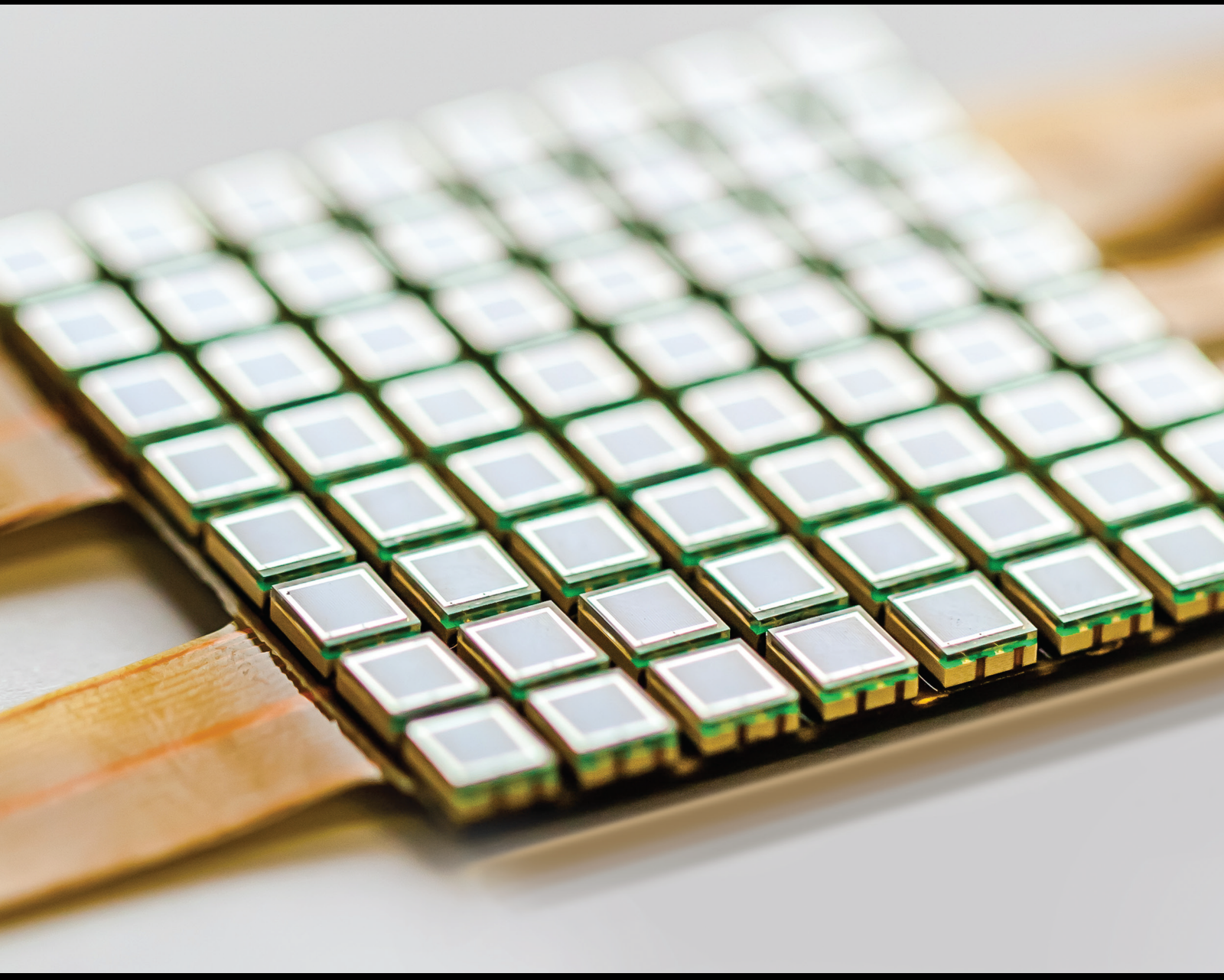
Journal of Sensors

Advanced Sensor Technologies in Geospatial Sciences and Engineering 2021

Lead Guest Editor: Hyung-Sup Jung
Guest Editors: Sang-Hoon Hong and Lei Zhang



Advanced Sensor Technologies in Geospatial Sciences and Engineering 2021

Journal of Sensors

Advanced Sensor Technologies in Geospatial Sciences and Engineering 2021

Lead Guest Editor: Hyung-Sup Jung
Guest Editors: Sang-Hoon Hong and Lei Zhang

Contents

Hindawi
Journal of Sensors
Volume 2021, Article ID 7331459, 18 pages
https://doi.org/10.1155/2021/7331459

Research Article

Bit-Error Aware Lossless Image Compression with 2D-Layer-Block Coding

Jungan Chen,[1] **Jean Jiang**,[2] **Xinnian Guo**,[3] **and Lizhe Tan**[4]

[1]*Department of Electronic and Computer Science, Zhejiang Wanli University, Ningbo, China*
[2]*College of Technology, Purdue University Northwest, Indiana, USA*
[3]*Department of Electronic Information Engineering, Huaiyin Institute of Technology, Huaian, China*
[4]*Department of Electrical and Computer Engineering, Purdue University Northwest, Indiana, USA*

Correspondence should be addressed to Jungan Chen; friendcen21@hotmail.com

Received 23 June 2021; Revised 10 October 2021; Accepted 12 October 2021; Published 3 November 2021

Academic Editor: Sang-Hoon Hong

With IoT development, it becomes more popular that image data is transmitted via wireless communication systems. If bit errors occur during transmission, the recovered image will become useless. To solve this problem, a bit-error aware lossless image compression based on bi-level coding is proposed for gray image compression. But bi-level coding has not considered the inherent statistical correlation in 2D context region. To resolve this shortage, a novel variable-size 2D-block extraction and encoding method with built-in bi-level coding for color image is developed to decrease the entropy of information and improve the compression ratio. A lossless color transformation from RGB to the YCrCb color space is used for the decorrelation of color components. Particularly, the layer-extraction method is proposed to keep the Laplacian distribution of the data in 2D blocks which is suitable for bi-level coding. In addition, optimization of 2D-block start bits is used to improve the performance. To evaluate the performance of our proposed method, many experiments including the comparison with state-of-the-art methods, the effects with different color space, etc. are conducted. The comparison experiments under a bit-error environment show that the average compression rate of our method is better than bi-level, Jpeg2000, WebP, FLIF, and L3C (deep learning method) with hamming code. Also, our method achieves the same image quality with the bi-level method. Other experiments illustrate the positive effect of built-in bi-level encoding and encoding with zero-mean values, which can maintain high image quality. At last, the results of the decrease of entropy and the procedure of our method are given and discussed.

1. Introduction

With cloud computing and Internet of Things (IoT) development, the requirement for data transmission and storage is increasing. Fast and efficient compression of data plays a very important role in many applications. For instance, image data compression has been used in many areas such as medical, satellite remote sensing, and multimedia.

There are many methods to compress image data including prediction-based, transformation-based, and other methods such as fractal image compression and deep learning with Auto Encoder (AE) [1, 2], Recurrent Neural Network (RNN), Convolutional Neural Network (CNN) [3], and Residual Neural Network (RestNet) [4]. The transformation-based method includes Discrete Cosine Transform (DCT), Karhunen-Loeve Transform (KLT), Hadamard transform, Slant transform, Haar transform, and singular value decomposition [5]. Usually, transformation-based or deep learning methods are used in lossy compression while prediction-based methods are used for lossless compression.

In some cases, lossless compression must be applied when data acquisition is expensive. For example, lossless image compression must be applied in aerial, medical, and space images [6, 7]. In industry, many engineered lossless compression methods including Portable Network Graphics (PNG), WebP [8], and Free Lossless Image Format (FLIF) [9] are used. Also, some deep learning-based lossless

compression methods [10–12] are early researched. As one classical method for lossless compression, the prediction-based method takes into account the difference between pixel values and their predicted values, which are generally smaller numbers than the pixel values themselves. Thus, each difference value needs a smaller number of bits to encode [13]. It mainly has three kinds of methods including context-based, least-square (LS)-based, and spatial structure-based. Among these methods, the method based on spatial structure with 2D context region is an effective solution to improve the compression ratio (CR) because of considering the inherent statistical correlation using block-based methods such as quadtree-based block [14], reference block [15], template-matching[16], and hierarchical decomposition [17]. Quadtree-based block and hierarchical decomposition methods split image to many subimages. And the reference block method considers the phenomenon that a physical object is constructed from numbers of structure components. Inspired from these methods, splitting image into many blocks where each block has similar color is taken as an effective method used by this work.

With IoT development, it becomes more usual that image data is transmitted through wireless communication systems and lossless image compression is used to improve transmission throughput. However, if bit errors occur in a wireless noisy channel during transmission, the recovered image will be damaged or become useless. So, lossless image compression must resolve the problem and keep the recovered image be useful. Most methods including engineered lossless compression methods and deep learning based methods are not suitable to transmission in noise channel; to the best of our knowledge, fewer researches have worked on this case except our previous work [7]. By protecting the key information bits with error control coding, our work proposed a bit-aware lossless image compression based on bi-level coding for gray image as a one-dimensional signal. In the coding method, only the linear predictive bi-level block coding parameters are encoded using (7,4) Hamming codes and residue sequences are left as they are to improve the performance of compression rate (CR). One reason for the efficiency of bi-level coding is that it uses the sparsity property of data which required fewer encoding bits.

In this work, we will use bi-level coding [7] for natural images with red (R), green (G), and blue (B) components. As R, G, and B are highly correlated, a linear transformation is applied to map RGB to other color space and achieve better CR [17, 18]. As discussed above, the spatial structure-based method with 2D context region is taken as an effective solution to improve CR. Therefore, image is split into many 2D blocks which has sparsity property and be suitable to be encoded with bi-level coding. Finally, a novel variable-size 2D-block extraction and encoding method with built-in bi-level coding is proposed to improve CR for color image and robust to bit-error environment. An important 2D-layer-block extraction method is used to split the image to many 2D blocks with similar color and keep the Laplacian or Gaussian distribution of data in one 2D block, which has sparsity property.

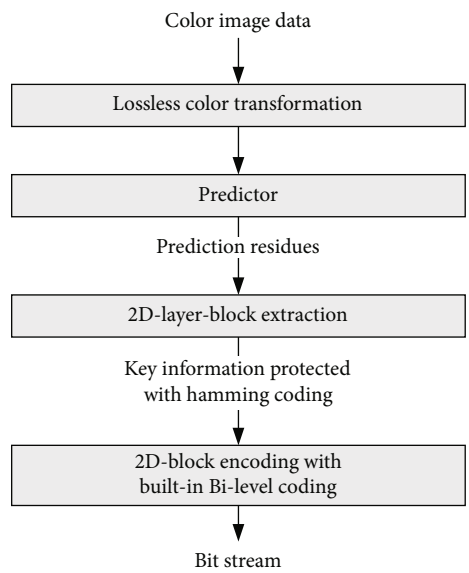


Figure 1: The proposed method.

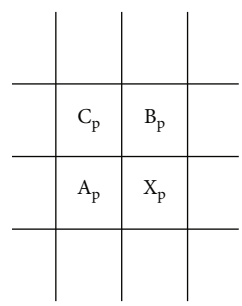


Figure 2: Neighboring pixels for the predictor.

