DeblurGAN-v2-Based Landslide Recognition Method for Blurred Images

Zichen Song [1,*], Dongqing Fang [2]

[1] College of Civil and Hydraulic Engineering, Hefei University of Technology, Hefei, China, 230009

[2] TongLing Nonferrous Metals Group Holding Co, Tongling, China, 244000

* Corresponding Author Email: song_zc0517@163.com

Abstract. Aiming at the problem of motion-blurred images due to camera shake, which affects the performance of landslide recognition when using computer vision technology for landslide recognition, a fuzzy image landslide recognition method combining DeblurGAN-v2 and YOLOv5 is proposed. Firstly, the custom camera image landslide dataset is trained by YOLOv5 deep learning object detection algorithm and a landslide recognition model is established; then the DeblurGAN-v2 algorithm is used to deblur the blurred image and obtain the deblurred image; finally, the deblurred image is inputted into the landslide recognition model to identify and locate the landslide area in the image. The experimental results show that the accuracy of landslide recognition of blurred image is 35%, and the accuracy of landslide recognition of deblurred image is 82.5%. Meanwhile, the statistical landslide confidence is found that the mean and variance of confidence are 0.26 and 0.13 under blurred image, and 0.67 and 0.11 under deblurred image, respectively, and the proposed method is able to effectively improve the accuracy of blurred image landslide recognition and the confidence. confidence, and has better stability, which provides an important technical solution for real-time monitoring, early warning and rescue decision-making of landslide disasters.

Keywords: Landslide Recognition, Deep Learning, YOLOv5, DeblurGAN-v2.

1. Introduction

Landslide is one of the common geological disasters, which not only threatens the life and property safety of the residents, but also may damage the infrastructure, resulting in a series of serious consequences, so the identification of landslide disaster is the key link of disaster prevention and mitigation of losses. At present, the identification of landslide disaster is the key link of disaster prevention and loss mitigation. The identification of existing landslide hazards is mainly realized by remote sensing images of existing landslides, for example, Pang et al. [1] produced a sample dataset of landslides using remote sensing images of the Hokkaido region in Japan, and constructed a landslide identification model through the YOLOv3 algorithm to realize the target identification of landslides; Wang et al. [2] integrated deep learning with remote sensing imagery, employing SHAP interpretability analysis and the entropy-weighted TOPSIS method to investigate causative factors and identification outcomes of landslides in Nyingchi City. Tang et al. [3] used high-resolution satellite images to construct a landslide sample dataset and demonstrated that the selection of the YOLOv5s model can obtain high recognition accuracy with fewer samples. However, the acquisition frequency of remote sensing images is not enough to support real-time recognition, which makes it difficult for landslides, a rapidly developing geohazard, to be recognized immediately after their occurrence. For this reason, this paper establishes a landslide recognition model for camera images through the YOLOv5 deep learning target recognition algorithm. At the same time, taking into account the camera in the field image acquisition process, by the wind and vibration and many other factors, resulting in camera shake, and then make the image appear motion blur, affecting the effect of landslide identification. Therefore it is necessary to de-blur the image first to improve the image quality.

Earlier conventional deblurring algorithms such as Wiener filtering, Richardson-Lucy deconvolution recovered clear images by mathematical optimization under the assumption that the blurring kernel is known [4]. In practice, the blur kernel is often unknown or difficult to estimate

accurately. With the development of deep learning algorithms, people are beginning to investigate the use of deep learning and neural networks to solve the problem of image blurring. Kong et al. [5] proposed a frequency-domain based Transformer method (FFTformer) to enhance the image deblurring effect by Frequency-domain Self-attention Solver and Discriminative Frequency-domain Feed-forward Network , but the method is mainly optimized for general blurring scenarios, and it is weak for extreme blurring; Kupyn et al. [6] introduced Feature Pyramid Network to image deblurring for the first time on the basis of DeblurGAN proposing DeblurGAN-v2, which achieves excellent deblurring results when paired with a complex backbone like Inception-ResNet-v2; Zamir et al. [7] proposes MPRNet. This algorithm progressively restores features of blurred images through a multi-stage network architecture, but its multi-stage design and complex modules result in increased computational resource requirements, limiting its applications in resource-constrained environments. [8] proposes the Sharpformer model to remove motion blur by directly learning image global features and adaptive local features through the Transformer module. However, this algorithm is more sensitive to the noise in the image, which is prone to result in less accurate recovery of image details.

