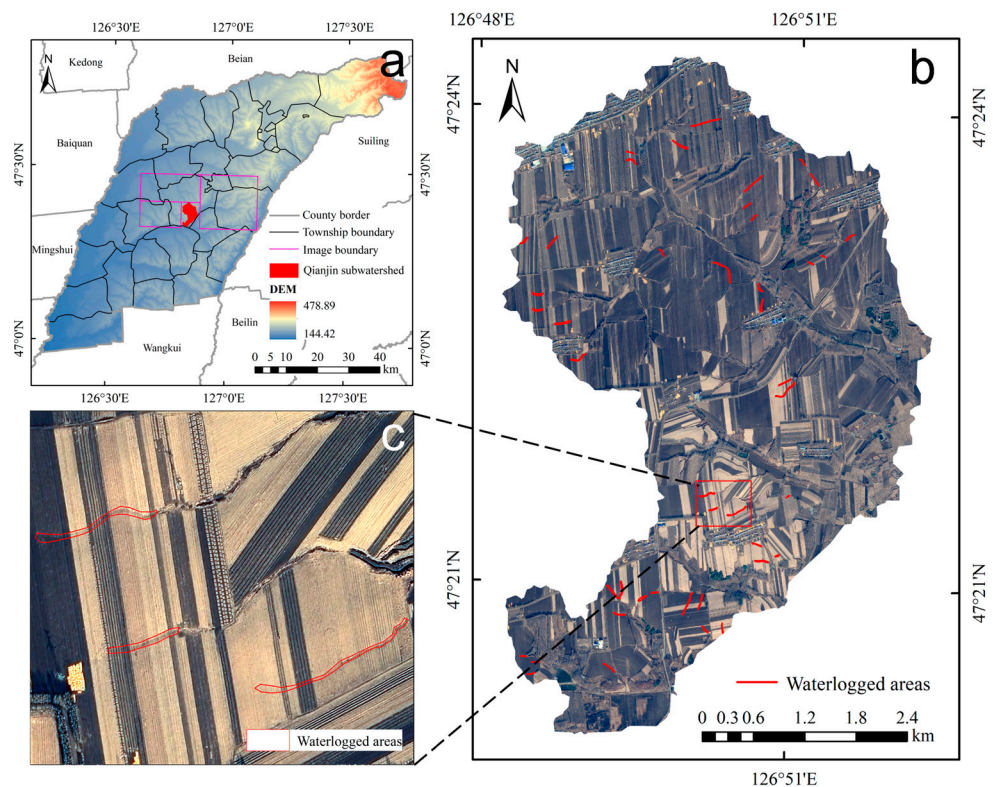


Figure 1. Overview of the study area: (a) Hailun city, Heilongjiang Province, China; (b) Jilin-1KF01B image of the Qianjin subwatershed; (c) boundary labeling of some waterlogged areas in the study area.

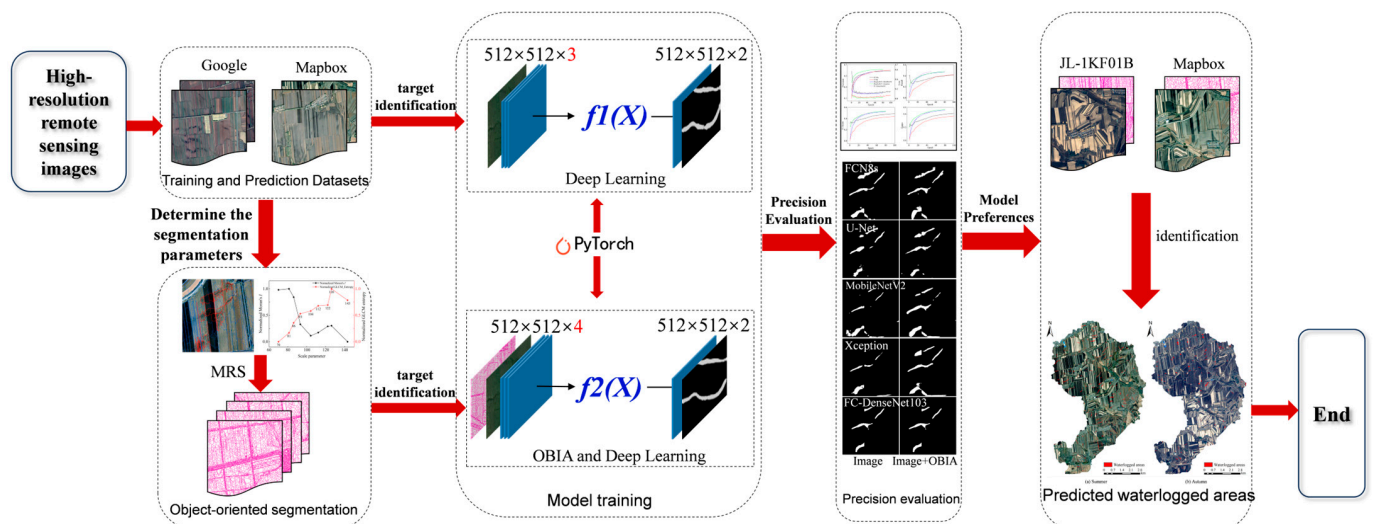


Figure 2. Flowchart of this study.

2.3. Remote Sensing Data Sources and Data Processing

The remote sensing image data used in this study were obtained from BIGEMAP software, version 25.5.0.1 (<http://www.bigemap.com/>, accessed on 23 March 2024) and the official website of the Jilin-1 satellite (<https://www.jl1mall.com/>, accessed on 23 March 2024). BIGEMAP integrates multiple mapping resources, including Google Maps, Tianditu, and Baidu Maps. The map data used for this study was chosen from a collection of remote sensing images captured at different locations and in different seasons with varied resolutions. The objective was to ensure that the trained model could be used in most remote sensing imaging application scenarios, thereby improving the model's universality. Finally,

the dataset comprised high-resolution images obtained from Google Maps, Mapbox, and JL-1KF01B, including a total area of over 1000 km². The detailed information of maps used in this study is described in Table 1.

Table 1. Data sources and information.

Data Sources	Time	Spatial Resolution (m)	Area (km ²)	Center Coordinate	Application Description
Google Maps	October 2021	0.20	105.51	126°42'36" E, 47°22'12" N	Training and validation data
Mapbox	July 2022	0.30	171.72	126°45'02" E, 47°27'05" N	
Google Maps	July 2022	0.30	316.26	127°00'00" E, 47°24'36" N	
Mapbox	October 2021	0.30	316.26	127°00'00" E, 47°24'36" N	
Jilin-1	October 2022	0.50	52.41	126°50'24" E, 47°22'12" N	Prediction data
Mapbox	July 2022	0.30	52.41	126°50'24" E, 47°22'12" N	

2.4. Object-Based Image Analysis

Object-based image analysis (OBIA) operates on the principle of using image objects, or segments, as the fundamental units of analysis. Image segmentation, the fundamental process of OBIA, employs certain algorithms and scales to divide an entire image region into collections of pixels with high feature similarity, creating “image objects” [24]. The object-based technique is more effective than traditional pixel-by-pixel analysis at reducing noise in the classification process and differences within classes and similarities between classes, further improving classification accuracy. MRS is specifically developed for high-resolution remote sensing imagery, utilizes the layered properties of images and caters to the extraction requirements of features at various scales. This ensures that the segmentation outcomes closely correspond to the real boundaries of a property.

The segmentation scale, shape, and compactness are the three factors that most significantly affect image segmentation effectiveness in MRS. Because different features have varied spatial characteristics, each feature type necessitates specific segmentation criteria. The segmentation scale determines the quantity and dimensions of the segmented entities. A smaller scale parameter leads to a greater quantity of entities with more fragmented boundaries, whereas a larger scale parameter has the opposite effect. The compactness parameter affects the degree of fragmentation in the segmentation process, where a lower compactness value results in more fragmented segmented objects. The shape parameter influences the degree of size variation observed among segmented objects. A greater shape parameter leads to a reduction in the size variability among segmentation results.

eCognition Developer software, version 10.3, which is extensively used for object-oriented image analysis, includes the Estimation Scale Parameter 2 (ESP 2) plugin [25]. This plugin was specifically developed to compute the local variance (LV) and its rate of change (ROC) at various segmentation scales. According to the rate of LV change, the best segmentation scale can be found at points where the ROC changes substantially. The extreme points on an ROC curve show where the best segmentation points are for each feature. In addition, we utilized the global Moran’s *I* to evaluate the variation across objects and used the gray level co-occurrence matrix entropy (GLCM_ENTROPY, GE) to quantify the consistency within objects. The optimal segmentation scale in this study was determined by identifying the intersection point between the scale values of both Moran’s *I* and GLCM entropy following normalization. Previous research demonstrated that the shape parameter for fine gully segmentation was between 0.3 and 0.5 and that the compactness parameter was between 0.5 and 0.7 [26]. Referring to these research findings and the visual interpretation of segmentation results, the optimal combination of shape and compactness factors was determined on the premise of optimal segmentation scale. Then, object-oriented image segmentation was implemented via MRS based on these determined segmentation parameters.