The contributions of this paper are summarized as follows:

(1) For color image compression, a lossless color transformation from RGB to the YCrCb color space is used for the decorrelation of color components. The prediction-based method is used to remove data correlation and produce residue sequence

(2) To keep the data distribution with the sparsity property and be suitable for bi-level coding, a novel 2D-layer-block extraction method is proposed to keep the Laplacian or Gaussian distribution of data in 2D blocks. Furthermore, by rearranging the order of data encoded, the extraction method can decrease the entropy of data and improve CR

(3) A novel variable-size 2D-block encoding method with built-in bi-level is proposed to improve CR and robust to bit-error environment just as the bi-level coding method. The mean or min value in one 2D block and key information bits in built-in bi-level coding are protected with hamming code. So, the image can be recovered and useful

The rest of this paper is organized as follows. In Section 2, related works on lossless compression are discussed. In

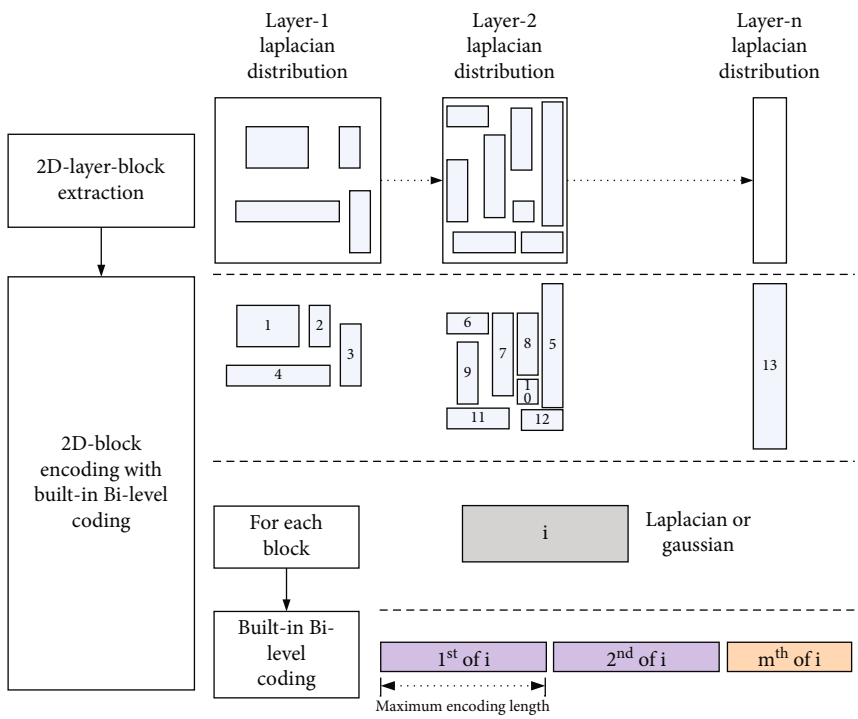


Figure 3: The description of 2D-block extraction and encoding.

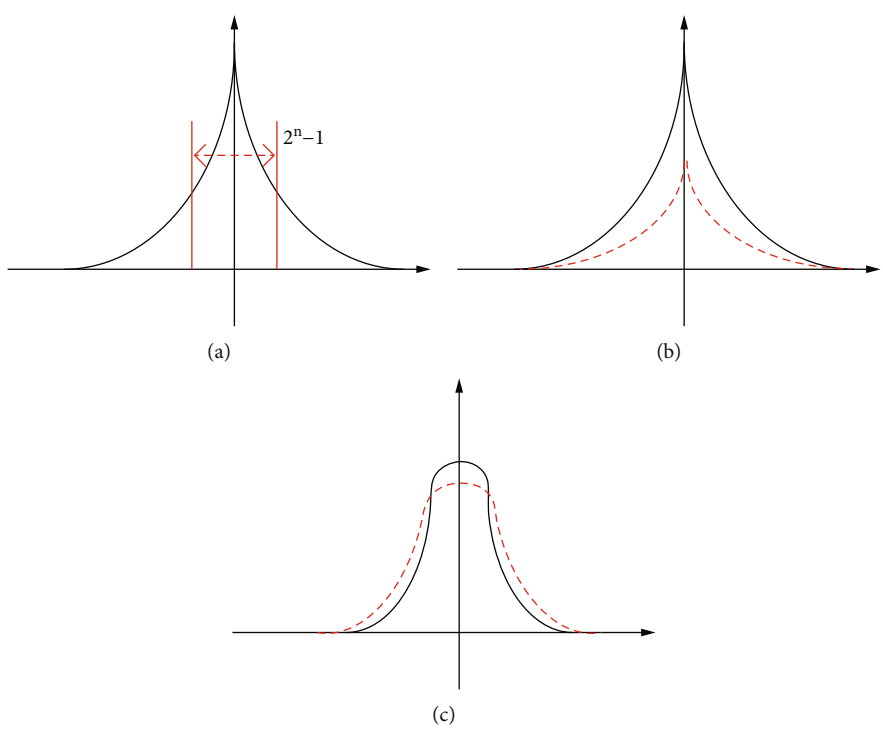


Figure 4: Data distribution. (a) data distribution in blocks; (b) data distribution after 2D-block extraction; (c) data in one block.

Section 3, the details of the proposed method are briefly introduced. In Section 4, several experiments including comparison and analysis of basic principles are conducted. Finally, the conclusion and future researches are drawn in Section 5.

2. Related Work on Lossless Compression

2.1. Prediction-Based Methods. The context-based adaptive prediction method is based on a static predictor which is usually a switching predictor able to adapt to several types

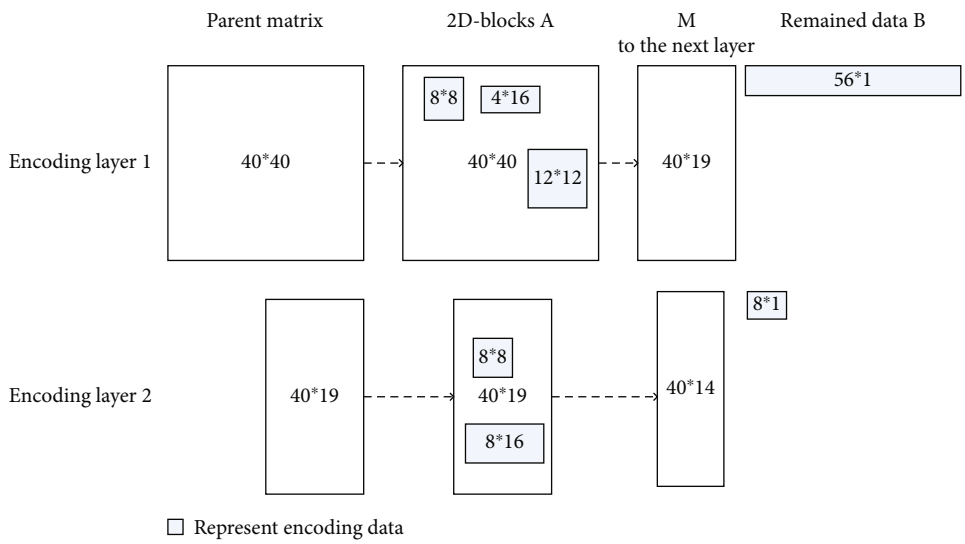


Figure 5: Example of extracting 2D blocks.

M = matrix of one channel

BLK-A = [] // blocks in every layer

ARY-B = [] // remaindata in every layer

For **bits** = 1:8

Extract blocks **A** with max(:)-min(:) = 2^**bits**-1 from **M**

M = reshaped matrix of data except the data (A and B)

BLK-A = { **BLK-A**, A}

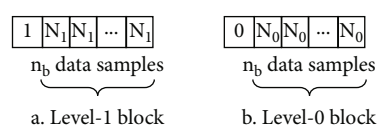
ARY-B = { **ARY-B**, B}

Figure 6: Extraction of 2D block.

of contexts, like horizontal edge, vertical edge, or smooth area. Many static predictors can be found in [6, 19]. Median edge detector (MED) used in LOCO-I uses only three causal pixels to determine a type of pixel area which is currently predicted [20]. LOCO-I is further improved and standardized as the JPEG-LS lossless compression algorithm, which has eight different predictive schemes including three one-dimensional and four tow-dimensional predictors [21]. To detect edges, Gradient Adjusted Predictor (GAP) embedded in the CALIC algorithm uses local gradient estimation and three heuristic-defined thresholds [22]. Gradient edge detection (GED) predictor combines simplicity of MED and efficiency of GAP [23]. In [19], the prediction errors are encoded using codes adaptively selected from the modified Golomb-Rice code family. To enable processing of images with higher bit depths, a simple context-based entropy coder is presented [6].

LS-based optimization is proposed as an approach to accommodate varying statistics of coding images. To reduce computational complexity, edge-directed prediction (EDP) initiates the LS optimization process only when the prediction error is beyond a preselected threshold [24]. In [25],

1. Divide a line of residue data with a length of $n_s = m \times n_b$ into m blocks, in which each block consists of n_b samples; that is, n_b is the block size. There are two types of blocks: the level-0 block and the level-1 block.



2. For a level-1 block, any sample in the block requires only N_1 bits ($N_1 < N_0$[original sample size]) to encode. Encode each sample using N_1 bits and add the prefix "1" to designate the block as the level-1 block.

3. For a level-0 block, at least one of the samples in the block needs more than N_1 bits to encode. Encode each sample in the block using N_0 bits and add the prefix "0" to indicate the level-0 block.

Figure 7: Bi-level coding rule.

$N^* = 8$ //the optimized value of start bits

For N = 1:7

Estimate p_0(more than N bits)

If $p_0 \times \lceil 1/\sqrt{(8-N)p_0} \rceil \leq 0.3$

stop and return $N^* = N$

Figure 8: Optimization of Start Bits.

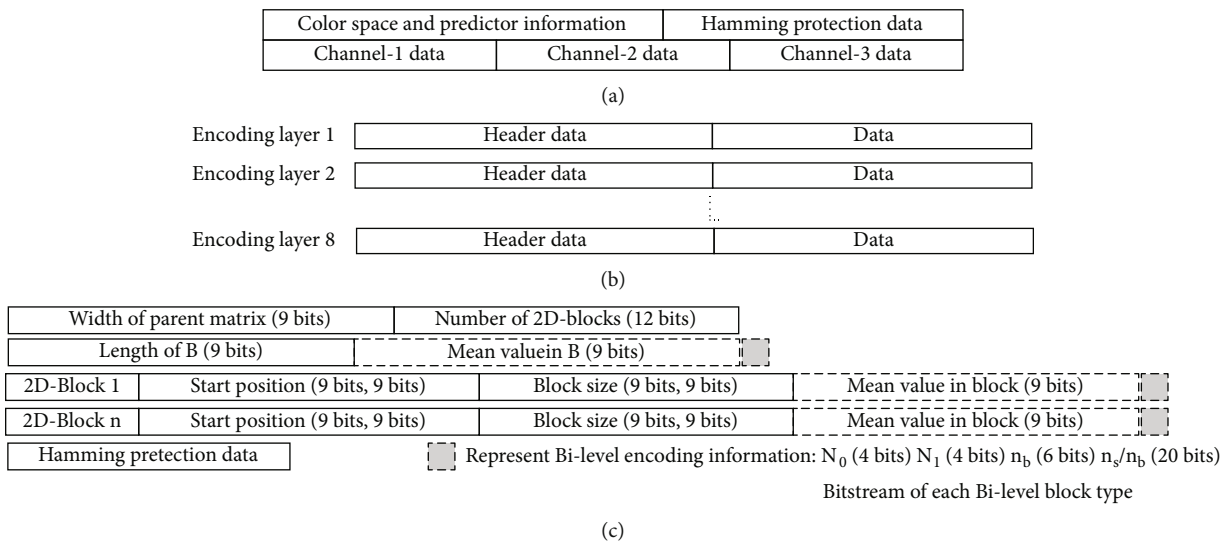


Figure 9: 2D-block encoding scheme: (a) data encoding of one image; (b) data encoding in one channel of image; (c) header data encoding of one layer.

Table 1: Images from Open Images and CLIC.

FileId	FileName	th_n	FileId	FileName	th_n
(1)	077a21a70e1b7f86 	102400	(7)	IMG_20160522_165515 	2024
(2)	3632d975bd623167 	10240	(8)	IMG_20170114_204505 	1536
(3)	488d4412c71bf3f1 	20480	(9)	IMG_20170725_123034 	3584
(4)	4e9226ecb7124541 	20480	(10)	IMG_20170730_133144 	20480
(5)	IMG_20131224_110518010_HDR 	20480	(11)	IMG_20170916_004924 	2024
(6)	IMG_20160312_095817 	4096	(12)	c1bfd9c6e7279456 	20480

the LS optimization is processed only when the coding pixel is around an edge or when the prediction error is large. And a switching coding scheme is further proposed that combines the advantages of both run-length and adaptive linear predictive coding [26]. Minimum Mean Square Error (MMSE) predictor uses least mean square principle to adapt k-order linear predictor coefficients for optimal prediction of the current pixel, from a fixed number of m causal neighbors [27]. The paper [28] presents a lossless coding method based on blending approach with a set of 20 blended predictors, such as recursive least squares (RLS) predictors and Context-Based Adaptive Linear Prediction (CoBALP+).