In this paper, YOLOv5 target recognition algorithm is used to train the homemade camera image landslide dataset and construct a landslide recognition model, while taking into account the problem of motion blur affecting landslide recognition caused by camera shake, a DeblurGAN-v2 based fuzzy image landslide recognition method is proposed, and a series of experimental analyses are carried out to verify the effectiveness of the proposed method.

2. Principles of the methodology

2.1. DeblurGAN-v2 Deblurring Algorithm

DeblurGAN-v2 is an end-to-end deblurring method based on Generative Adversarial Networks (GAN). The method follows the basic architecture of traditional GAN, which consists of two core components, Generator (G) and Discriminator (D). In the traditional GAN framework, the Generator is responsible for generating realistic samples from the input data, while the Discriminator is used to distinguish the generated samples from the real ones. The objective function underlying the framework can be expressed as:

$$\min_G \max_D V(D,G) = \tilde{a}_{x\sim p_{data}(x)}\left[\log D(x)\right] + \tilde{a}_{z\sim p_z(z)}\left[\log(1-D(G(z)))\right] \tag{1}$$

Where $V(D,G)$ denotes the adversarial goal of the Discriminator and Generator, x and z represent the real data and the input noise of the generator, respectively. $\tilde{a}_{x\sim p_{data}(x)}$ denotes the expectation of the real data distribution $p_{data}(x)$, $\tilde{a}_{z\sim p_z(z)}$ denotes the expectation of the noise prior distribution $p_z(z)$.

This objective function enables the generator to learn mappings that approximate the real image through adversarial training. However, traditional GANs are prone to problems such as gradient vanishing and pattern collapse during training, for this reason DeblurGAN-v2 introduces Relativistic Conditional GAN (RCGAN) and employs Least-Square Loss (LSGAN) in the discriminator to obtain smoother gradient, producing the new RaGAN-LS loss:

$$L_D^{RaLSGAN} = \tilde{a}_{x\sim P_{data}(x)}\left[(D(x)-\tilde{a}_{z\sim p_z(z)}D(G(z))-1)^2\right] + \tilde{a}_{z\sim P_z(z)}\left[(D(G(z))-\tilde{a}_{x\sim p_{data}(x)}D(x)+1)^2\right] \tag{2}$$

In terms of network structure design, the generator of DeblurGAN-v2 consists of two main parts: an encoder-decoder architecture for extracting global context information, and a high-resolution branch for preserving high-resolution local details. To realize multi-scale feature fusion, DeblurGAN-v2 introduces Feature Pyramid Network (FPN) into the generator for the first time. Assuming that the feature maps generated at different scales are $F_1, F_2, \ldots, F_N$, the process of fusion after upsampling can be written as:

$$F_{\text{agg}} = Concat(Upsample(F_1), Upsample(F_2), \ldots, Upsample(F_N)) \tag{3}$$

The fused features are then further processed using the convolution module $\mathrm{H}(\cdot)$ and the input image I_{in} is summed with the residuals learned by the network using a direct jump join to obtain the deblurred output image I_{out}:

$$I_{out} = I_{in} + \mathrm{H}(F_{agg}) \tag{4}$$

Also, to ensure that the generated image is pixel and perceptually consistent with the real image, a hybrid loss function is introduced during the training process. The total loss function of the generator is:

$$L_G = 0.5 * L_p + 0.006 * L_X + 0.01 * L_{adv} \tag{5}$$

Where L_p is the mean square error (MSE) loss, L_{x} is the perceptual loss, and item L_{adv} contains the global and local discriminator losses.