2.5. Dataset Production

Before establishing the training samples, this study employed visual interpretation techniques on remote sensing imagery to demarcate waterlogged areas. This process

entailed pinpointing the exact positions of the waterlogged areas in the images and transforming these regions into foreground data in raster format while marking other non-waterlogged features as background. As depicted in Figure 3, for summer imagery, the boundaries of erosion gullies in sloping cropland exhibit limited visual clarity and areas that have not yet formed into erosion gullies are also difficult to be accurately defined due to vegetation cover. In autumn, poor crop growth and soil erosion usually occur in waterlogged areas, which lead to the presenting distinct black areas. Considering the limited extent of the waterlogged areas in the study area, ArcGIS software, version 10.8, was used to label the waterlogged areas in the remote sensing data. In total, 638 waterlogged areas were identified via visual interpretation. Data labels were obtained by cropping these marked waterlogged areas.

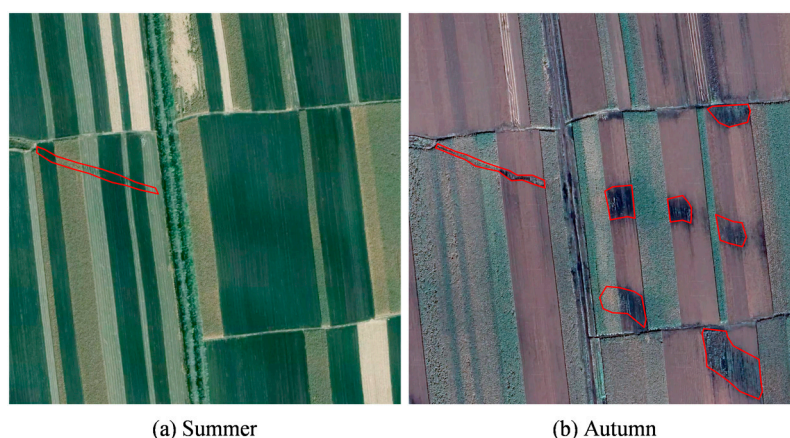


Figure 3. Seasonal changes in waterlogged areas (marked in red: waterlogged area boundary).

In this study, two types of training samples were constructed based on original remote sensing imagery and OBIA imagery for training deep learning models. A sliding window with an overlap ratio of 0.4 was employed for the remote sensing imagery. To utilize the computer and model effectively, the photos were chopped into pairs with 512×512 pixels to create training examples. By removing labels with very small area proportions, a total of 4000 samples were acquired. All samples contained both foreground and background parts, and they were divided at a ratio of 3:1 for the training and validation of the deep learning model. The un-cropped remote sensing images were processed using the MRS method to generate MRS vector data, which were then converted into raster binary data corresponding to the pixels, with boundary pixels marked as 1 and other pixels marked as 0. This ensured that the MRS data corresponded one to one with the pixels of the remote sensing images, and sliders were used for cropping at the same locations as the remote sensing images to obtain a data set with a matching boundary. The OBIA and deep learning dataset was created by adding a single-band binary image from the MRS dataset to the original remote sensing image's three RGB bands. Since the cropping locations of both images were exactly the same, they could be overlaid to form four-channel image data. The two types of training samples were based on the same image foundation, with identical cropping methods and locations, ensuring that both sets had the same number of samples and the same precision.

2.6. Building Deep Learning Model

In this study, five widely used semantic segmentation models were selected as the main frameworks for deep learning, namely SE-FCN8s, U-Net, DeepLabV3+-Xception, DeepLabV3+-MobileNetV2, and FC-DenseNet103. SE-FCN8s is developed from FCN8s by integrating the channel attention mechanism squeeze and excitation network (SENet) into each of the four upsampling layers of the FCN8s. U-Net, a different convolutional neural network, employs an encoding and decoding approach to maximize the utilization of information within the network. DeepLabV3+ has a unique atrous convolution module

that, when combined with the atrous spatial pyramid pooling (ASPP) module, makes it possible to extract features at different scales [27]. We used the DeepLabV3+ model with two different backbones: MobileNetV2 and Xception. The MobileNetV2 backbone is less complex than the Xception backbone and requires only half the training time [28]. FC-DenseNet enhances the structure of FCNs by augmenting the model depth, thereby enhancing accuracy.

We employed weighted cross-entropy as a loss function to address the problem of imbalanced positive and negative samples in image classification. This method amplifies the importance of positive samples in the loss function, effectively improving the recognition accuracy of the model. The Adam optimizer was chosen for binary classification tasks. Due to the intricate nature of the model architecture and the constraints on the number of training samples, data augmentation involving flipping images horizontally and vertically and rotating them randomly was implemented to enhance the model's ability to generalize. During the process of feeding data into the model, data augmentation was applied to all samples inputted in each epoch.

The models were trained in a Python 3.9 environment using the PyTorch 1.12 framework and took advantage of CUDA 11.3 for GPU (A5000, 24 GB) acceleration to improve computational performance. The learning rate was uniformly set to 2×10^{-4} . The batch size, which determined the number of samples input into the model at one time, was adjusted to the maximum capacity of the GPU memory for all models. During the deep learning process, the training epochs were extended until the model accuracy reached a stable state, ensuring consistent model training. At the same time, the parameter configurations were maintained consistently before and after the introduction of OBIA to mitigate potential errors arising from parameter discrepancies. Table 2 shows the detailed information and parameter settings of models.

Table 2. Model parameters and settings.

Approach	Models	Backbones	Epochs	Time (min/epoch)
Deep Learning	SE-FCN8s	VGG19	100	4.4
	U-Net	—	100	4.1
	DeepLabV3+	MobileNetV2	80	3.9
	DeepLabV3+	Xception	80	6.4
	FC-DenseNet103	—	50	41.7
OBIA and Deep Learning	SE-FCN8s	VGG19	100	5.1
	U-Net	—	100	4.2
	DeepLabV3+	MobileNetV2	80	4.7
	DeepLabV3+	Xception	80	7.5
	FC-DenseNet103	—	50	42.7

2.7. Accuracy Evaluation

To assess the reliability of the model in identifying waterlogged areas on sloping cropland, we employed precision assessment measures to analyze the outcomes. The evaluation indicators include precision, recall, the *F1-score* and the kappa coefficient. Precision represents the proportion of correctly identified positive samples to all samples identified as positive; recall reflects the proportion of correctly predicted positive samples to the actual ground truth positive samples; the *F1-score* is the harmonic mean of precision and recall, providing a measure of their combined performance and the kappa coefficient assesses the consistency of model predictions with actual classifications, calculated based on a confusion matrix. Area accuracy refers to the ratio of the area correctly predicted by the model to the total area, used to indicate the model's accuracy in predicting areas. The values of precision, recall, area accuracy, and the *F1-score* range from 0 to 1, with values closer to 1 indicating better model performance. The kappa coefficient ranges from -1 to 1 , where -1 indicates total inconsistency and 1 indicates complete consistency, making it a key indicator of model prediction accuracy. The formulas for these metrics are as follows:

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$F1\text{-score} = \frac{2Precision \cdot Recall}{Precision + Recall} \quad (3)$$

$$Kappa = \frac{P_o - P_e}{1 - P_e} \quad (4)$$

$$Area\ accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (5)$$

$$P_o = \frac{\sum_{i=1}^C T_i}{n} \quad (6)$$

$$P_e = \frac{\sum_{i=1}^C a_i * b_i}{n^2} \quad (7)$$

In these formulas, TP refers to the number of waterlogged area grids correctly identified by the model, FP represents the number of grids erroneously predicted to be waterlogged areas, FN represents the number of waterlogged area grids that the model failed to identify, and TN refers to the number of correctly identified non-waterlogged area grids. P_o is calculated as the sum of correctly classified samples in each class divided by the total number of samples. C represents the total number of categories, which was set to two in this study. T_i represents the number of accurately categorized samples in each category. The actual number of samples in each class is referred to as a_1 and a_2 , while the predicted number of samples in each class is denoted as b_1 and b_2 .