Although individual prediction is favored, the morphology of 2D context region would be destructed accordingly and inherent statistical correlation among the correlated region gets obscure. As an alternative, spatial structure has been considered to compensate the pixel-wise prediction [29]. In [14], quadtree-based variable block-size partitioning is introduced into the adaptive prediction technique to remove spatial redundancy in a given image and the resulting prediction errors are encoded using context-adaptive arithmetic coding. Inspired by the success of prediction by partial matching (PPM) in sequential compression, the paper [30] introduces the probabilistic modeling of the encoding symbol based on

Table 2: Comparison with prediction-based methods.

Image	Ref. [17]				Ref. [16]			Best CR(7,4)	Our
	Calic	jpeg2000	jpeg-xr	Ref. method	MRP	CoBaLP2	LOCO-I		
Lena	13.1787	13.5848	14.0942	13.6461	11.872	12.399	13.173	1.155179053	1.3876
Peppers	13.8661	14.8	15.3245	15.2102				0.989051407	1.2764
Mandrill	18.1511	18.0939	18.2553	18.5305	16.041	17.039	17.822	0.854952043	1.13
Barbara	14.9567	11.1612	12.1408	11.4575				1.228746525	1.2406
Airplane					10.121		11.06	1.355032676	1.5302
Average								1.116592	1.31296

Table 3: Comparison with engineering and deep learning methods.

	Our CR	Bilevel	PNG(7,4)	Jpeg2000(7,4)	WebP(7,4)	FLIF(7,4)	L3C(7,4)
(1)	1.5141	1.4958	1.201423	1.195981	1.209832	1.26147	1.201423
(2)	1.8985	1.9171	1.589528	1.591757	1.629015	1.697848	1.589528
(3)	2.0597	1.9877	1.762672	1.724237	1.767789	1.811668	1.762672
(4)	1.8084	1.7509	1.350212	1.362216	1.343821	1.373226	1.350212
(5)	1.4849	1.4478	1.098277	1.11419	1.104521	1.131249	1.098277
(6)	1.7418	1.7397	1.578617	1.500231	1.544284	1.571556	1.578617
(7)	1.4761	1.4772	1.122815	1.098124	1.086112	1.100787	1.122815
(8)	1.7614	1.7707	1.667265	1.566052	1.611617	1.709558	1.667265
(9)	1.295	1.2774	0.918406	0.904853	0.911615	0.928585	0.918406
(10)	1.5852	1.5504	1.19649	1.176481	1.216922	1.201159	1.19649
(11)	1.807	1.8237	1.740748	1.670904	1.780918	1.795483	1.740748
(12)	1.7631	1.7602	1.528354	1.461382	1.506129	1.576181	1.528354
Average	1.682933	1.66655	1.396234	1.363867	1.392715	1.429898	1.396234

its previous context occurrences. In [15], superspatial structure prediction is proposed to find an optimal prediction of the structure components, e.g., edges, patterns, and textures, within the previously encoded image regions instead of the spatial causal neighborhood. The paper [17] presents a lossless color image compression algorithm based on the hierarchical prediction and context-adaptive arithmetic coding. By exploiting the decomposition and combinatorial structure of the local prediction task and making the conditional prediction with multiple max-margin estimation in a correlated region, a structured set prediction model with max-margin Markov networks is proposed [29]. In [16], the image data is treated as an interleaved sequence generated by multiple sources and a new linear prediction technique combined with template-matching prediction and predictor blending method is proposed. Our method uses a variable-size 2D-block extraction and encoding method with built-in bi-level to improve the compression rate.

2.2. Engineered Lossless Compression Algorithms. PNG remove redundancies from the RGB representation with autoregressive filters and then the deflate algorithm based on the LZ77 algorithm and Huffman coding is used for data compression. Lossless WebP compression uses many types of transformation including spatial transformation, color transformation, green subtraction transformation, color indexing transformation, and color cache coding and then performs the entropy coding which uses a variation of LZ77 Huffman coding [8]. FLIF use Adam7 interlacing and YCoCg interleaving to traverse the image and perform entropy coding with "meta-adaptive near-zero integer arithmetic coding" (MANIAC) based on context-adaptive binary arithmetic coding CABAC [9].

2.3. Deep Learning-Based Lossless Compression. Huffman, arithmetic coding, and asymmetric numeral systems are the algorithms for implementing lossless compression, but they do not cater for latent variable models, so bits back with asymmetric numeral systems (BB-ANS) are proposed to solve the issue [10]. But BB-ANS become inefficient when the number of latent variables grows, to improve its performance on hierarchical latent variable models, Bit-Swap is proposed [11]. In contrast to these works focusing on smaller datasets, a fully parallel hierarchical probabilistic model (termed L3C) to enable practical compression on superresolution images [12].

3. Our Proposed Method

The proposed method is shown in Figure 1. In color image data, R, G, and B are highly correlated. So, their straightforward encoding is not efficient. Therefore, a linear transformation from RGB to the YCrCb color space is used for the

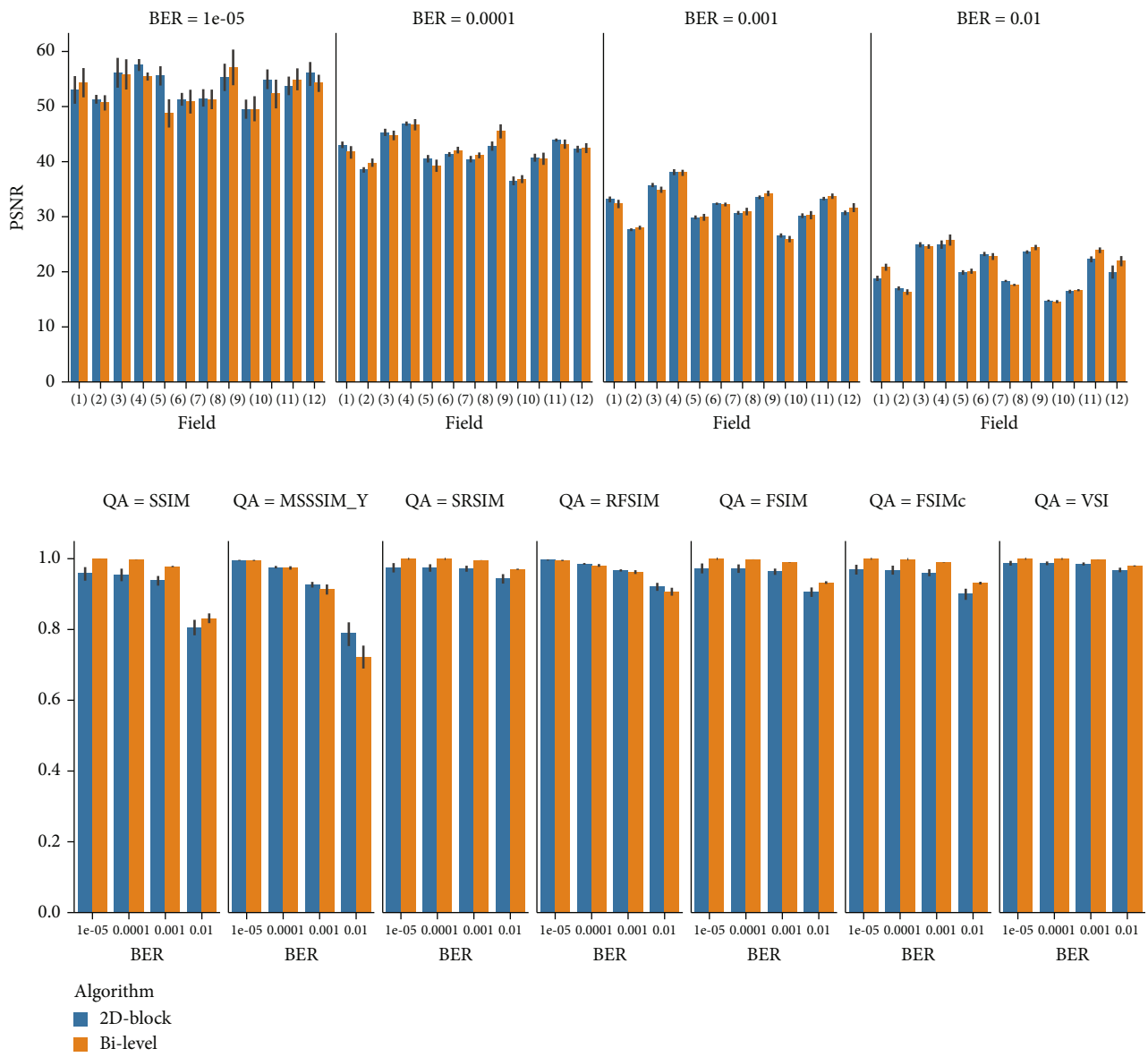


Figure 10: Comparison of recovered image quality: (a) PSNR; (b) other assessment methods.

decorrelation of color components [31]. To approximate the original color transform well enough, the lossless color transform with equations (1) and (2) in [32] is adopted in our algorithm. As [7, 19] mentioned, the prediction residues have reduced amplitudes and are assumed to be statistically independent with an approximate Laplacian distribution. Therefore, a predictor in Figure 1 is employed to further remove data correlation in Y, Cb, and Cr channels, respectively. The predictor value of X_p can be obtained with equation (3), where A_p, B_p, and C_p are the pixel value and their location is illustrated in Figure 2. After the prediction step, variable-size 2D blocks are extracted and key information about the blocks are encoded with Hamming code. Finally, these 2D blocks are separately encoding with built-in bi-level coding to make use of the sparsity property of Laplacian distribution and achieve better signal quality and robust to bit errors [7].

$$\begin{cases} Y = G + \left\lfloor \frac{86}{256}\mathrm{Cr} + \frac{29}{256}\mathrm{Cb} \right\rfloor, \\ \mathrm{Cr} = R - G, \\ \mathrm{Cb} = B - \left\lfloor \frac{87}{256}R + \frac{169}{256}G \right\rfloor, \end{cases} \tag{1}$$

$$\begin{cases} G = Y - \left\lfloor \frac{86}{256}\mathrm{Cr} + \frac{29}{256}\mathrm{Cb} \right\rfloor, \\ R = \mathrm{Cr} + G, \\ B = \mathrm{Cb} + \left\lfloor \frac{87}{256}R + \frac{169}{256}G \right\rfloor, \end{cases} \tag{2}$$

$$\mathrm{Predictor}(X_p) = \frac{3A_p + 3B_p - 2C_p}{4}. \tag{3}$$

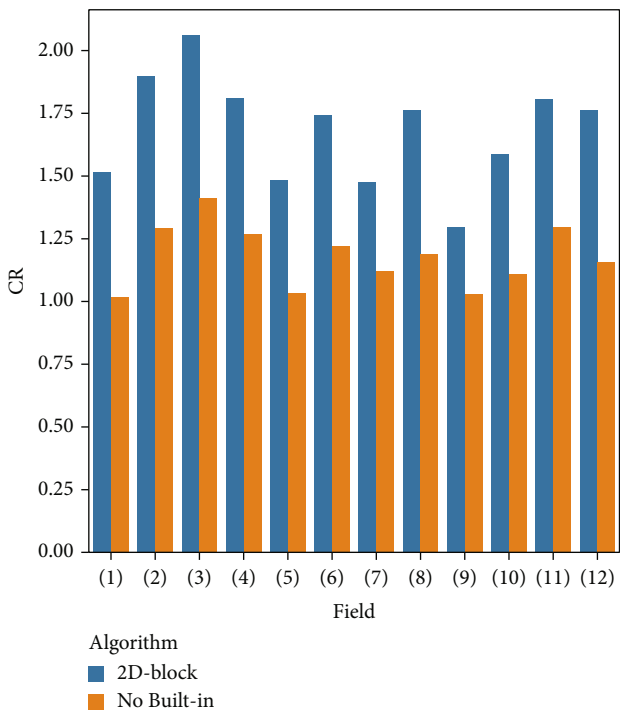


Figure 11: Effect of bi-level on CR (BER = 0.001, optimal th_n).

The procedure of 2D-block extraction and encoding is further shown in Figure 3; the 2D-layer-block extraction method is used to keep the Laplacian or Gaussian distribution of data in Layer-1~n or 2D blocks, which have the sparsity property and are suitable for bi-level coding. n in Layer-n represents n bits required to encode in the extracted blocks, and the remaining data not belonging to any blocks is left to the next layer for extraction. The 2D-block encoding method with built-in bi-level is used to improve CR and keeps robust to bit errors. The built-in bi-level procedure split the 2D block into many one-dimension signals, and each signal is encoding separately. It is because the bi-level method has the maximum encoding length, which is normally the same as the width of image.