In terms of discriminator design, a dual-scale discriminator structure is adopted, where one branch performs global discrimination for the whole image and the other branch performs discrimination at the local patch level to fully utilize the global structure of the image and the local detail information, so as to more accurately distinguish the real image from the generated image.

2.2. YOLOv5 target detection algorithm

YOLOv5 algorithm is a single-stage real-time target recognition algorithm based on the mechanism of having anchor frames [9], similar to other versions of the YOLO series, it still inputs the whole image into the network through convolutional neural network (CNN) for regression prediction without using candidate frames or region proposal networks (RPNs) like traditional methods. The YOLOv5 algorithm flow is as follows:

1) Different types of landslide images are labeled to get actual boxes of different sizes.

2) Clustering is performed based on the aspect ratio to obtain the a priori frame (c, x_p, y_p, w_p, h_p) with greater probability of occurrence that has a good match to the target object and use it as a benchmark [10]. The training is done so that the confidence level of the grid where the center of the target is located is 1, and the confidence level of the other grids is set to 0. The target identified in this paper focuses only on 1 category, landslides, and the value of c is taken as 1.

3) Perform $S \times S$ gridding on the input image. Assuming that the model clustering yields N prior frames, each grid corresponds to have N prior frames (bounding boxes of different sizes and different aspect ratios).

4) Based on the actual frames, a priori frames, Eqs. (6) and (7), the target detection frames can be obtained.

$$\begin{cases} x_D = \sigma(x_t) + x_C \\ y_D = \sigma(y_t) + y_C \\ w_D = w_P e^{wt} \\ h_D = h_P e^{ht} \end{cases} \tag{6}$$

Among them,

$$\begin{cases} x_t = \log \dfrac{x_G - x_C}{1-(x_G - x_C)} \\ y_t = \log \dfrac{y_G - y_C}{1-(y_G - x_C)} \\ w_t = \log \dfrac{w_G}{w_P} \\ h_t = \log \dfrac{h_G}{h_P} \end{cases} \tag{7}$$

Where (x_C, y_C) are the coordinates of the upper left corner point of the grid where the center of the target actual frame (x_G, y_G) is located; $\sigma(\bullet)$ is the sigmod function [11].

5) Iterate through all the images and grids to calculate the loss, and select the recognition box with the smallest loss value as the optimal recognition box.