3. Results

3.1. Optimal Segmentation Parameters for Remote Sensing Images

We used the ESP 2 tool to investigate the LV and ROC at different segmentation scales within the imagery. Nine segmentation scale values were calculated by selecting the extremum points of the variance change rates to analyze the homogeneity and heterogeneity of the segmented objects, as shown in green in Figure 4a. During the initial selection step, the shape and compactness parameters were set to 0.5 and 0.7, respectively. This design made it easier to collect MRS results from the nine segmentation scales. The normalized Moran's I and GE were subsequently calculated (Figure 4b).

Figure 4 shows that the normalized Moran's I typically decreased but the normalized GE increased, with the segmentation scale increasing. These two variables intersected at a segmentation scale of 90.25, where the balance between how similar things were within each segmented object and how dissimilar they were from one another struck a sweet spot. However, this point was useful for the overall view; it was not always ideal for identifying waterlogged areas. According to this intersected point, we selected four scales around 90.25, which were 88, 89, 90, 91, and 92 to compare the segmentation results at different scales with the delineation of waterlogged area boundaries. We determined 89 as the segmentation scale for identifying waterlogged areas in our study zone.

Under a segmentation scale of 89, we set different shape and compactness factors. Segmentation results with different parameters can be seen in Figure 5. Decreasing the values of the shape and compactness parameters contributed to a fragmented outcome. When the shape parameter was set to 0.3, the segmentation result showed an excessive number of pieces, surpassing our requirements (Figure 5, #1). Conversely, excessively increasing the shape parameter and enhancing compactness leads to under-segmentation, resulting in loss of details at a shape parameter value of 0.5 (Figure 5, #2). After testing several combinations, a shape parameter of 0.4 and a compactness parameter of 0.6 yielded optimal results in segmentation quality and accuracy.

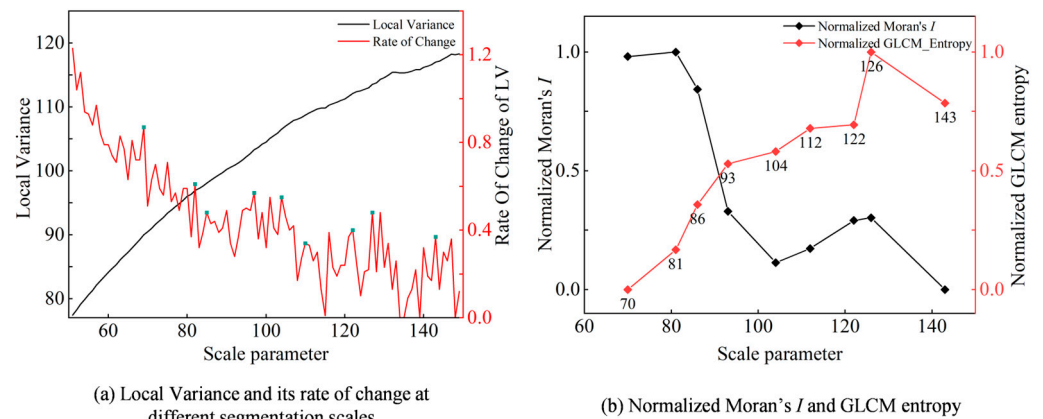


Figure 4. Determination of the optimal segmentation scale.

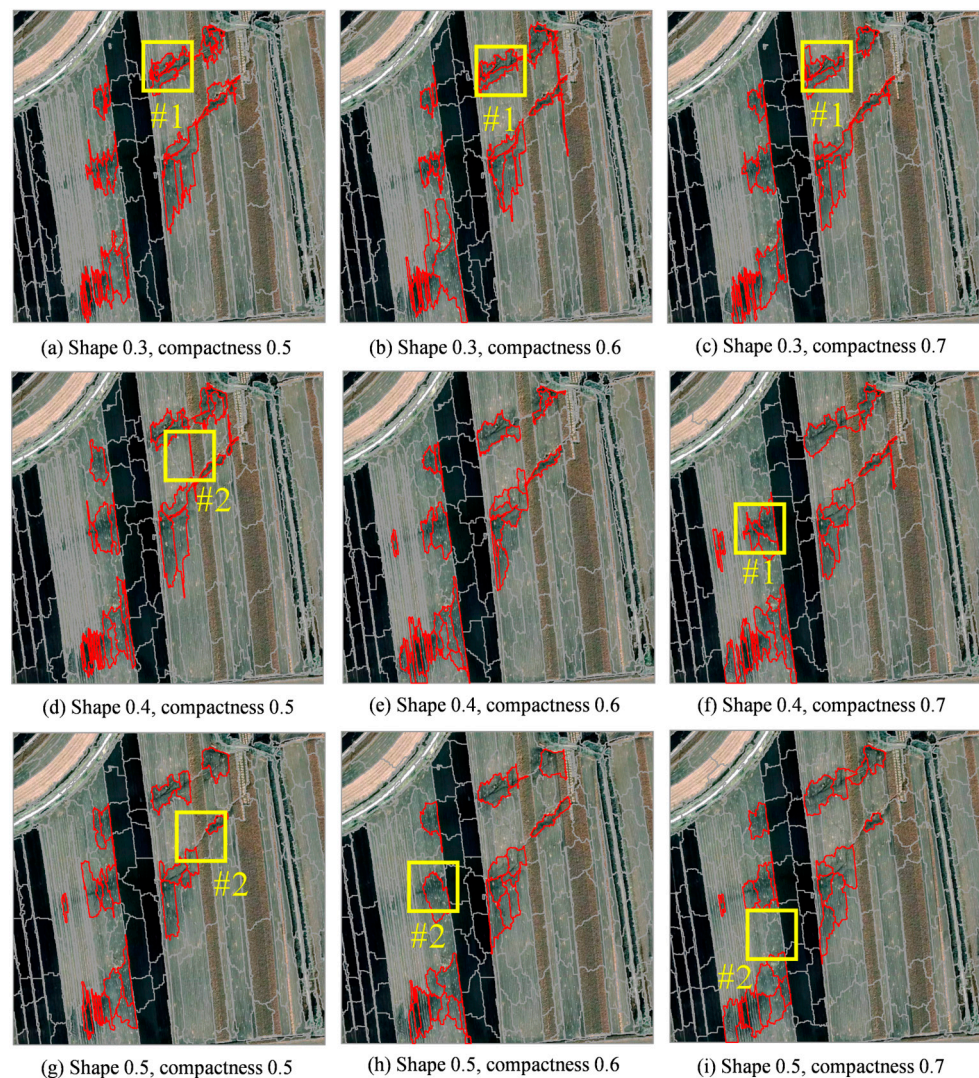


Figure 5. Segmentation results of localized images with different parameters (figures contain the following labels: #1, more redundancy in segmentation; #2, lack of detail).

3.2. Conventional Deep Learning Algorithms for Waterlogged Area Recognition

The results of conventional deep learning-based identification show significant differences in validation accuracy among the various models for the task of recognizing

waterlogged areas, as detailed in Table 3 and illustrated in Figure 6. In precision assessment, the Xception model led the pack with a precision rate of 60.9%, while SE-FCN8s, U-Net, and MobileNetV2 hovered around 59%. FC-DenseNet103 had inferior performance compared to the other models. On the other hand, analysis of the recall rates demonstrated that SE-FCN8s and U-Net had much poorer performances than the other three models, which might be attributed to the simpler architectures of these two models and their constrained ability to perform comprehensive picture recognition. The results for area accuracy reveal that all models achieved a prediction accuracy exceeding 95%, indicating strong classification performance. Notably, the Xception model demonstrated the highest level of precision, distinguishing itself as the most accurate among the tested models. Figure 6 emphasizes the drawbacks that SE-FCN8s and U-Net encounter when learning and fitting models. When evaluating the overall performance of the models, the Xception model had superior classification accuracy compared to the other models, as evidenced by its high *F1-score* and kappa coefficient. In contrast, the performance of the U-Net model was the poorest.

Table 3. Recognition results and accuracy statistics of different models.