3.1. 2D-Layer-Block Extraction Method

3.1.1. Principle of the Extraction Method. In the proposed algorithm, to keep the data distribution have the sparsity property and be suitable for bi-level coding, a novel 2D-layer-block extraction method is proposed to keep the Laplacian or Gaussian distribution of data in 2D blocks. In addition, the extraction method can rearrange the order of data encoded and the entropy of data is decreased, so CR can be improved. The principle of the method is introduced as follows.

For encoding residues, if a two-dimension block, called 2D block, can be encoded with n bits per residue, all of these datum x in the block must be satisfied with condition shown in (4). Therefore, it is feasible to find these blocks for $n = 1$ bits, then $n = 2, ..8$.

$$\max(x) - \min(x) \leq 2^n - 1. \tag{4}$$

Let us consider all these datum x in the block governed by a probability density $f(x)$, and the entropy is calculated by (6) [33].

$$f(x) = \frac{1}{2\pi\sigma} e^{-\left((x-\mu)/2\sigma^2\right)^2}, \tag{5}$$

$$H(x) = -\int_{-\infty}^{\infty} f(x) \log f(x) dx. \tag{6}$$

By inserting (5) into (6), the entropy for a Gaussian distribution is expressed as

$$H(x) = \log \sqrt{2\pi e \sigma}^2 (Gaussian). \tag{7}$$

Since the residue sequence with Gaussian distribution has maximum entropy, the following inequality holds in general.

$$H(x) \leq \log \sqrt{2\pi e \sigma}^2 (\text{in } general). \tag{8}$$

According to (4), the fixed standard deviation σ is less than $(2^n - 1)/2$ and (9) can be deducted when we assume $\mu = 0$; L is the sample size in all of these datum x in one block. By substituting (9) to (8), equation (10) can be obtained. When blocks for n bits are found starting from 1 to 8, the entropy of these data in blocks is increased later and later according to (10). So, the entropy is decided by n and it is possible to improve the compression ratio with this method.

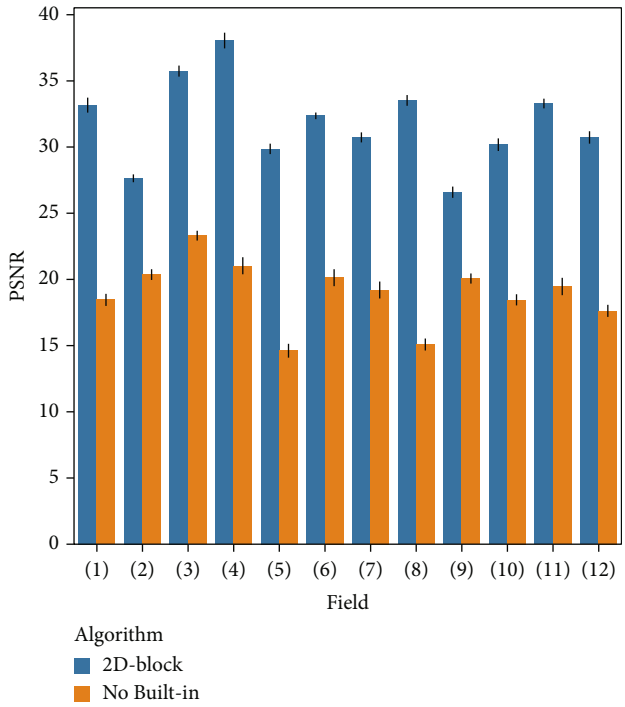
$$\sigma = \sqrt{\frac{1}{L}\sum_{i=1}^{L}(x_i - \mu)^2} \leq \sqrt{\frac{1}{L}\sum_{i=1}^{L} x_i^2} \leq \frac{(2^n - 1)}{2}, \tag{9}$$

$$H(x) \leq \log \sqrt{2\pi e} \frac{(2^n - 1)}{2}. \tag{10}$$

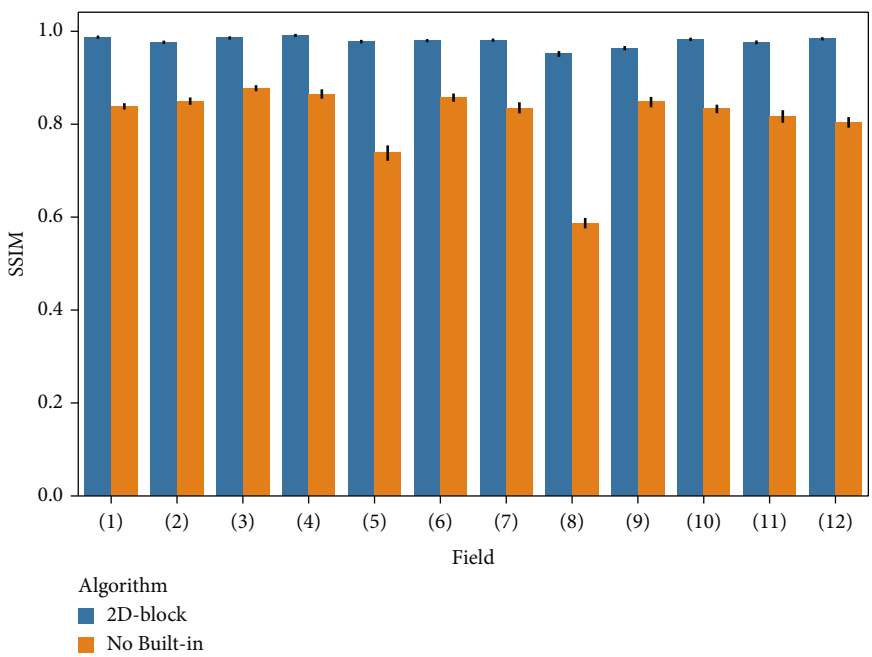
According to the discussion above, we assume $\mu = 0$. After performing prediction and making it zero-mean by removing the average, many residue values are close to zero and the residues follow a Laplacian distribution as shown in Figure 4(a). That is, all the data in one of these encoding 2D blocks will satisfy (11). Note that the sample size L in the block is above a threshold value of th_n and data in the block possess a Laplacian or Gaussian distribution approximately.

$$\text{mean}(x) \approx 0, L \geq th_n. \tag{11}$$

To proceed, if all of these 2D blocks with $n = 1$ bits are found, they will be extracted from residue data. The rest of the residue data consists of three portions. The first portion has values bigger than $(2^n - 1)/2$, and the second portion has values smaller than $-(2^n - 1)/2$, while the third portion contains data which size is smaller than th_n. It is noted that after the residues $\in[-(2^n - 1)/2, (2^n - 1)/2]$ shown in Figure 4(a) are extracted, the rest of the residue data will nearly keep the Laplacian distribution. When the extraction is

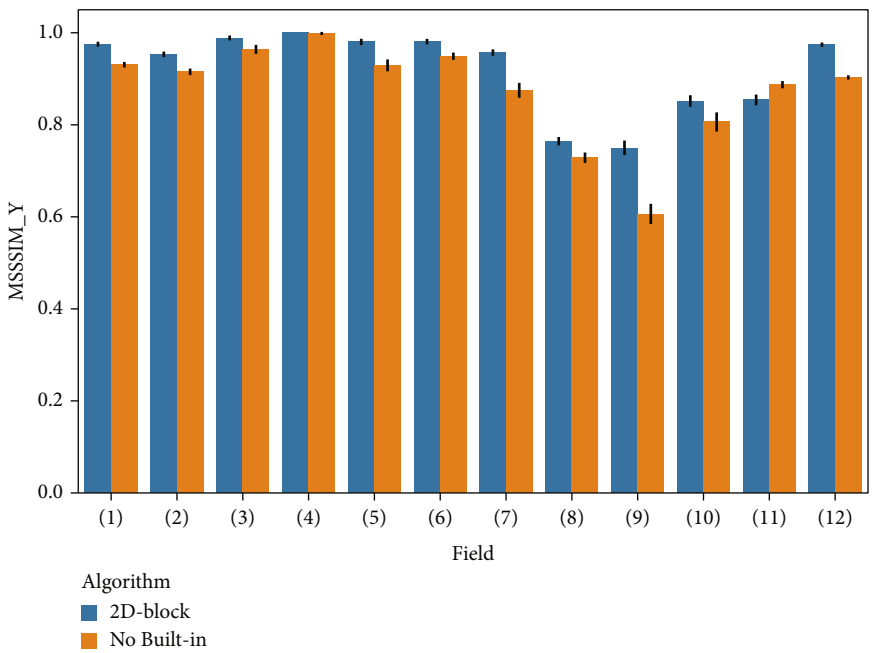


(a)



(b)

Figure 12: Continued.



(c)

Figure 12: Effect of bi-level on PSNR (BER = 0.001, optimal th_n).

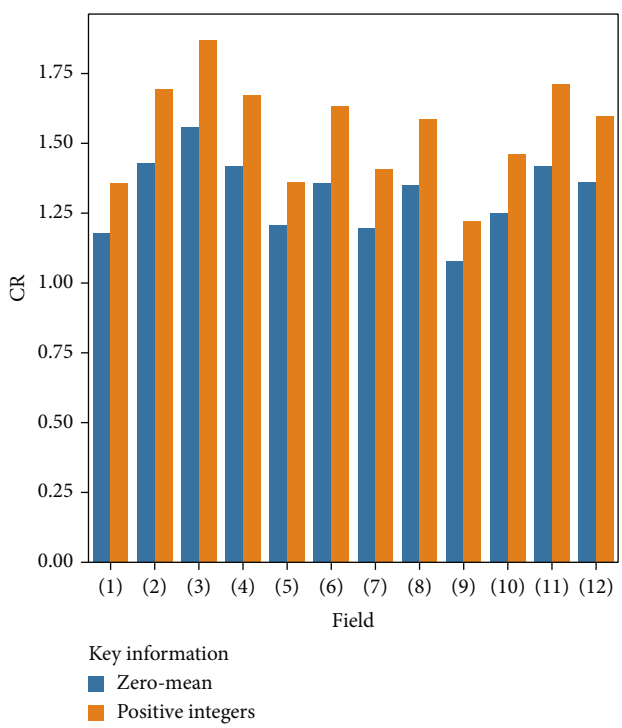


Figure 13: Effect of zero-mean on CR (BER = 0.001, $th_n = 128$, no bi-level).

repeated from $n = 1$ to 8, the Laplacian distribution of the remained residue data will change with decreasing the probability density around zeros as shown in Figure 4(b). In addition, the Laplacian or Gaussian distribution in these 2D blocks will be flattened as depicted in Figure 4(c) because of increasing value of $(2^n - 1)$. In this paper, the procedure is called as layer extraction.

3.1.2. Procedure of the Extraction Method. According to (11), 2D-layer blocks each having a sample size above the threshold of th_n are extracted repeatedly. For example, in Figure 5, image residue data is given as a matrix $40 * 40$ and many 2D blocks belonging A are extracted. The data which are not included in blocks are reshaped as a matrix M with the same height as the original residue image, while other remaining data are collected as an array of B. After the first layer is finished, matrix M is processed similarly in the next layer. With the extraction and matrix reshape operations, many edge values will be merged with other data and have less effect on compression [24]. The pseudocode of block extraction procedure is shown in Figure 6.

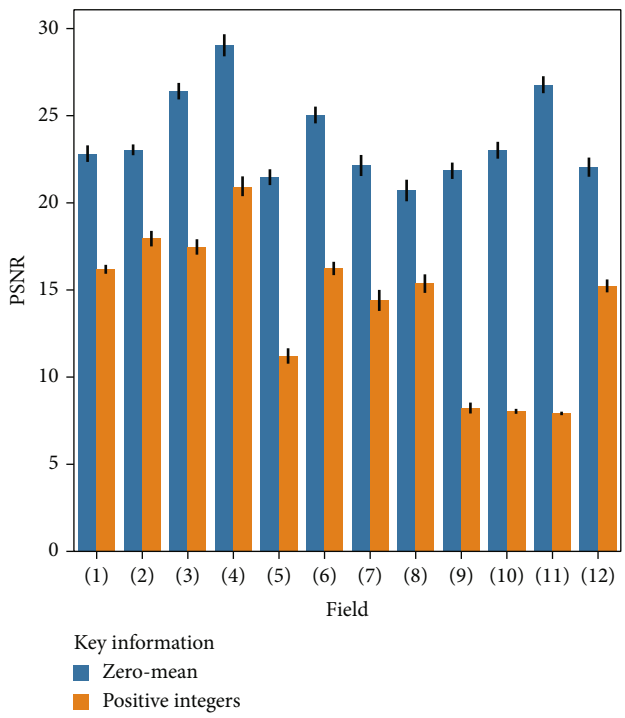
3.2. Built-In Bilevel Coding

3.2.1. Bilevel Coding. As most of these data in one of 2D blocks has a sparse distribution discussed above, a bi-level coding scheme proposed by our previous works [7, 34] in Figure 7 can be applied.