3. Automatic Landslide Identification Model

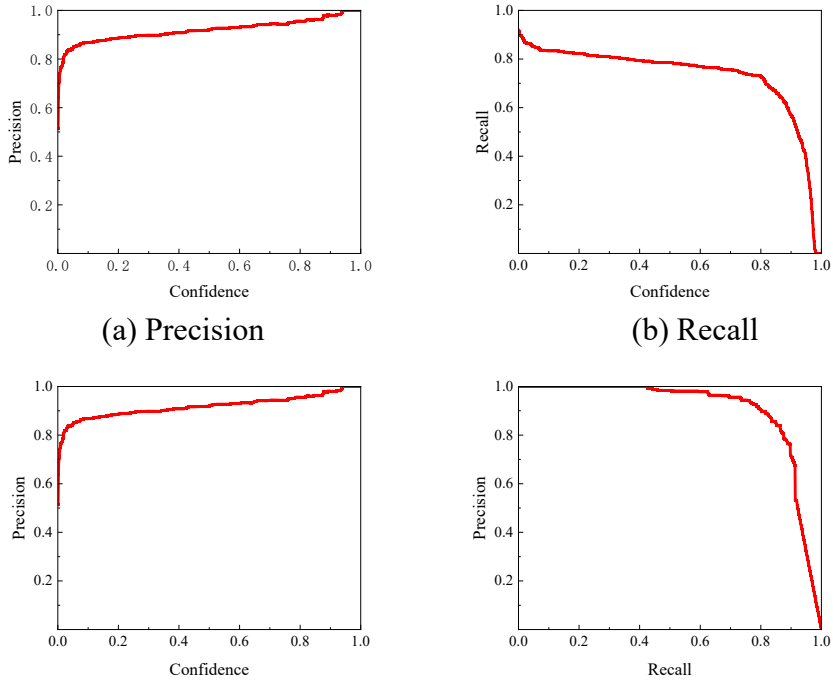
3.1. Dataset and experimental platform

The dataset used was obtained by expanding 386 photos of landslides around a railroad in Inner Mongolia taken by a Canon EOS5D MARK II camera. By rotating, flipping, adding noise, adjusting brightness, random pixel zeroing, and combining different ways of pre-processing the landslide images, we can expand the scale of the landslide dataset, improve the generalization ability of the model, and effectively improve the stability of the recognition method, and the dataset consists of 2797 landslide images after the expansion. Among them, the dataset is randomly divided into training set and validation set in the ratio of 7:3.

The experiments use YOLOv5s architecture as the benchmark model, set the batch size to 32, the number of training rounds to 300, and the initial learning rate to 0.001. The hardware platform configuration includes a 12-core Intel Xeon Platinum 8352V processor, an NVIDIA RTX 4090 graphics card (24GB video memory) and 90GB memory space, and the software environment is Windows 10 operating system with Python 3.8 and CUDA 11.0 computing framework.

3.2. Model performance analysis

In this experiment, four core indicators, Precision (P, Precision), Recall (R, Recall), F1Score and mean Average Precision (mAP), are used to evaluate the performance of the landslide identification model from different perspectives, respectively [12]. The analysis results are presented in Figure 1. The precision reaches 1.0 at a confidence threshold of 0.942, achieving zero false positives for landslide detection, while the recall drops to 0.424, resulting in a significant increase in missed detections. When the confidence threshold is set to 0, the recall achieves 0.919 with minimal missed detections, but the precision decreases to 0.516, leading to more false positives. The F1-score peaks at 0.854 with a confidence threshold of 0.217 and maintains stable performance (F1 > 0.8) across a wide confidence range [0.005-0.836], demonstrating the model's strong robustness to threshold variations within this interval - this range should be considered for practical deployment. The model achieves mAP@0.5 of 0.905 and mAP@0.5:0.95 of 0.775, indicating excellent landslide detection capability, though with limited performance in precisely quantifying landslide boundaries.