Models	Epochs	Precision (%)	Recall (%)	<i>F1-score</i> (%)	Kappa	Area Accuracy (%)
SE-FCN8s	100	59.1	41.5	48.8	0.41	95.5
U-Net	100	59.1	34.8	43.8	0.37	95.4
MobileNetV2	80	59.4	47.2	52.6	0.46	96.9
Xception	80	60.9	47.5	53.4	0.47	97.0
FC-DenseNet103	50	54.6	48.0	51.1	0.43	96.3

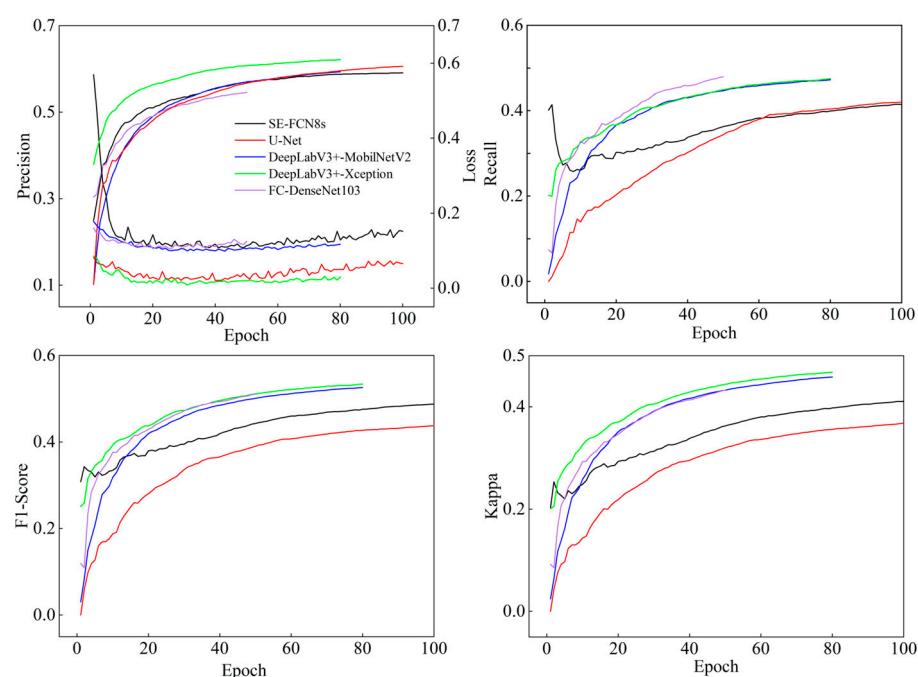


Figure 6. Training accuracy curves for different models.

When assessing the trained models, we found significant differences in the efficacy of recognizing two distinct types of waterlogged areas: erosion areas in sloping areas (Figure 7) and erosion risk areas (Figure 8). Generally, the models were able to identify the approximate locations of waterlogged areas. However, the accuracy of their border definitions required improvement. Specifically, the models performed significantly better when recognizing sloping cropland erosion zones than when recognizing erosion risk zones. The U-Net model had the smallest range of recognition in identifying cropland

erosion zones (Figure 7). All the models struggled to maintain the continuity of the sloping cropland erosion zones; especially SE-FCN8s, U-Net, and FC-DenseNet103 had significant issues, as shown in Figure 7 (#1). The MobileNetV2 model, however, displayed more severe fragmentation with a strong influence of noise (Figure 7d). In terms of accuracy, SE-FCN8s had the highest level of precision in identifying waterlogged areas (Figure 7b), while the other models occasionally misidentified passageways between fields as waterlogged areas (Figure 7, #2). However, SE-FCN8s had significant issues with incorrect classification and information omission, which resulted in subpar recognition outcomes in identifying erosion risk zones (Figure 8, #1 and #3). Additionally, some models incorrectly recognized roadways as waterlogged areas (Figure 8, #2). MobileNetV2 exhibited superior recognition efficacy compared to the other models, with fewer errors and more accurate border identification (Figure 8d). This indicated its superior adaptability in identifying waterlogged areas. These findings emphasize that the performance of deep learning models in recognizing waterlogged areas varies depending on the model and feature. This implies that further enhancements are necessary for enhancing recognition accuracy and boundary precision.

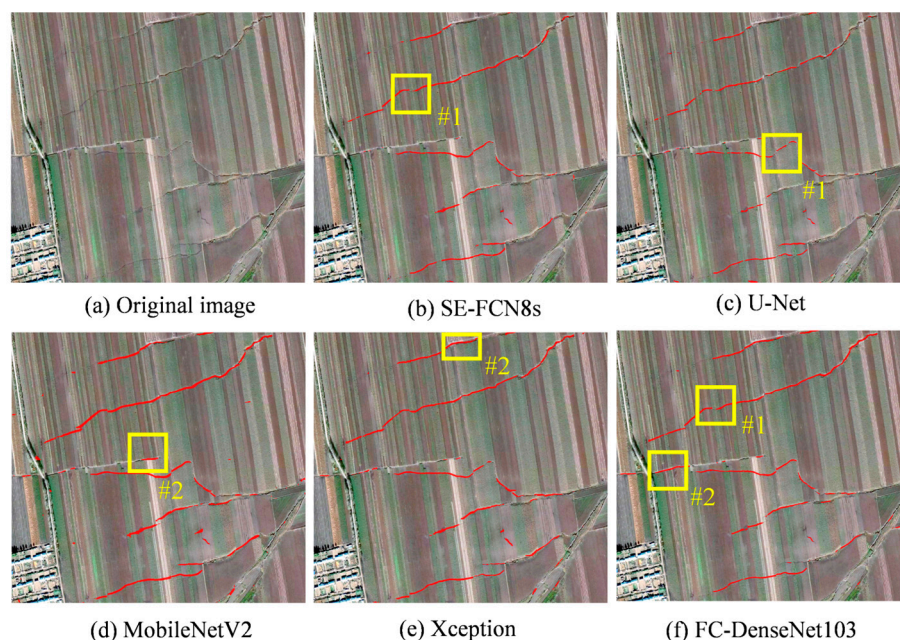


Figure 7. Results of the recognition of erosion areas on sloping cropland (figures contain the following labels: #1, missed score issue; #2, road identification issue).

3.3. The Integration of OBIA and Deep Learning for Waterlogged Area Recognition

To improve the accuracy of the models in terms of recognizing objects, we integrated the results of object-oriented image segmentation into the deep learning models. As depicted in Figure 9, when incorporating object-oriented data, all models exhibited substantial enhancements in both *F1-score* and kappa coefficient compared to the model trained on image information alone. The improvement in area accuracy is relatively limited. Notably, the SE-FCN8s and U-Net models displayed particularly significant gains, with SE-FCN8s achieving a remarkable increase in accuracy of approximately 10%. By implementing this novel approach, MobileNetV2 demonstrated superior recognition accuracy, followed closely by SE-FCN8s. Both achieved *F1-scores* approaching 60%.

The inclusion of the object-oriented approach not only improved the overall accuracy of the models but also effectively resolved the problem of fragmentation caused by characteristics that were challenging to identify or pixel-based semantic segmentation noise. The SE-FCN8s model, upon including object-oriented data, effectively reduced the noise in the images, while the U-Net model introduced more noise during this procedure (Figure 10, #1). Additionally, omission errors were reduced in all the models (#2). The MobileNetV2 and

Xception models had similar tendencies to erroneously classify road edges as waterlogged areas (Figure 10, #3), while the FC-DenseNet103 model demonstrated a more conservative recognition threshold, leading to a greater number of omission mistakes. In terms of the recognition results, the MobileNetV2 model demonstrated superior performance. These results suggest that integrating object-oriented image segmentation outcomes might significantly enhance the precision and robustness of deep learning models in identifying waterlogged areas.

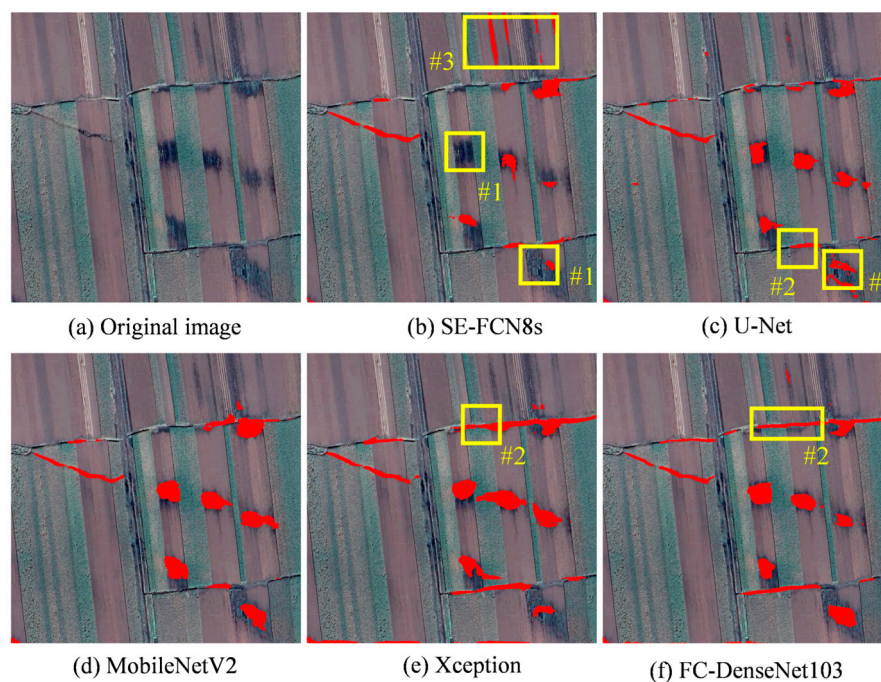


Figure 8. Results of recognizing erosion risk areas (figures contain the following labels: #1, missing score problem; #2, road identification problem; and #3, wrong score problem).