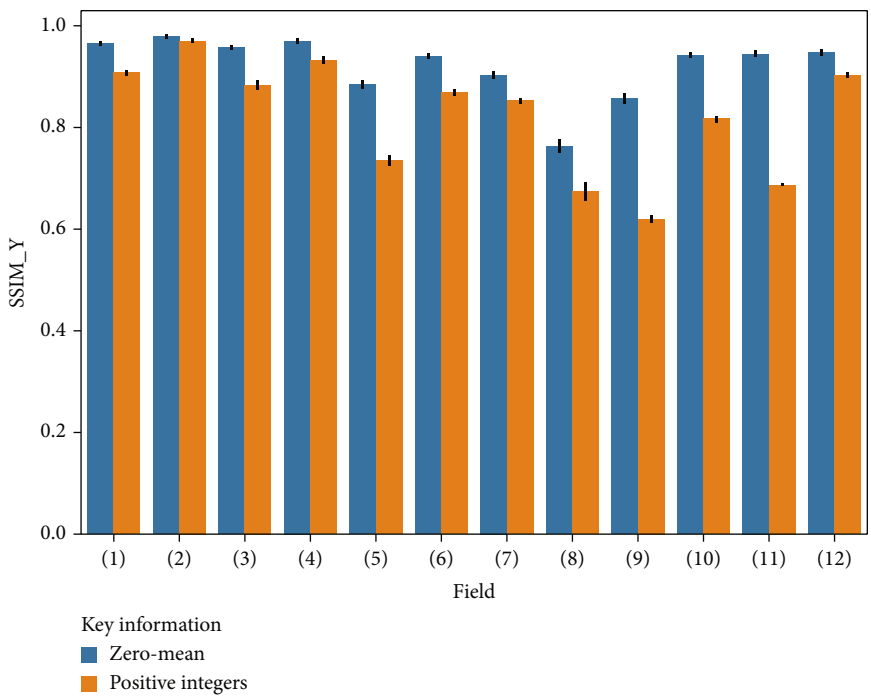
Let p_0 as the probability of a data sample requiring more than N_1 bits and less or equal to N_0 bits to encode. Assuming that $n_b p_0 \leq 0.3$ [34], the average total length is expressed in the following:

$$L_{\text{ave}} = \frac{n_s}{n_b} + n_s N_1 + (N_0 - N_1) n_s n_b p_0, \quad n_b p_0 \leq 0.3. \quad (12)$$

For a given 2D block for n bits, $N_0 = n$, the original total length is $N_0 * n_s$. When bi-level block coding is applied, the compression ratio will be improved according to (12).

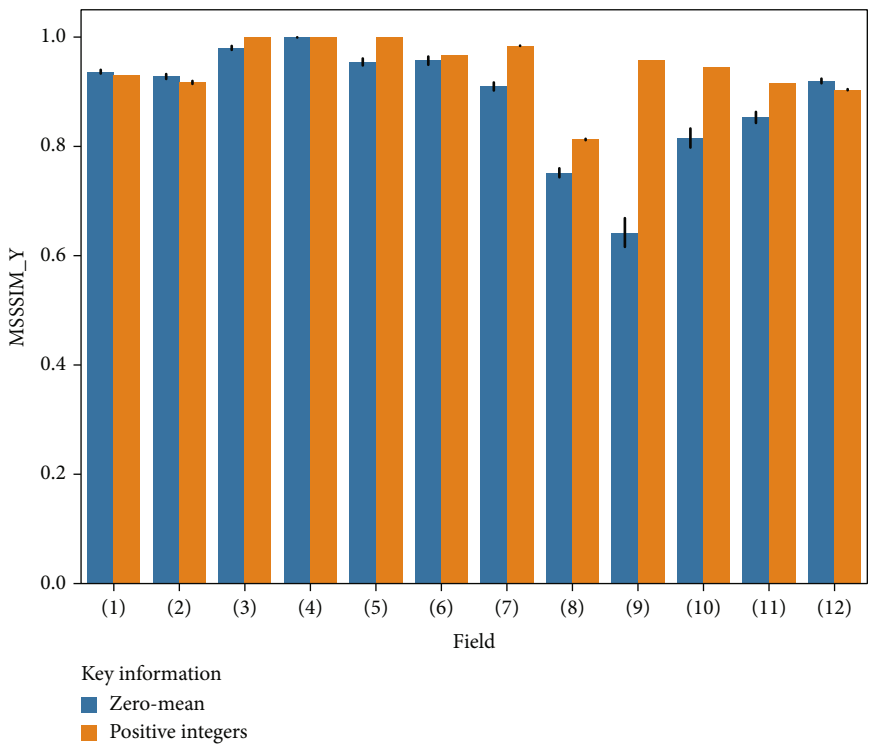


(a)



(b)

Figure 14: Continued.



(c)

Figure 14: Effect of zero-mean on PSNR (BER = 0.001, $th_n = 128$, no bi-level).

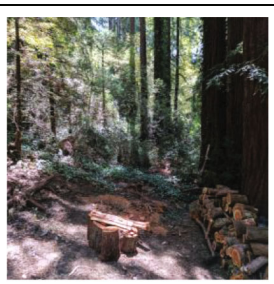
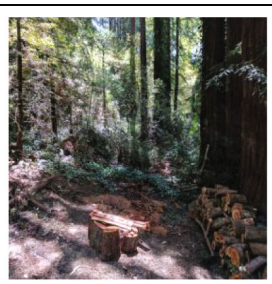
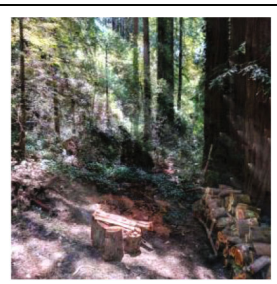
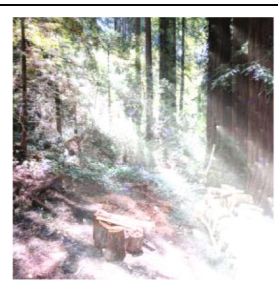
File Id	Original	Built-in Bi-level (zero mean)	No Bi-level (zero mean)	No Bi-level (positive integer)
(4)				
(9)				

Figure 15: Effects on images with different methods.

3.2.2. Optimization of 2D-Block Start Bits. According to (12), for an 8-bit gray image data, n_s is a constant, $N_0 = 8$. Given N_1, p_0, which can be estimated, optimal n_b can be determined to achieve a minimum length of L_{ave}. By taking the derivative of (12) and setting it equal to zero, the optimized block size x can be calculated by equation (13). So, the minimum N_1 satisfying $n_b p_0 \leq 0.3$ can be found through Figure 8. Finally, the "bits" value in Figure 7 can start from the minimum N_1 and the efficiency of 2D-block extraction will be improved.

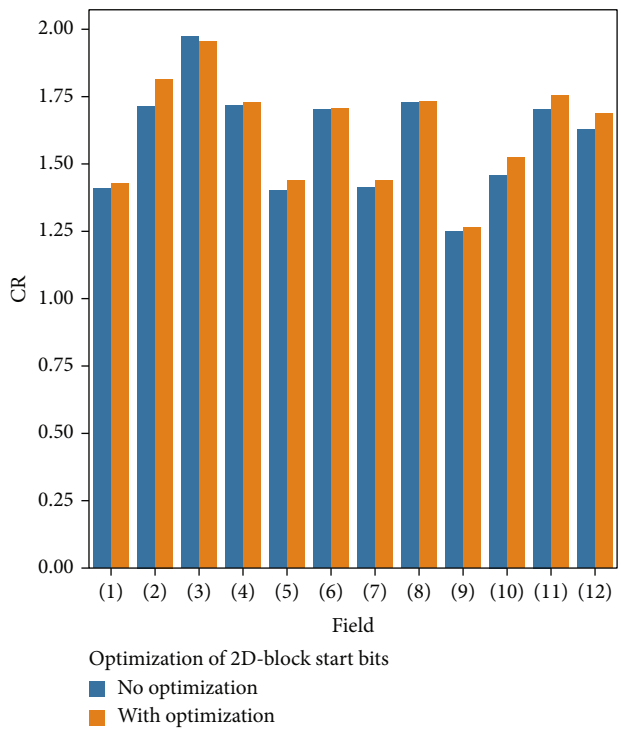


Figure 16: Effects on CR with optimization (BER = 0.001, $th_n = 128$).

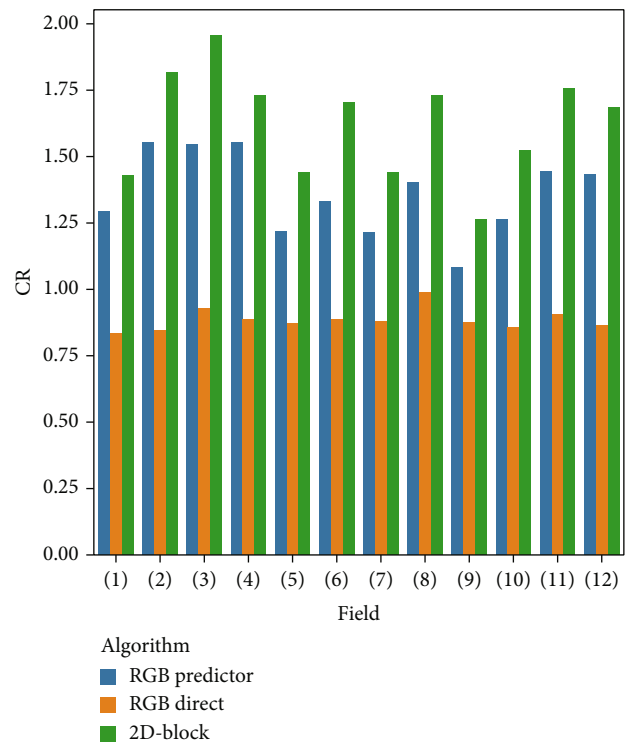


Figure 17: Comparison with different color space and predictor (BER = 0.001, $th_n = 128$).

$$x^* = \frac{1}{\sqrt{(N_0 - N_1)p_0}} = \frac{1}{\sqrt{(8 - N_1)p_0}}. \tag{13}$$

3.3. 2D-Block Encoding. Figured 9(a)–9(c) show the details of the encoding scheme. When one color image is given, three channels are separately encoded and the head information including the color space, predictor information, and their hamming coding in (a). In each channel, 2D blocks are extracted with extraction method layer by layer so encoding is implemented recursively layer by layer as well and each layer is encoded separately. In each layer of (b), head data and image data are separately encoded. (c) shows the encoding scheme of head data. Width of parent matrix $*$ the height of image is the size of M, and length of B is the length of the remaining data in Figures 5 and 6. Every block has start position (x, y), its size (w, h), mean value of the data in block, the maximum bits used of each data in block, and the key information of built-in bi-level coding including N_0, N_1, n_b, the number of block n_s/n_b, and bitstream of block type. Particularly, the mean value of data in block has two functions. One is used to improve the capability of robust to bit error because the mean value keeps the key information of one block. Another is used to ensure zero-mean of the block data, which is the feature in bi-level coding.

4. Experiments

To validate our proposed algorithm, Open Images from http://data.vision.ee.ethz.ch/mentzerf/validation_sets_lossless/val_oi_500_r.tar.gz, CLIC mobile dataset from https://data.vision.ee.ethz.ch/cvl/clic/mobile_valid_2020.zip shown in Table 1, and many classic images from http://sipi.usc.edu/database/ and http://homepages.cae.wisc.edu/~ece533/images/ are used.

In the extraction of 2D blocks, an optimal threshold of sample size th_n is given in Table 1. All these data in one layer will be split into many pieces with 512 samples, which is the same as the width of image. To evaluate the effect of 2D-block encoding, built-in bi-level coding and color space, etc., experiments with different combinations are implemented. All the results are the average values from 10 runs. A bit-error rate (BER) is default set as 0.001. The bi-level method is applied for RGB color image [7].