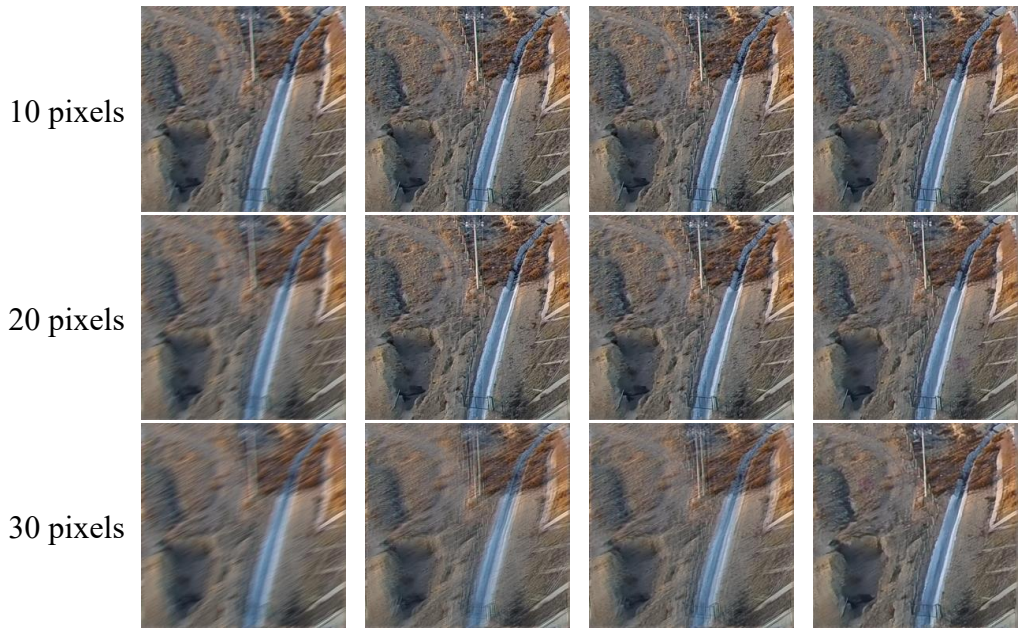
(a) Precision (b) Recall

(c) F1 score (d) Precision-recall curve

Figure 1. Graph of evaluation indexes of landslide model

4. Experimental analysis

4.1. Deblurring Quality Evaluation



(a) FFTformer (b) MPRNet (c) DeblurGAN-v2

Figure 2. De-blurring effect of different algorithms

In order to compare the deblurring performance of different algorithms, deblurring experiments were carried out in this study. The experiments first synthesized three images with different blurring levels using original clear images, corresponding to blur radii of 10 pixels, 20 pixels, and 30 pixels, respectively. Subsequently, these blurred images were processed using various deblurring algorithms such as FFTformer, MPRNet, DeblurGAN-v2, etc., and the results are shown in Fig. 2. The experimental results show that all the algorithms are effective in improving the image quality and

restoring the image features to varying degrees. When slightly blurring the image (blurring 10, 20 pixels), the de-blurring effect of each algorithm is relatively close, and it is difficult to distinguish the difference with the naked eye. However, when dealing with images with a high degree of blurring (blurring 30 pixels), DeblurGAN-v2 shows a significant advantage in that the recovered image is richer in detail and clearer in structure, and the visual quality is significantly better than that of the other algorithms.

To quantitatively assess the performance of different deblurring algorithms, this study employs three standard evaluation metrics: Mean Square Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM) [13]. MSE measures the difference between deblurred images and their original sharp counterparts, with smaller values indicating superior reconstruction quality. PSNR quantifies the level of image distortion, where higher values correspond to better image quality. SSIM evaluates the structural similarity between images, with values approaching 1 representing better structural preservation. As presented in Table.1, FFTformer demonstrates superior performance in all three metrics (MSE, PSNR, and SSIM) when processing images with blur radii of 10 and 20 pixels, confirming its enhanced restoration capability for mildly blurred images. However, for images with a 30-pixel blur radius, FFTformer's performance is exceeded by DeblurGAN-v2, which exhibits superior deblurring effectiveness under severe blurring conditions.

Table 1. Image quality evaluation metrics

Evaluation Indicators	Blur Radius	Blurred Image	FFTformer	MPRNet	DeblurGAN-v2
MSE	10	560.4180	176.8998	187.6210	358.7600
	20	767.1749	266.6087	333.3644	522.2297
	30	878.7505	1039.1681	1079.1841	647.1818
SSIM	10	0.5162	0.8560	0.8429	0.6876
	20	0.3512	0.7680	0.7009	0.5305
	30	0.2895	0.2224	0.2081	0.4189
PSNR	10	20.6457	25.6535	25.3980	22.5828
	20	19.2819	23.8721	22.9016	20.9522
	30	18.6921	17.9639	17.7998	20.0205

4.2. Confidence Analysis

(a)Blurred Image (b)FFTformer (c) MPRNet (d)DeblurGAN-v2

Figure 3. Schematic comparison of landslide identification results

In order to further validate the effectiveness of the proposed method, a landslide identification confidence comparison experiment is conducted in this paper, and in order to maintain the consistency of the calculation, the confidence level of missed landslides is regarded as zero. With the IoU loss threshold set at 0.45 and confidence threshold at 0.50, landslide recognition was performed on both 40 originally blurred landslide images and 40 deblurred landslide images. The model generated six distinct sets of recognition results, with partial landslide recognition outcomes illustrated in Figure 3.