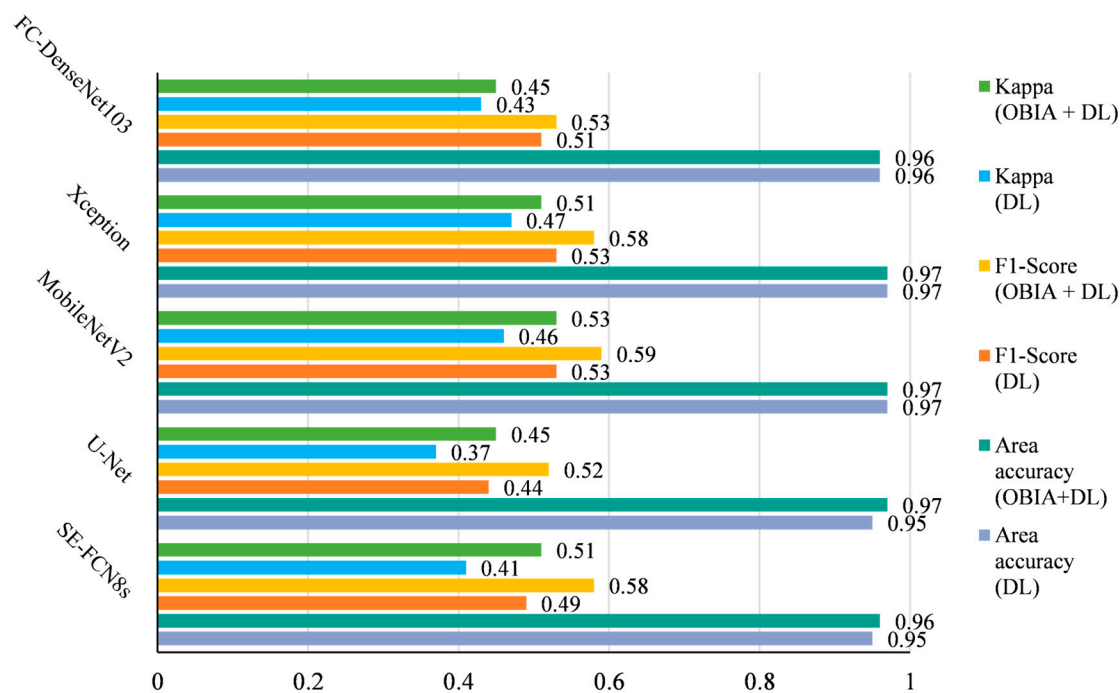


Figure 9. Comparison of the accuracy of different models after the integration of OBIA.

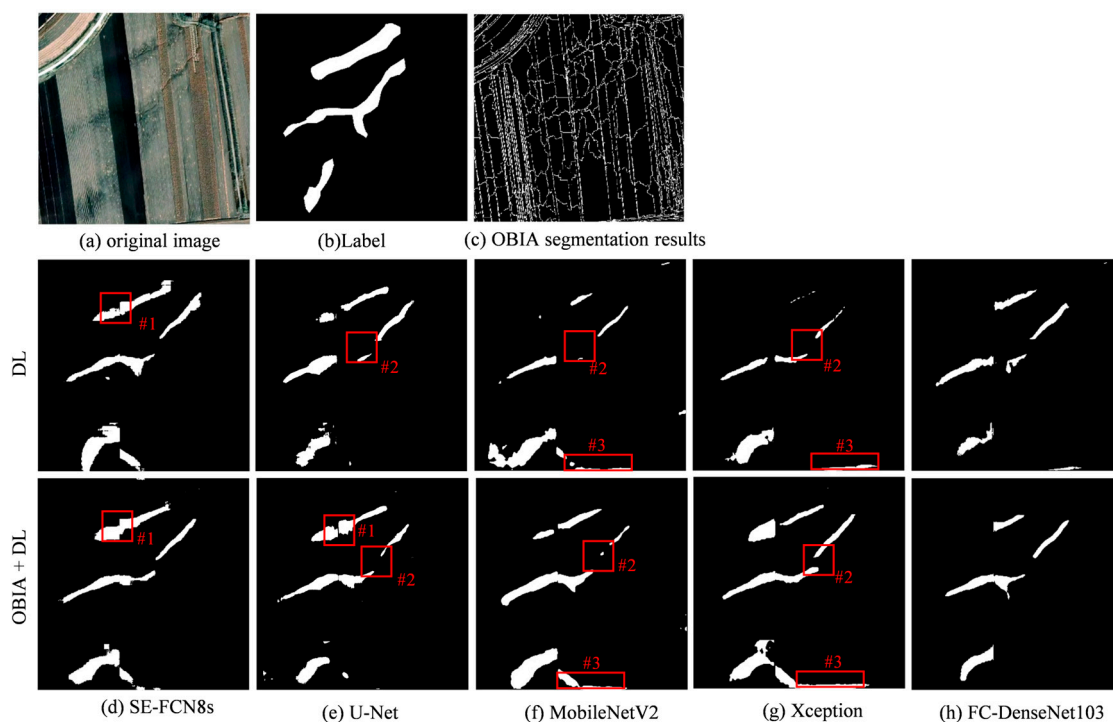


Figure 10. Model recognition results (figures contain the following labels: #1, noise performance; #2, continuity issues; and #3, road misidentification issues).

When analyzing the impact of integrating OBIA data on model accuracy, we observed variations in accuracy when MobileNetV2 was used, as shown in Figure 11. During the initial training phase, which covered the first 10 epochs, both the model utilizing the OBIA and deep learning imagery and the model relying purely on the original imagery had comparable rates of convergence. However, the former had higher initial accuracy. This phenomenon may be attributed to the heightened intricacy of the OBIA and the deep learning model, which led to diminished accuracy during the initial stages of the learning process. However, after 20 epochs, the rate at which the model relying only on the original images was fitting started to fall, whereas the rate at which the model including the object-oriented strategy was learning decreased more slowly. Ultimately, the proposed algorithm outperformed the original image-based algorithm, attaining superior and consistent accuracy in the final outcome. This illustrates that while the incorporation of OBIA data may initially hinder the development of accuracy due to the heightened complexity of the model, in the long run, this technique can significantly enhance the efficiency of the model's learning process and ultimately improve its accuracy. This approach greatly enhances the efficiency of deep learning models for specific tasks, demonstrating that an enhanced deep learning model is a good approach for enhancing outputs.

Finally, we employed an enhanced MobileNetV2 model to identify the waterlogged area in the Qianjin subwatershed. The outcomes are displayed in Figure 12. Upon examining the identification of waterlogged areas in summer and autumn, we observed that the waterlogged areas in summer primarily consisted of sloping cropland erosion zones; however, there was a mixture of both sloping cropland erosion zones and areas at risk of erosion in autumn. The summer recognition results surpassed those in autumn in terms of both range and precision. The primary factors contributing to the decreased performance during autumn were the heightened presence of bare soil and residual straw, which diminished the color contrasts in the images and diminished the overall differentiation between terrain characteristics. This finding highlighted the significance of selecting imaging data that correspond to seasonal patterns to increase the efficiency of identifying waterlogged areas. The identifying of waterlogged areas is helpful to make runoff regu-

lation strategy. The goal is to prevent the future growth of waterlogged areas and their accompanying hazards.

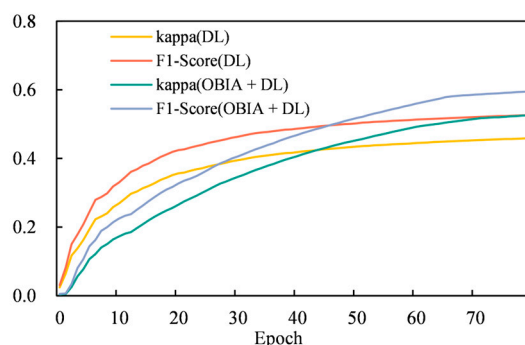


Figure 11. Accuracy curves of MobileNetV2 under different treatments.