To evaluate the performances for the bit-error environment, the Peak Signal to Noise Ratio (PSNR in dB), Structural Similarity (SSIM) as (16) [35], Multi-Scale SSIM(MSSSIM_Y), Spectral Residual-based Similarity (SRSIM), Riesz-transform-based Feature SIMilarity (RFSIM), Feature Similarity (FSIM), FSIMc, and Visual Saliency-based Index (VSI) [36, 37] are used as error metrics to measure the recovered image quality. In (15), $I(x, y)$ represents the original pixel, while $\widehat{I}(x, y)$ is the recovered pixel. In (17) and (18), $K_1 = 0.01$, $K_2 = 0.03$, and $L = 255$.

$$\mathrm{PSNR} = 20 \times \log_{10}\left(\frac{255}{\mathrm{RMSE}}\right), \tag{14}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{M \times N}\sum_{x=1}^{M}\sum_{y=1}^{N}\left[I(x, y) - I\wedge(x, y)\right]^2}, \tag{15}$$

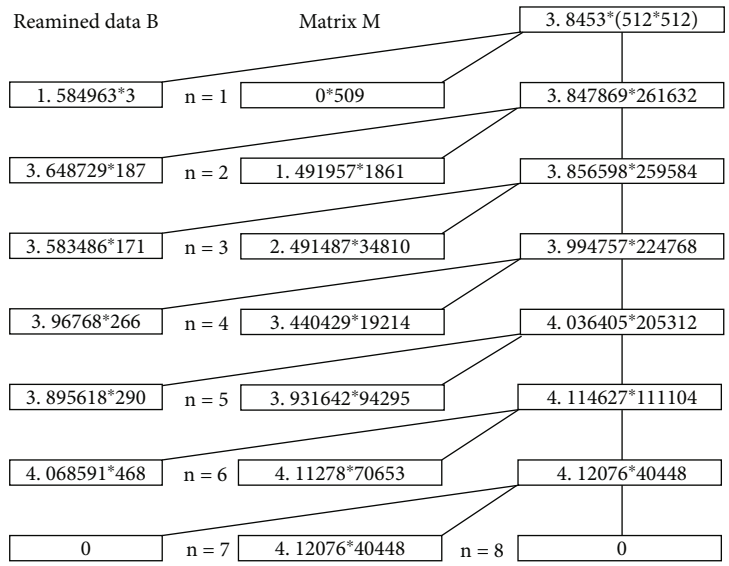


Figure 18: The entropy of 2D blocks for n bits (entropy $*$ number of samples, $\mathrm{th}_n = 128$). Left is the entropy of blocks extracted.

$$\mathrm{SSIM}(x, y) = \frac{\left(2\mu_x\mu_y + C_1\right)\left(2\sigma_{xy} + C_2\right)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\sigma_x^2 + \sigma_y^2 + C_2\right)}, \tag{16}$$

$$C_1 = (K_1 L)^2, \tag{17}$$

$$C_2 = (K_2 L)^2. \tag{18}$$

4.1. Comparison. In this experiment, our proposed method is compared with many state-of-the-art methods from Refs. [16, 17], engineered lossless compression algorithms including PNG, Jpeg2000, WebP, FLIF, and deep learning-based lossless compression algorithm L3C. As all of these methods are not suitable to be applied in bit-error situation, these methods with hamming code (7,4) are supposed a solution robust to bit-error environment. The results are given in Tables 2 and 3 and Figure 10.

In Table 2, the results are taken from Refs. [15, 16] and the best results of CR with hamming code are listed in the last second column. The average CR of our method is 1.31296 and better than 1.116592. In Table 3, the results of Jpeg2000, WebP, and FLIF are achieved through the compression tools including OpenJpeg, WebP from Google, and FLIF from Cloudinary. As a deep learning method, the result of L3C is achieved by using the neuron model trained with Open Images dataset to compress images. The average CR of our method is 1.682933 and better than others such as bi-level (1.66655), FLIF (1.429898), and L3C (1.396234). In addition, it is noticed that CR of L3C with images (5) and (9) from the CLIC dataset get the worst results, which are 1.098277 and 0.918406. One reason that leads to the result is that the neuron model used to compress is trained with Open Images dataset and L3C do not perform well on the images from a different dataset CLIC.

Figure 10 shows the image quality assessment results. PSNR, SSIM and MSSSIM_Y can better reflect the situation of bit-error channel. Therefore, only these three assessment results are discussed in the late section.

According to the comparison results in Tables 2 and 3 and Figure 10, the compression ratio of our proposed method is higher than bi-level coding although 2D-block encoding requires more header bits to encode the information about the position and size of block. And similar image quality with bi-level coding is kept in Figure 10. The reason is that the 2D-layer-block extraction method rearranges the data order to decrease the entropy and the data distribution of one-layer blocks nearly keeps as Laplacian distribution which is suitable for bi-level coding as discussed before. In addition, the analysis will be further discussed in Section 4.6–4.7.

4.2. The Effect of Built-In Bilevel Encoding. To investigate the advantage of the bi-level coding method, two experiments including "built-in bi-level" and "no bi-level" are implemented. The results are shown in Figures 11 and 12.

When the center of the Laplacian distribution is located at zero, bi-level coding can require less bits to encode. So, the built-in bi-level encoding method achieves the best compression ratio as shown in Figure 11. As bi-level coding is proposed for noisy channel [7], it gets higher PSNR, SSIM, and MSSSIM_Y and maintains a better image quality just as shown in Figure 12.

4.3. Comparison between Zero-Mean and Positive Integers. As we know, min value can also be as the key information to improve image quality just as mean value. In this experiment, positive integer values by removing min value are encoded. The results are shown in Figures 13 and 14.

As discussed in Section 4.2, the compression ratio with built-in bi-level coding is higher when zero-mean values are used. While without the bi-level coding used, the compression ratio with positive integer values by removing the min value is higher than with zero-mean values in

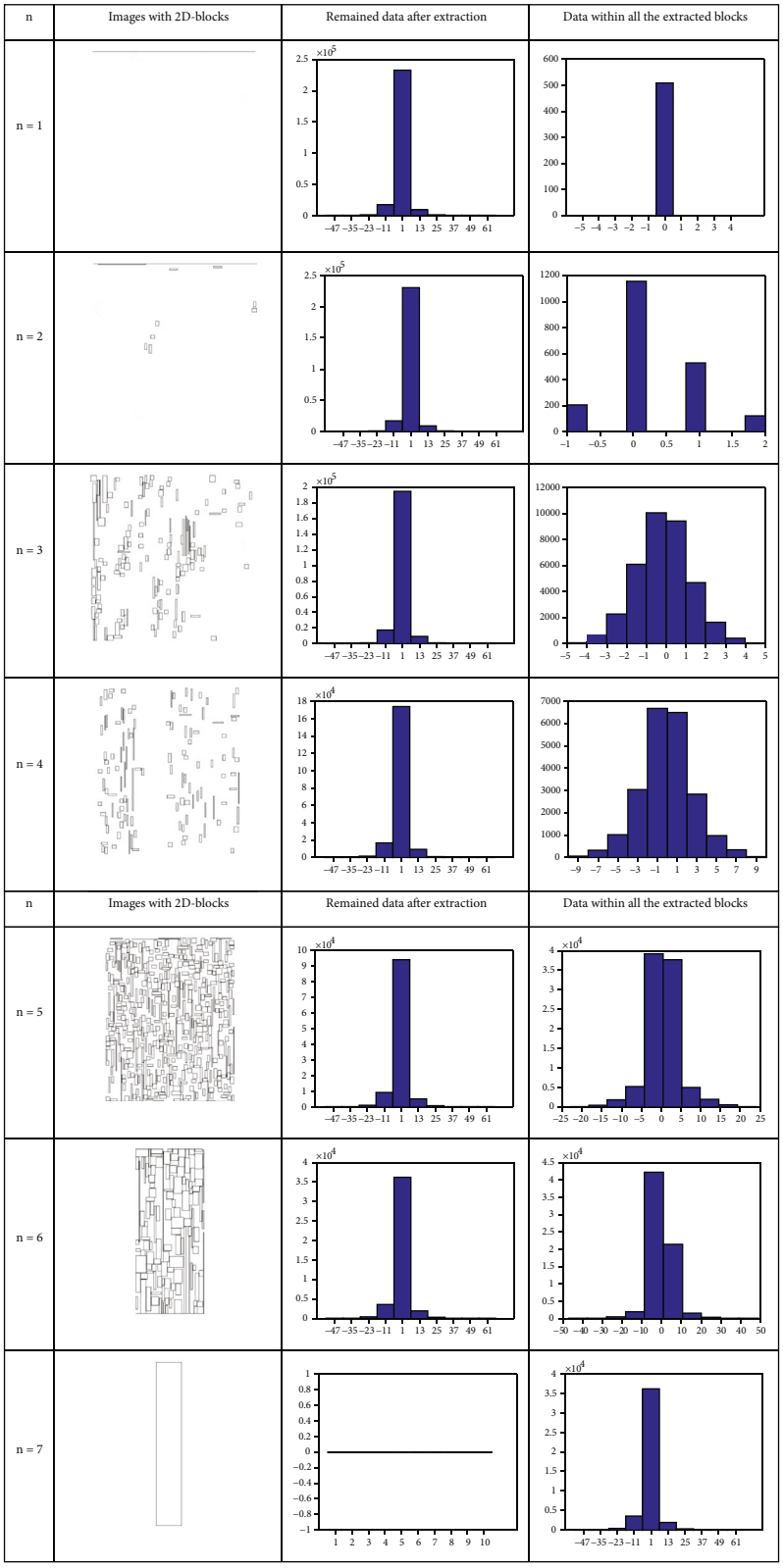


Figure 19: Change of data distribution ($Th_n = 128$).

Figure 13. The reason is that encoding with zero-mean values requires the sign bit. But the usage of sign bit can result in that encoding with zero-mean values has smaller amplitude around the mean value, so encoding with zero-mean values has less effect on PSNR and SSIM than with positive integer values. Consequently, according to the

PSNR and SSIM shown in Figures 14(a) and 14(b), encoding with zero-mean values can achieve better PSNR and SSIM than with positive integer, which is consistent with Section 3.3.

Figure 15 gives the reconstructed images in different methods. According to the results of image (9) at the last row, encoding with "positive integer" values shows worse image quality than others. But in Figure 14(c), the MSSSIM value of "positive integer" is higher than "zero-mean." Therefore, the MSSSIM result cannot be consistent with the real image quality in some cases.

4.4. Evaluation with Optimization of 2D-Block Start Bits. In the 2D-block coding method, the excessive number of blocks can lead to a decrease of compression ratio. So, the experiment based on optimization is conducted and the results are given in Figure 16. It is observed that the optimization does work and improves the compression ratio.

4.5. Evaluation with Different Color Spaces and Predictors. In Figure 17, "RGB Direct" denotes that RGB image is directly encoded without a predictor while "RGB Predictor" designates that the RGB residue with a predictor is used. It shows that color space YCrCb in "2D block" performs better than color space RGB and "RGB Direct" is the worst, which validates that the difference between pixel values and their predictions is generally smaller numbers than the pixel values themselves [13].

4.6. The Decrease of Entropy. In Figure 18, the original entropy of gray image of (4) is 3.8453 while the average entropy with the 2D-block encoding method is 3.772696. It is indicated that the compression ratio is improved and less than the value according to the information theory, which is coincident with the principle in Section 3.1.

4.7. The Change of Data Distribution in Every Layer. Figure 19 shows all the images with 2D blocks and the data distribution of remained data or all the data in blocks from gray image of (4) without the optimization of 2D-block start bits. The images indicated reshape operator has changed the distribution of edge cared about by many predictors [24]. The remained data distribution is close to the Laplacian distribution shown in the left histogram, and the data distributions in blocks are close to Gaussian or Laplacian distribution shown in the right histogram. All of these are coincident with the analysis in Section 3.1.

4.8. Discussion. Through these experiments, the results of comparison proved that our method performs better than state-of-the-art methods, engineering lossless compression algorithms and deep learning methods under bit-error situation. There are four main reasons.

First, the 2D block extraction method extracts the data encoded with smaller bits layer by layer; thus, the entropy is decreased as Section 4.6 show.