The experiments evaluated the practical application of different algorithms by comparing the landslide recognition confidence level of the original blurred images with that of the deblurred processed images. As shown in Table 2, the image processed with DeblurGAN-v2 algorithm performs optimally: the landslide recognition model recognizes a total of 33 landslide targets with an accuracy of 82.5% and an average confidence level of 0.67, which is higher than that of other algorithms. At the same time, the confidence variance is 0.11, indicating that its recognition results have better stability. Combining all the experimental results, it can be seen that DeblurGAN-v2 shows better adaptability and practicality in the deblurring task, and can better meet the needs of practical applications.

Table 2. Confidence analysis

Image Type	Number of Labels	Accuracy	Mean Confidence	Confidence Variance
Blurred Image	14	35%	0.26	0.13
FFTformer	16	40%	0.32	0.16
MPRNet	29	72.5%	0.58	0.14
DeblurGAN-v2	33	82.5%	0.67	0.11

5. Conclusion

Aiming at the problem of landslide detection using computer vision technology, where motion blur caused by camera shake degrades detection accuracy, a deblurring and landslide detection method based on DeblurGAN-v2 is proposed. Firstly, a landslide detection model is constructed via the YOLOv5 object detection algorithm; secondly, the blurred images are restored using DeblurGAN-v2 to generate clear images; finally, the processed images are fed into the detection model to achieve precise localization of landslide areas. The proposed method significantly improves detection performance on blurred images and provides efficient technical support for disaster rescue decision-making. However, two major challenges remain: the limited scale and diversity of landslide image datasets, and the predominance of close-range shots in existing data. In long-range scenarios, the typical morphological features of landslides exhibit low saliency in images due to insufficient camera resolution, leading to detection failures. Future work may focus on expanding multi-scale landslide data and optimizing model architecture to enhance performance.

References

[1] Pang D, Liu G, He J, et al. Automatic remote sensing identification of co-seismic landslides using deep learning methods [J]. Forests, 2022, 13 (8): 1213.

[2] Wang Y, Gao H, Liu S, et al. Landslide detection based on deep learning and remote sensing imagery: A case study in Linzhi City [J]. Natural Hazards Research, 2025, 5 (1): 95-108.

[3] Tang Fengshun, Hao Lina, Song Yuyang, et al. Application of target recognition algorithm in landslide identification [J]. Journal of Lanzhou University (Natural Science Edition), 2024, 60 (02): 229-234.

[4] Richardson W H. Bayesian-based iterative method of image restoration [J]. Journal of the optical society of America, 1972, 62 (1): 55-59.

[5] Kong L, Dong J, Ge J, et al. Efficient frequency domain-based transformers for high-quality image deblurring [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 5886-5895.

[6] Kupyn O, Martyniuk T, Wu J, et al. Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better [C] // Proceedings of the IEEE/CVF international conference on computer vision. 2019: 8878-8887.

[7] Zamir S W, Arora A, Khan S, et al. Multi-stage progressive image restoration [C] // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021: 14821-14831.

[8] Yan Q, Gong D, Wang P, et al. SharpFormer: Learning local feature preserving global representations for image deblurring [J]. IEEE Transactions on Image Processing, 2023, 32: 2857-2866.

[9] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection [C] // Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 779-788.

[10] Ou S, Gao Y, Zhang Z, et al. Polyp-yolov5-tiny: A lightweight model for real-time polyp detection [C] // 2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA). IEEE, 2021, 2: 1106-1111.

[11] Zhang H, Tian M, Shao G, et al. Target detection of forward-looking sonar image based on improved YOLOv5 [J]. IEEE Access, 2022, 10: 18023-18034.

[12] Xu Siqing, Ke deping, Xu jian. Wind power creep event recognition based on YOLOv5s [J]. Journal of Wuhan University (Engineering Edition), 2022, 55 (09): 910-918.

[13] Cha Tibo, Lin Luo, Yang Kai, et al. Image reconstruction algorithm based on improved super-resolution generative adversarial network [J]. Advances in Lasers and Optoelectronics, 2021, 58 (08): 93-103.