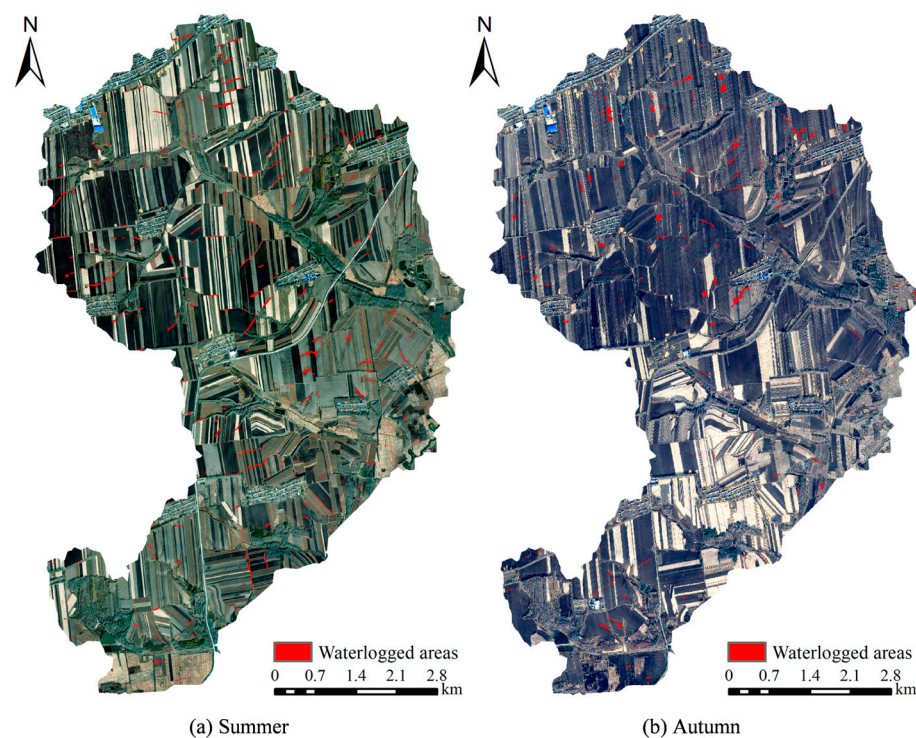


Figure 12. Recognition results for the study area.

4. Discussion

Deep learning enhances the ability to identify waterlogged areas by acquiring advanced identification skills through the extraction of deep characteristics from images. However, this approach still relies on pixel-based categorization determinations, where classification is based on probabilities of different categories, potentially leading to the inclusion of noise and indistinct borders. Combining deep learning with OBIA, which focuses on identifying features based on objects rather than pixels, might enhance model accuracy [21,29,30]. Based on such assumptions, a series of studies was conducted. The efficacy of this combination exhibits substantial variation contingent upon the object being identified and the model employed [31,32]. The primary objective of this work was to investigate changes in accuracy and their influence on outcomes when diverse deep learning models incorporated object-oriented methodologies. CNN and U-Net models were often combined with OBIA techniques in present studies because their architectures are not excessively complicated. Our result also showed that these two models provided

significant accuracy enhancements when utilized in conjunction with object-oriented data (Figure 9), and their lower computational requirements made them easy to apply. However, when OBIA was combined with other advanced models such as FC-DenseNet103 and DeepLabV3+-Xception, they exhibited less significant improvement, even though they were better at extracting high-level characteristics. In this study, we observed that DeepLabV3+-Xception performed best, while DeepLabV3+-MobileNetV2 showed the best performance after applying OBIA images (Figure 9). Both Xception and MobileNetV2 utilize depthwise separable convolution technology, but Xception aims to increase model parameters [33], whereas MobileNetV2 aims to compress the model and increase speed. It therefore has fewer model parameters. Many studies indicate that, given sufficient computational resources, Xception's accuracy surpasses that of MobileNetV2 as Xception can provide more boundary information, which is consistent with our results of using original images [34]. Though after applying OBIA images, MobileNetV2 showed a more significant improvement, signifying that the added boundary information helps MobileNetV2 capture some complex textures which are difficult for it to recognize on its own. However, the results were very similar after 100 iterations for the two models (Figure 9), indicating further exploration needed to determine whether there is still room for improvement in Xception. Above all, to obtain the highest possible accuracy in recognition by deep learning, it is important to consider the model's complexity and how well it works with OBIA techniques.

When dealing with a high-resolution image, an increase in resolution has both advantages and disadvantages. On the one hand, it reduces mixed pixels and facilitates the discernment of intricate details. On the other hand, it accentuates the similarity between different features and noise. It was particularly evident in the task of identifying waterlogged areas with lack distinct boundaries. Using OBIA techniques helped with some problems that come with pixel-based semantic identification, such as blurred edges and the presence of noise. OBIA and deep learning approach led to better consistency for related features, as shown in Figure 10. After implementing object-oriented approaches, the DeepLabV3+-MobileNetV2 model achieved a significant increase in accuracy, with a maximum improvement of 6%, even though its *F1-score* only reached 59%. A comparison with research conducted by Guirado et al. demonstrated that the utilization of object-oriented analysis with Mask R-CNN resulted in 81% increase in the accuracy of vegetation cover identification [29]. This finding implied a potential inverse correlation between the complexity of the model and the improvement degree in accuracy. Our results emphasized the efficacy and promise of OBIA techniques for improving the ability of deep learning models to handle intricate feature recognition tasks, especially in situations involving high-resolution imagery.

Research of waterlogged areas focuses on the erosion gully in sloping cropland, but research on erosion risk areas which have not developed erosion gullies is quite limited. The primary reason is that regions of erosion on sloping cropland are more prevalent and have more distinct boundaries, rendering them easier to identify. Nevertheless, erosion risk areas pose greater potential risks and are easier to manage in their early stages. We conducted a statistical comparison of the contributions of our study and those of existing technologies, and the results are presented in Table 4. While deep learning models have been refined in numerous studies, resulting in improved identification accuracy ranging from 70% to 84%, it is worth noting that these studies typically utilize low-resolution images [35,36]. Our focus was on investigating the impact of integrating OBIA with various deep learning models and high-resolution images. There is great potential for enhancing the performance of models through better tuning and processing of data. Subsequent studies will prioritize the exploration of improved methodologies for the optimal utilization of high-resolution remote sensing imagery in order to accurately identify erosion sites and determine their associated danger zones.

Table 4. Results from different studies.

Research	Method	Accuracy	Application Scenarios	Image Resolution
This study	OBIA + multiple deep learning models	<i>F1-score</i> : 53–60%	Waterlogged area identification	0.3–0.5 m
Wei et al. [21]	OBIA + FCN8s	<i>F1-score</i> : 83%	Land use identification	0.5 m
Zhao et al. [37]	OBIA + CNN	Precision: 70%	Land use identification	0.5m
Ghorbanzadeh et al. [38]	OBIA + ResU-Net	<i>F1-score</i> : 76%	Landslide detection	10 m
Guan et al. [23]	OBIA + Gaussian Mixture Model (GMM)	<i>F1-score</i> : 84%	Flooded crops identification	10 m

Recent research on enhancing recognition accuracy has shown two primary pathways for improving accuracy by combining OBIA methodologies and deep learning. The first approach is conducting OBIA to select and combine characteristics prior to inputting these data into a deep learning model for training [38,39], which was used in our study. The second approach implied employment of deep learning for initial image analysis, followed by merging the results with OBIA to enhance the ultimate recognition results [19,21]. In order to investigate more effective utilization of object-oriented approaches, we combined the two methods described above. Specifically, we integrated the MobileNetV2 model with OBIA image segmentation twice. The identification results of waterlogged areas were initially obtained using the method developed in this study. Subsequently, the boundary information acquired from OBIA was employed to determine the proportion of identified waterlogged areas within the boundaries defined by OBIA to optimize the recognition results. As illustrated in Figure 13, this approach further mitigated problems associated with indistinct boundaries. Notably, integrating the selected segmented objects with the model's recognized objects greatly improved the precision of the recognition of borders. This is demonstrated in the image where the model accuracy was improved when the model's outcomes encompassed 50% and 25% of the segmented objects. This method improved the boundary information by integrating OBIA data before and after the deep learning framework. Yin and his colleagues enhanced the accuracy of boundary recognition by incorporating a Spatial Transformation Network (STN) into their approach [40]. This modification resulted in favorable outcomes and demonstrated potential for future investigation.

**Figure 13.** OBIA and MobileNetV2 recognition results and segmentation data fusion (the 50% share area is displayed in the figure's box).

Our study mostly used the shape cues of objects acquired from images to recognize waterlogged areas, without considering other factors that could affect the accuracy of detection. Prior research has shown that the digital elevation model (DEM), an essential part of topography, was crucial for identifying features beyond what can be seen in images alone [41,42]. Researchers such as Chen Rong have successfully enhanced the accuracy of models by approximately 2–5% by integrating topographic elements into their data sources, as opposed to models that exclusively depend on photography [43]. Nevertheless, our study did not consider the impact of terrain or other factors on the creation of waterlogged areas, as we lacked DEM data with a matching resolution and relied solely on high-spatial-resolution remote sensing photos. In light of this, it is imperative for future studies to explore in greater detail the influence of many factors, such as topography and vegetation, on the precision of models used to identify waterlogged areas [44,45]. It will not only increase the overall recognition accuracy of the models, but will also augment their understanding and identification capacities for intricate terrain characteristics.