Second, the edge data always cause poor compression rate but the 2D-block extraction method has changed the edge data distribution. And the data distribution of each layer block nearly keeps as Laplacian distribution which is suitable for bi-level coding as Figure 19 of Section 4.7.

Third, built-in bi-level coding with zero-mean value can preserve high image quality under bit-error environment as Section 4.2 and 4.3 discussed.

At last, optimization of 2D-block start bits and color space used in "2D block" is an important mechanism to improve the compression rate as the discussion in Sections 4.4 and 4.5.

5. Conclusions

When image data is transferred through wireless communication systems, bit errors may occur and will cause corruption of image data. To reducing the bit-error effect, a bit-aware lossless image compression algorithm based on bi-level coding can be applied. But bi-level coding is one of the one-dimension coding methods and has not considered the inherent statistical correlation in 2D context region. So, to resolve this shortage, a novel 2D-layer-block extraction and encoding method with built-in bi-level coding is proposed to improve the compression ratio. With the layer extraction method, the data distribution is close to the Laplacian distribution after each layer extraction, which is suitable for bi-level coding. For color image, a lossless color transformation from RGB to the YCrCb color space is used for the decorrelation of color components. Through experiments, it is demonstrated that our proposed method obtains the better lossless compression ratio and keeps the same image quality with the bi-level method under noise transmission channel. Although it is not as efficient when compared to state-of-the-art methods in terms of lossless compression ratio sometimes, it is more robust to bit errors caused by noisy channel. Furthermore, after applying the feed-forward error control scheme, different predictor, and coding method, we can achieve better compression efficiency, since the bi-level block coder requires a smaller number of bits by the bit-error protection algorithm than the amount required by the entropy coder. Also, it is noted that deep learning methods are trained with Open Images dataset but perform poor on the images from a different dataset. Therefore, the generalization ability of deep learning methods is required to be improved in the future.

Data Availability

The [CLIC mobile dataset], [Open Images] and other classic images used to support the findings of this study are included within the article.

Ethical Approval

If the images or other third-party material in this article are not included in the article's Creative Commons license, it is required to obtain permission directly from the copyright holder.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (No. 61502423), Zhejiang Provincial Natural Science Foundation (No. LY16G020012), and Zhejiang Province Public Welfare Technology Application Research Project (Nos. LGF19F010002, LGN20F010001, LGF20F010004, and LGG21F030014).

References

[1] J. Ballé, D. Minnen, S. Singh, S. J. Hwang, and N. Johnston, "Variational image compression with a scale hyperprior," 2018, arXiv preprint arXiv:1802.01436.

[2] J. Lee, S. Cho, and M. Kim, *An End-to-End Joint Learning Scheme of Image Compression and Quality Enhancement with Improved Entropy Minimization*, 2019, https://arxiv.org/pdf/1912.12817.

[3] F. Jiang, W. Tao, S. Liu, J. Ren, X. Guo, and D. Zhao, "An end-to-end compression framework based on convolutional neural networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 10, pp. 3007–3018, 2018.

[4] Z. Cheng, H. Sun, M. Takeuchi, and J. Katto, "Learned image compression with discretized Gaussian mixture likelihoods and attention modules," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, June 2020.

[5] K. N. Shruthi, B. M. Shashank, Y. S. Saketh et al., "Comparison analysis of a biomedical image for compression using various transform coding techniques," in *2016 IEEE 6th International Conference on Advanced Computing (IACC)*, pp. 297–303, Bhimavaram, India, February 2016.

[6] A. Avramović and G. Banjac, "On predictive-based lossless compression of images with higher bit depths," *Telfor Journal*, vol. 4, no. 2, pp. 122–127, 2012.

[7] L. Tan and L. Wang, "Bit-error aware lossless image compression," *International Journal of Modern Engineering*, vol. 11, no. 2, pp. 54–59, 2011.

[8] "WebP image format," https://developers.google.com/speed/webp/.

[9] J. Sneyers and P. Wuille, "FLIF: free lossless image format based on MANIAC compression," in *2016 IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, USA, September 2016.

[10] J. Townsend, T. Bird, and D. Barber, "Practical lossless compression with latent variables using bits back coding," The International Conference on Learning Representations, Louisiana, USA, 2019, arXiv preprint arXiv:1901.04866.

[11] F. Kingma, P. Abbeel, and J. Ho, "Bit-swap: recursive bits-back coding for lossless compression with hierarchical latent variables," in *In International Conference on Machine Learning*, pp. 3408–3417, California, USA, 2019.

[12] F. Mentzer, E. Agustsson, M. Tschannen, R. Timofte, and L. Van Gool, "Practical full resolution learned lossless image compression," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, June 2019.

[13] R. Jovanovic and R. A. Lorentz, "Adaptive lossless prediction based image compression," *Applied Mathematics & Information Sciences*, vol. 8, no. 1, pp. 153–160, 2014.

[14] I. Matsuda, N. Ozaki, Y. Umezu, and S. Itoh, "Lossless coding using variable block-size adaptive prediction optimized for each image," *EUSIPCO*, pp. 818–821, 2005.

[15] X. Zhao and Z. He, "Lossless image compression using super-spatial structure prediction," *IEEE Signal Processing Letters*, vol. 17, no. 4, pp. 383–386, 2010.

[16] T. Strutz, "Context-based predictor blending for lossless color image compression," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 4, pp. 687–695, 2016.

[17] S. Kim and N. Cho, "Hierarchical prediction and context adaptive coding for lossless color image compression," *IEEE Transactions on Image Processing*, vol. 23, no. 1, pp. 445–449, 2014.

[18] H. S. Malvar, G. J. Sullivan, and S. Srinivasan, "Lifting-based reversible color transformations for image compression," *Proceedings of SPIE-The International Society for Optical Engineering*, 2008.

[19] R. Starrosolski, "Simple fast and adaptive lossless compression algorithm," *Softwware: Practice and Experience*, vol. 37, no. 1, pp. 65–91, 2007.

[20] M. Wienberger, G. Seroussi, and G. Sapiro, "LOCO-I: a low complexity, context-based, lossless image compression algorithm," in *Proceedings of the Conference on Data Compression*, pp. 140–148, Snowbird, UT, USA, 31 March-3 April 1996.

[21] S. D. Rane and G. Sapiro, "Evaluation of JPEG-LS, the new lossless and near-lossless still image compression standard for compression of high-resolution elevation data," *IEEE Transactions of Geosciences and Remote Sensing*, vol. 39, no. 10, pp. 2298–2306, 2001.

[22] X. Wu and N. Memon, "CALIC–a context based adaptive lossless image codec," in *1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings*, pp. 1890–1893, Atlanta, GA, USA, May 1996.

[23] A. Avramović and B. Reljin, "Gradient edge detection predictor for image lossless compression," in *In Proceedings ELMAR-2010*, pp. 131–134, Zadar, Croatia, September 2010.

[24] L. Xin and M. T. Orchard, "Edge-directed prediction for lossless compression of natural images," *IEEE Transactions on Image Processing*, vol. 10, no. 6, pp. 813–817, 2001.

[25] L. Kau and Y. Lin, "Adaptive lossless image coding using least-squares optimization with edge-look-ahead," *EEE Transactions on Circuits and Systems II: Express Briefs*, vol. 52, no. 11, pp. 751–755, 2005.

[26] L. Kau and Y. Lin, "Least-squares-based switching structure for lossless image coding," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 54, no. 7, pp. 1529–1541, 2007.

[27] F. Hsieh and K. Fan, "A high performance lossless image coder," in *in Conference on Computer Vision & Graphic Image Processing*, Taipei, Taiwan, 2005.

[28] G. Ulacha and R. Stasinski, "Performance optimized predictor blending technique for lossless image coding," in *In 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1541–1544, Prague, Czech Republic, 2011.

[29] W. Dai, H. Xiong, J. Wang, and Y. Zheng, "Large discriminative structured set prediction modeling with max-margin Markov network for lossless image coding," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 541–554, 2014.

[30] Y. Zhang and D. A. Adjeroh, "Prediction by partial approximate matching for lossless image compression," *IEEE Transactions on Image Processing*, vol. 17, no. 6, pp. 924–935, 2008.

[31] M. Domanski and K. Rakowski, "Color transformations for lossless image compression," in *IEEE Signal Processing Conference*, pp. 1–4, Tampere, Finland, 2000.

[32] S. C. Pei and J. J. Ding, "Improved reversible integer-to-integer color transforms," in *IEEE International Conference on Image Processing*, pp. 473–476, Cairo, Egypt, 2009.

[33] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, "Entropic uncertainty relations and their applications," *Reviews of Modern Physics*, vol. 89, no. 1, pp. 1–58, 2017.

[34] L. Tan, J. Jiang, and Y. Zhang, "Bit-error aware lossless compression of waveform data," *IEEE Signal Processing Letters*, vol. 17, no. 6, pp. 547–550, 2010.

[35] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error measurement to structural similarity," *IEEE Trans. on Image Processing*, vol. 13, no. 4, pp. 1–14, 2004.

[36] H. Amirpour, A. M. G. Pinheiro, M. Pereira, and M. Ghanbari, "Reliability of the most common objective metrics for light field quality assessment," in *ICASSP 2019-2019 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pp. 2402–2406, Brighton, United Kingdom, 2019.

[37] https://sse.tongji.edu.cn/linzhang/IQA/IQA.htm.

Hindawi
Journal of Sensors
Volume 2021, Article ID 7961306, 5 pages
https://doi.org/10.1155/2021/7961306

Research Article

Radar Circular Data Analysis Using a New Watson's Goodness of Test under Complexity

Muhammad Aslam [1] **and Muhammad Saleem** [2]

[1] *Department of Statistics, Faculty of Science, King Abdulaziz University, Jeddah 21551, Saudi Arabia*
[2] *Department of Industrial Engineering, Faculty of Engineering-Rabigh, King Abdulaziz University, Jeddah 21911, Saudi Arabia*

Correspondence should be addressed to Muhammad Aslam; aslam_ravian@hotmail.com

Received 11 July 2021; Accepted 3 September 2021; Published 15 September 2021

Academic Editor: Sang-Hoon Hong

Watson's test is applied to test either the given angular data follows the specified distribution or not. The existing Watson's test under classical statistics is applied when there is no uncertainty and indeterminacy in sample size or angular data. Under indeterminacy, the existing Watson's test cannot be applied for testing purposes. Neutrosophic statistics is an alternative to classical statistics for this kind of situation. The Watson's test under neutrosophic statistics is proposed in this paper. The test statistic of Watson's test is introduced first. The operational procedure of the proposed Watson's test is discussed with the help of radar data. From the data analysis and simulation study, it is concluded the proposed Watson's test is efficient than the existing Watson's test.

1. Introduction

The data obtained from devices or systems follows any statistical distribution. For efficient prediction and estimation, the decision-makers are interested to investigate which statistical distribution is suitable for the data. [1] suggested, "when we assume that our data follows a specific distribution, we take a serious risk. If our assumption is wrong, then the results obtained may invalid." Therefore, the goodness of fit tests is performed for testing either the data follows the specified theoretical distribution or not. These tests are performed using cumulative distribution function or probability density function of the theoretical distribution. [2–10] worked on the goodness of fit tests with applications using various types of data sets.

Usually, the statistical tests are applied to the data obtained from the linear scale. In practice, the data obtained from metrology, space, radars, and ecology are circular observations that are measured in radian or degree. [11] mentioned, "Circular data need special treatment in data analysis: consider that an angle of 355° is much nearer to an angle of 5° than it is to an angle of 330°, and so simple arithmetic mean for example can be quite misleading." The circular tests are applied for testing the randomness of circular data. [11–22] contributed in designing statistical tests for circular data.

The presence of uncertainty and indeterminacy in the sample size, parameters, and the observations lead to apply the fuzzy-based tests for testing purpose. The fuzzy-based tests are quite flexible and widely applied under the uncertain environment in a variety of fields. [23, 24–29] presented various tests using the fuzzy approach.

[30] introduced neutrosophic logic as a generalization of fuzzy logic. The efficiency of neutrosophic over fuzzy logic and interval-based analysis can be read in [31]. The applications of neutrosophic can be seen [32–36]. The classical statistics cannot be applied when vague, uncertain, and indeterminate observations are presented in the data. To overcome the issue, [37] used neutrosophic logic to introduce neutrosophic statistics. More information of neutrosophic statistics can be seen on the websites https://archive.org/details/neutrosophic-statistics?tab=about and https://archive.org/details/neutrosophic-statistics?tab=collection. The neutrosophic statistics is more informative than the

classical statistics. [38–44] worked on neutrosophic statistics with applications. Recently, [45] introduced a neutrosophic test for circular data.