5. Conclusions

This study integrated OBIA and deep learning approaches to efficiently identify waterlogged areas using high-resolution remote sensing images. By conducting a comparative investigation of various semantic segmentation models, including SE-FCN8s, U-Net, MobileNetV2, Xception, and FC-DenseNet103, we arrived at the following main conclusions.

The optimal parameters for segmenting waterlogged areas in object-oriented MRS were a segmentation scale of 89, a shape parameter of 0.4, and a compactness parameter of 0.6. Among all the deep learning models, the Xception model exhibited exceptional classification accuracy, with a precision of 60.9%, an area accuracy of 97% and an *F1-score* of 53.4%. This highlighted its remarkable ability to recognize waterlogged areas. More precisely, the SE-FCN8s model exhibited superior performance in detecting areas with sloping cropland erosion, whereas the MobileNetV2 model exhibited excellent performance in identifying regions with a high risk of erosion.

Among the models that incorporated OBIA images, the MobileNetV2 model demonstrated the highest *F1-score*, achieving a rate of 59%. This finding suggested that this model achieved optimal performance when utilized in conjunction with object-oriented approaches. The implementation of OBIA technology led to an increase in accuracy of approximately 6%. This improvement enhanced the model's ability to handle ambiguity at edges, reduced image noise, and improved the continuity between terrain objects. However, this approach did not have a significant effect on misclassification issues.

By employing high-resolution images and integrating OBIA and deep learning techniques, this study successfully detected waterlogged areas in sloping cropland. The recognition accuracy for sloping cropland erosion areas was notably superior to that for erosion risk areas. The recognition performance in the summer was better than that in the autumn, yet challenges such as indistinct borders and incorrect categorization persisted.

Author Contributions: Conceptualization, P.X., M.W., Z.T. and X.S.; Data curation, P.X.; Formal analysis, P.X.; Investigation, S.W.; Methodology, R.M.; Project administration, M.W. and Z.T.; Resources, M.W.; Software, P.X.; Supervision, S.W., M.W., Y.L. and X.S.; Validation, Y.L. and X.S.; Visualization, P.X.; Writing—original draft, P.X.; Writing—review and editing, M.W. and R.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Strategy Priority Research Program (Category A) of the Chinese Academy of Sciences (grant number XDA28010202).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: We thank the journal’s editors and reviewers for their valuable suggestions to improve this paper.

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

References

1. Zhang, S.; Han, X.; Cruse, R.M.; Zhang, X.; Hu, W.; Yan, Y.; Guo, M. Morphological Characteristics and Influencing Factors of Permanent Gully and Its Contribution to Regional Soil Loss Based on a Field Investigation of 393 Km² in Mollisols Region of Northeast China. *CATENA* **2022**, *217*, 106467. [\[CrossRef\]](#)
2. Zhou, P.; Guo, M.; Zhang, X.; Zhang, S.; Qi, J.; Chen, Z.; Wang, L.; Xu, J. Quantifying the Effect of Freeze–Thaw on the Soil Erodibility of Gully Heads of Typical Gullies in the Mollisols Region of Northeast China. *CATENA* **2023**, *228*, 107180. [\[CrossRef\]](#)
3. Maniruzzaman, M.; Sarangi, S.K.; Mainuddin, M.; Biswas, J.C.; Bell, R.W.; Hossain, M.B.; Paul, P.L.C.; Kabir, M.J.; Digar, S.; Mandal, S.; et al. A Novel System for Boosting Land Productivity and Income of Smallholder Farmers by Intercropping Vegetables in Waterlogged Paddy Fields in the Coastal Zone of the Ganges Delta. *Land Use Policy* **2024**, *139*, 107066. [\[CrossRef\]](#)
4. Deng, C.; Zhang, Y.; Bailey, R.T. Evaluating Crop–Soil–Water Dynamics in Waterlogged Areas Using a Coupled Groundwater–Agronomic Model. *Environ. Model. Softw.* **2021**, *143*, 105130. [\[CrossRef\]](#)
5. Kaur, G.; Singh, G.; Motavalli, P.P.; Nelson, K.A.; Orlowski, J.M.; Golden, B.R. Impacts and Management Strategies for Crop Production in Waterlogged or Flooded Soils: A Review. *Agron. J.* **2020**, *112*, 1475–1501. [\[CrossRef\]](#)
6. Qamer, F.M.; Abbas, S.; Ahmad, B.; Hussain, A.; Salman, A.; Muhammad, S.; Nawaz, M.; Shrestha, S.; Iqbal, B.; Thapa, S. A Framework for Multi-Sensor Satellite Data to Evaluate Crop Production Losses: The Case Study of 2022 Pakistan Floods. *Sci. Rep.* **2023**, *13*, 4240. [\[CrossRef\]](#)
7. Lesschen, J.P.; Kok, K.; Verburg, P.H.; Cammeraat, L.H. Identification of Vulnerable Areas for Gully Erosion under Different Scenarios of Land Abandonment in Southeast Spain. *CATENA* **2007**, *71*, 110–121. [\[CrossRef\]](#)
8. Guan, Y.; Yang, S.; Zhao, C.; Lou, H.; Chen, K.; Zhang, C.; Wu, B. Monitoring Long-Term Gully Erosion and Topographic Thresholds in the Marginal Zone of the Chinese Loess Plateau. *Soil Tillage Res.* **2021**, *205*, 104800. [\[CrossRef\]](#)
9. Gao, Y.; Zhu, Y.; Chen, J.; Yang, X.; Huang, Y.; Song, F.; He, Y.; Tian, Z.; Lin, L.; Cai, C.; et al. Temporal and Spatial Distribution and Development of Permanent Gully in Cropland in the Rolling Hill Region (Phaeozems Area) of Northeast China. *CATENA* **2024**, *235*, 107625. [\[CrossRef\]](#)
10. Chen, R.; Zhou, Y.; Wang, Z.; Li, Y.; Li, F.; Yang, F. Towards Accurate Mapping of Loess Waterworn Gully by Integrating Google Earth Imagery and DEM Using Deep Learning. *Int. Soil Water Conserv. Res.* **2024**, *12*, 13–28. [\[CrossRef\]](#)
11. Yang, T.; Jiang, S.; Hong, Z.; Zhang, Y.; Han, Y.; Zhou, R.; Wang, J.; Yang, S.; Tong, X.; Kuc, T. Sea-Land Segmentation Using Deep Learning Techniques for Landsat-8 OLI Imagery. *Mar. Geod.* **2020**, *43*, 105–133. [\[CrossRef\]](#)
12. Lu, Y.; Chen, Y.; Zhao, D.; Chen, J. Graph-FCN for Image Semantic Segmentation. In *Advances in Neural Networks—ISNN 2019, Proceedings of the 16th International Symposium on Neural Networks, ISNN 2019, Moscow, Russia, 10–12 July 2019*; Lu, H., Tang, H., Wang, Z., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 97–105.
13. Alalwan, N.; Abozeid, A.; ElHabshy, A.A.; Alzahrani, A. Efficient 3D Deep Learning Model for Medical Image Semantic Segmentation. *Alex. Eng. J.* **2021**, *60*, 1231–1239. [\[CrossRef\]](#)
14. Liu, B.; Zhang, B.; Feng, H.; Wu, S.; Yang, J.; Zou, Y.; Siddique, K.H.M. Ephemeral Gully Recognition and Accuracy Evaluation Using Deep Learning in the Hilly and Gully Region of the Loess Plateau in China. *Int. Soil Water Conserv. Res.* **2022**, *10*, 371–381. [\[CrossRef\]](#)
15. Singh, G.; Singh, S.; Sethi, G.K.; Sood, V. Detection and Mapping of Agriculture Seasonal Variations with Deep Learning–Based Change Detection Using Sentinel-2 Data. *Arab. J. Geosci.* **2022**, *15*, 825. [\[CrossRef\]](#)
16. Efthimiou, N.; Psomiadis, E.; Papanikolaou, I.; Soulis, K.X.; Borrelli, P.; Panagos, P. A New High Resolution Object-Oriented Approach to Define the Spatiotemporal Dynamics of the Cover-Management Factor in Soil Erosion Modelling. *CATENA* **2022**, *213*, 106149. [\[CrossRef\]](#)
17. Diaz-Varela, R.A.; Zarco-Tejada, P.J.; Angileri, V.; Loudjani, P. Automatic Identification of Agricultural Terraces through Object-Oriented Analysis of Very High Resolution DSMs and Multispectral Imagery Obtained from an Unmanned Aerial Vehicle. *J. Environ. Manag.* **2014**, *134*, 117–126. [\[CrossRef\]](#)
18. Hossain, M.D.; Chen, D. Segmentation for Object-Based Image Analysis (OBIA): A Review of Algorithms and Challenges from Remote Sensing Perspective. *ISPRS J. Photogramm. Remote Sens.* **2019**, *150*, 115–134. [\[CrossRef\]](#)
19. Ye, Z.; Yang, K.; Lin, Y.; Guo, S.; Sun, Y.; Chen, X.; Lai, R.; Zhang, H. A Comparison between Pixel-Based Deep Learning and Object-Based Image Analysis (OBIA) for Individual Detection of Cabbage Plants Based on UAV Visible-Light Images. *Comput. Electron. Agric.* **2023**, *209*, 107822. [\[CrossRef\]](#)
20. Saeed, M.U.; Bin, W.; Sheng, J.; Albarakati, H.M.; Dastgir, A. MSFF: An Automated Multi-Scale Feature Fusion Deep Learning Model for Spine Fracture Segmentation Using MRI. *Biomed. Signal Process. Control.* **2024**, *91*, 105943. [\[CrossRef\]](#)
21. Wei, S.; Luo, M.; Zhu, L.; Yang, Z. Using Object-Oriented Coupled Deep Learning Approach for Typical Object Inspection of Transmission Channel. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *116*, 103137. [\[CrossRef\]](#)
22. Kotera, A.; Nagano, T.; Hanittin, P.; Koontanakulvong, S. Assessing the Degree of Flood Damage to Rice Crops in the Chao Phraya Delta, Thailand, Using MODIS Satellite Imaging. *Paddy Water Environ.* **2016**, *14*, 271–280. [\[CrossRef\]](#)

23. Guan, H.; Huang, J.; Li, L.; Li, X.; Miao, S.; Su, W.; Ma, Y.; Niu, Q.; Huang, H. Improved Gaussian Mixture Model to Map the Flooded Crops of VV and VH Polarization Data. *Remote Sens. Environ.* **2023**, *295*, 113714. [\[CrossRef\]](#)
24. Keyport, R.N.; Oommen, T.; Martha, T.R.; Sajinkumar, K.S.; Gierke, J.S. A Comparative Analysis of Pixel- and Object-Based Detection of Landslides from Very High-Resolution Images. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *64*, 1–11. [\[CrossRef\]](#)
25. Drăguț, L.; Tiede, D.; Levick, S.R. ESP: A Tool to Estimate Scale Parameter for Multiresolution Image Segmentation of Remotely Sensed Data. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 859–871. [\[CrossRef\]](#)
26. Wang, B.; Zhang, Z.; Wang, X.; Zhao, X.; Yi, L.; Hu, S. Object-Based Mapping of Gullies Using Optical Images: A Case Study in the Black Soil Region, Northeast of China. *Remote Sens.* **2020**, *12*, 487. [\[CrossRef\]](#)
27. Chen, L.-C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018.
28. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.-C. MobileNetV2: Inverted Residuals and Linear Bottlenecks. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 4510–4520.
29. Guirado, E.; Blanco-Sacristán, J.; Rodríguez-Caballero, E.; Tabik, S.; Alcaraz-Segura, D.; Martínez-Valderrama, J.; Cabello, J. Mask R-CNN and OBIA Fusion Improves the Segmentation of Scattered Vegetation in Very High-Resolution Optical Sensors. *Sensors* **2021**, *21*, 320. [\[CrossRef\]](#)
30. Azeez, O.S.; Shafri, H.Z.M.; Alias, A.H.; Haron, N.A.B. Integration of Object-Based Image Analysis and Convolutional Neural Network for the Classification of High-Resolution Satellite Image: A Comparative Assessment. *Appl. Sci.* **2022**, *12*, 10890. [\[CrossRef\]](#)
31. Robson, B.A.; Bolch, T.; MacDonell, S.; Hölbling, D.; Rastner, P.; Schaffer, N. Automated Detection of Rock Glaciers Using Deep Learning and Object-Based Image Analysis. *Remote Sens. Environ.* **2020**, *250*, 112033. [\[CrossRef\]](#)
32. Guirado, E.; Tabik, S.; Alcaraz-Segura, D.; Cabello, J.; Herrera, F. Deep-Learning Versus OBIA for Scattered Shrub Detection with Google Earth Imagery: Ziziphus Lotus as Case Study. *Remote Sens.* **2017**, *9*, 1220. [\[CrossRef\]](#)
33. Chollet, F. Xception: Deep Learning with Depthwise Separable Convolutions. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition CVPR, Honolulu, HI, USA, 21–26 July 2017; pp. 1800–1807.
34. Sutaji, D.; Yıldız, O. LEMOXINET: Lite Ensemble MobileNetV2 and Xception Models to Predict Plant Disease. *Ecol. Inform.* **2022**, *70*, 101698. [\[CrossRef\]](#)
35. Hasanuzzaman, M.; Pratim Adhikary, P.; Kumar Shit, P. Gully Erosion Susceptibility Mapping and Prioritization of Gully-Dominant Sub-Watersheds Using Machine Learning Algorithms: Evidence from the Silabati River (Tropical River, India). *Adv. Space Res.* **2024**, *73*, 1653–1666. [\[CrossRef\]](#)
36. Bammou, Y.; Benzougagh, B.; Abdessalam, O.; Brahim, I.; Kader, S.; Spalevic, V.; Sestras, P.; Ercişli, S. Machine Learning Models for Gully Erosion Susceptibility Assessment in the Tensift Catchment, Haouz Plain, Morocco for Sustainable Development. *J. Afr. Earth Sci.* **2024**, *213*, 105229. [\[CrossRef\]](#)
37. Zhao, W.; Du, S.; Emery, W.J. Object-Based Convolutional Neural Network for High-Resolution Imagery Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 3386–3396. [\[CrossRef\]](#)
38. Ghorbanzadeh, O.; Shahabi, H.; Crivellari, A.; Homayouni, S.; Blaschke, T.; Ghamisi, P. Landslide Detection Using Deep Learning and Object-Based Image Analysis. *Landslides* **2022**, *19*, 929–939. [\[CrossRef\]](#)
39. Sicard, P.; Coulibaly, F.; Lameiro, M.; Araminien, V.; De Marco, A.; Sorrentino, B.; Anav, A.; Manzini, J.; Hoshika, Y.; Moura, B.B.; et al. Object-Based Classification of Urban Plant Species from Very High-Resolution Satellite Imagery. *Urban For. Urban Green.* **2023**, *81*, 127866. [\[CrossRef\]](#)
40. Yin, L.; Wang, L.; Li, T.; Lu, S.; Yin, Z.; Liu, X.; Li, X.; Zheng, W. U-Net-STN: A Novel End-to-End Lake Boundary Prediction Model. *Land* **2023**, *12*, 1602. [\[CrossRef\]](#)
41. Tiwari, A.; Silver, M.; Karnieli, A. A Deep Learning Approach for Automatic Identification of Ancient Agricultural Water Harvesting Systems. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *118*, 103270. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Huang, D.; Su, L.; Fan, H.; Zhou, L.; Tian, Y. Identification of Topographic Factors for Gully Erosion Susceptibility and Their Spatial Modelling Using Machine Learning in the Black Soil Region of Northeast China. *Ecol. Indic.* **2022**, *143*, 109376. [\[CrossRef\]](#)
43. Chen, W.; Lei, X.; Chakraborty, R.; Chandra Pal, S.; Sahana, M.; Janizadeh, S. Evaluation of Different Boosting Ensemble Machine Learning Models and Novel Deep Learning and Boosting Framework for Head-Cut Gully Erosion Susceptibility. *J. Environ. Manag.* **2021**, *284*, 112015. [\[CrossRef\]](#)
44. Li, J.; Chen, Y.; Jiao, J.; Chen, Y.; Chen, T.; Zhao, C.; Zhao, W.; Shang, T.; Xu, Q.; Wang, H.; et al. Gully Erosion Susceptibility Maps and Influence Factor Analysis in the Lhasa River Basin on the Tibetan Plateau, Based on Machine Learning Algorithms. *CATENA* **2024**, *235*, 107695. [\[CrossRef\]](#)
45. Cheng, Y.; Wang, W.; Ren, Z.; Zhao, Y.; Liao, Y.; Ge, Y.; Wang, J.; He, J.; Gu, Y.; Wang, Y.; et al. Multi-Scale Feature Fusion and Transformer Network for Urban Green Space Segmentation from High-Resolution Remote Sensing Images. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *124*, 103514. [\[CrossRef\]](#)

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.