The existing Watson's test under classical statistics cannot be applied when uncertainty is presented. By exploring the literature and best of our knowledge, no work on Watson's test under neutrosophic statistics is done so for. In this paper, we will introduce Watson's test under neutrosophic statistics. The proposed Watson's test is mainly aimed at testing either the given circular data follows the given theoretical distribution. The application of the proposed Watson's test will be given with the help of data measured from radars. It is expected that the proposed Watson's test will be efficient, flexible, and informative than the existing Watson's test.

2. Neutrosophic Watson's U^2_{Nn} Test

The existing Watson's U^2_n test under classical statistics is applied to test either the given random sample of angular values is fitted to the specified distribution. The existing Watson's U^2_n test is applied when all circular observations in the data are determined, exact, and uncertain. In the case when the circular observations in the data are uncertain, imprecise, and are in intervals, the existing Watson's U^2_n test under classical statistics cannot be applied for fitting the given distribution. In this section, neutrosophic Watson's U^2_{Nn} test will be introduced to fit neutrosophic angular values to the given specified distribution. The methodology of the proposed Watson's U^2_{Nn} test by following [14] is discussed as follows.

Suppose that $\Phi_{1_N}, \Phi_{2_N}, \Phi_{3_N}, \cdots, \Phi_{n_N}$ be a random sample of neutrosophic angular values of size n_N. The neutrosophic forms of angular values and neutrosophic sample size are given as follows: $\Phi_N = \Phi_L + \Phi_U I_{\mathrm{N}\Phi}\,; I_{\mathrm{N}\Phi}\epsilon[I_{\mathrm{L}\Phi}, I_{\mathrm{U}\Phi}]$ and $n_N = n_L + n_U I_{n_N}\,; I_{n_N}\epsilon[I_{n_L}, I_{n_U}]$, respectively. Note that Φ_L, n_L and $\Phi_U I_{\mathrm{N}\Phi}$, $n_U I_{n_N}$ denote the determined and indeterminate parts, respectively, and $I_{\mathrm{N}\Phi}\epsilon[I_{\mathrm{L}\Phi}, I_{\mathrm{U}\Phi}]$ and $I_{n_N}\epsilon[I_{n_L}, I_{n_U}]$ are the associated measure of uncertainty. To implement the proposed test, the first step is to arrange neutrosophic angular data $\Phi_{1_N}, \Phi_{2_N}, \Phi_{3_N}, \cdots, \Phi_{n_N}$ in ascending order $\Phi_{1_N} \le \Phi_{2_N} \le \Phi_{3_N} \le \cdots \le \Phi_{n_N}$ with respect to midvalues of each interval of $\Phi_N \epsilon[\Phi_L, \Phi_U]$. Let $F_N(\Phi_N)$ be a neutrosophic cumulative distribution function (ncdf) of the given theoretical distribution.

Let

$$V_{\mathrm{iN}} = F_N(\Phi_N), i = 1, 2, 3, \cdots, n_N. \tag{1}$$

The neutrosophic average of $\bar{V}_N \epsilon[\bar{V}_L, \bar{V}_U]$ is computed as follows:

$$\bar{V}_N = \frac{\sum_{i=1}^{n_N} V_{\mathrm{iN}}}{n_N}\,; n_N \epsilon[n_L, n_U], \bar{V}_N \epsilon\left[\bar{V}_L, \bar{V}_U\right]. \tag{2}$$

The test statistic of the proposed Watson's $U^2_{Nn}\epsilon[U^2_{Ln}, U^2_{Un}]$ is given by

$$U^2_{Nn} = \sum_{i=1}^{n_N} V^2_{\mathrm{iN}} - \sum_{i=1}^{n_N}\left(\frac{C_{\mathrm{iN}} V_{\mathrm{iN}}}{n_N}\right) + n_N\left[\frac{1}{[3,3]} - \left(\bar{V}_N - \frac{1}{[2,2]}\right)^2\right]; n_N \epsilon[n_L, n_U], \bar{V}_N \epsilon\left[\bar{V}_L, \bar{V}_U\right], \tag{3}$$

where $C_{\mathrm{iN}} = 2i - 1$.

The statistic $U^2_{Nn}\epsilon[U^2_{Ln}, U^2_{Un}]$ in neutrosophic form is given by

$$U^2_{Nn} = U^2_{Ln} + U^2_{Un} I_{U^2_{Nn}}\,; I_{U^2_{Nn}} \epsilon\left[I_{U^2_{Ln}}, I_{U^2_{Un}}\right]. \tag{4}$$

In the given neutrosophic form, the statistic U^2_{Ln} presents the test statistic under classical statistics. The value of statistic $U^2_{Un} I_{U^2_{Nn}}$ shows the indeterminate value under uncertainty and $I_{U^2_{Nn}} \epsilon[I_{U^2_{Ln}}, I_{U^2_{Un}}]$ is a measure of uncertainty associated with $U^2_{Nn}\epsilon[U^2_{Ln}, U^2_{Un}]$. The proposed statistic reduces to U^2_{Ln} when $I_{U^2_{Ln}} = 0$.

3. Application

The application of the proposed Watson's U^2_{Nn} test is given with the aid of angle data obtained from the radar. The decision-makers are interested to investigate either the radar data obtained from the radar systems follow the given theoretical distribution or not. For testing the null hypothesis, H_0: radar data follows the given theoretical distribution vs. H_1: the radar data does not follow the given theoretical distribution. For testing this hypothesis, the decision-maker is uncertain about the sample size with the measure of uncertainty I_U=0.13 with n_L=13. The neutrosophic form of sample size is $n_N = 13 + 15 I_{n_N}\,; I_{n_N}\epsilon[0,0.13]$. The following radar data is obtained from [45].

$$n_L = 13,$$

$$\Phi_1 = 2500, \Phi_2 = 275\circ, \Phi_3 = 285\circ, \Phi_4 = 2850, \Phi_5 = 290\circ, \Phi_6 = 290\circ,$$

$$\Phi_7 = 295\circ, \Phi_8 = 300\circ, \Phi_9 = 305\circ, \Phi_{10} = 310\circ, \Phi_{11} = 315\circ, \Phi_{12} = 320\circ, \Phi_{13} = 330\circ,$$

$$n_L = 15,$$

$$\Phi_1 = 2500, \Phi_2 = 275\circ, \Phi_3 = 285\circ, \Phi_4 = 2850, \Phi_5 = 290\circ, \Phi_6 = 290\circ,$$

$$\Phi_7 = 295\circ, \Phi_8 = 300\circ, \Phi_9 = 305\circ, \Phi_{10} = 310\circ, \Phi_{11} = 315\circ, \Phi_{12} = 320\circ, \Phi_{13} = 330\circ,$$

$$\Phi_{14} = 330\circ, \Phi_{15} = 5\circ. \tag{5}$$

Table 1: Effect of measure of indeterminacy on n_N and U^2_{Nn}.

$I_{n_L} = I_{U^2_{Ln}}$	n_N	U^2_{Nn}
0	[13,13]	[0.7221,0.7221]
0.001	[13,13]	[0.7221,0.7228]
0.005	[13,13]	[0.7221,0.7257]
0.010	[13,13]	[0.7221,0.7294]
0.10	[13,15]	[0.7221,0.7950]
0.20	[13,16]	[0.7221,0.8678]
0.30	[13,18]	[0.7221,0.9407]
0.40	[13,19]	[0.7221,1.0136]
0.50	[13,21]	[0.7221,1.0865]
0.60	[13,22]	[0.7221,1.1593]
0.70	[13,24]	[0.7221,1.2322]
0.80	[13,25]	[0.7221,1.3051]
0.90	[13,27]	[0.7221,1.3779]
1.00	[13,28]	[0.7221,1.4508]

The proposed test for the given data is implemented as follows:

$$U^2_{Nn} = \sum_{i=1}^{n_N} V^2_{\mathrm{iN}} - \sum_{i=1}^{n_N}\left(\frac{C_{\mathrm{iN}} V_{\mathrm{iN}}}{n_N}\right) + n_N\left[\frac{1}{[3,3]} - \left(\bar{V}_N - \frac{1}{[2,2]}\right)^2\right] = [0.7221, 0.7287]. \tag{6}$$

Step 1. Arrange the angle data in ascending order and assign number i.

Step 2. Compute the values of $V_i = \Phi^0_N/360$ and $\overline{V}_N \epsilon [0.8226, 0.775]$.

Step 3. Generate the values of V^2_{iN} and C_{iN}. The values of $\sum_{i=1}^{n_N} V^2_{\mathrm{iN}} = [8.83, 9.67]$.

Step 4. Generate the values of $\sum_{i=1}^{n_N}(C_{\mathrm{iN}} V_{\mathrm{iN}}/n_N) = [11.09, 12.81]$.

Step 5. Finally, compute the values of the proposed test statistic as follows.

Let $\alpha = 0.05$, and the tabulated value from [14] is 0.184. By comparing the values of statistic U^2_{Nn} with the critical value, the null hypothesis H_0: radar data follows the given theoretical distribution is rejected. Based on the study, it is concluded that angle data obtained from the radar does not follow the given theoretical distribution.

4. Comparative Study

In this section, the efficiency of the proposed Watson's U^2_{Nn} test will be compared with the existing Watson's U^2_n test in terms of flexibility, the measure of uncertainty, and information. As mentioned before, the existing Watson's U^2_n test is a special case of the proposed Watson's U^2_{Nn} test. The proposed Watson's U^2_{Nn} test becomes the existing Watson's U^2_n test when $I_{U^2_{Ln}}$=0. The neutrosophic form of the statistic $U^2_{Nn} \epsilon [U^2_{Ln}, U^2_{Un}]$ for the given data is as follows: $U^2_{Nn} = 0.7221 + 0.7287 I_{U^2_{Nn}}; I_{U^2_{Nn}} \epsilon [0, 0.0091]$. From the neutrosophic form, it can be seen that the statistic $U^2_{Nn} \epsilon [U^2_{Ln}, U^2_{Un}]$ adopts the value in an indeterminate interval. Under uncertainty, the proposed statistic $U^2_{Nn} \epsilon [U^2_{Ln}, U^2_{Un}]$ takes the value from 0.7221 to 0.7287. On the other hand, the existing Watson's U^2_n test takes only a single value. Therefore, under indeterminacy, the proposed Watson's U^2_{Nn} test is flexible than the existing Watson's U^2_n test. In addition, the proposed Watson's U^2_n test gives information about the measure of uncertainty that the existing test cannot provide. For the radar data, the measure of uncertainty associated with Watson's U^2_{Nn} test is 0.0091. For testing H_0: radar data follows the given theoretical distribution vs. H_1: the radar data does not follow the given theoretical distribution when $\alpha = 0.05$, the proposed Watson's U^2_{Nn} test indicates that the probability of accepting H_0 is 0.95, the probability of committing type-1 error is 0.05, and the probability of in-decision is 0.0091. From the study, it can be seen that the proposed Watson's U^2_{Nn} test is informative than the existing Watson's U^2_n test.

5. Simulation Study

A simulation study is performed to see the effect of the indeterminacy parameters on the sample and the proposed Watson's U^2_{Nn} test. The various values of the indeterminacy parameters are considered to see the behavior of sample size and test statistic U^2_{Nn}. The values of n_N and U^2_{Nn} for various values of $I_{n_L} = I_{U^2_{Ln}}$ are shown in Table 1. From Table 1, it can be seen that when $I_{n_L} = I_{U^2_{Ln}} = 0$, the values of $n_N \epsilon [13, 13]$ and $U^2_{Nn} = [0.7221, 0.7221]$. It can be seen that the values of n_N do not change up to I_{n_L}=0.10. The indeterminate parts of n_N and U^2_{Nn} increase as the values of $I_{n_L} = I_{U^2_{Ln}}$ increase 0 to 1.00. From this simulation study, it can be observed that the measure of indeterminacy plays a significant role in determined n_N and U^2_{Nn}; therefore, the sample size should be selected keeping in mind the measure of indeterminacy.

6. Concluding Remarks

The existing Watson's test under classical statistics was applied when there is no uncertainty and indeterminacy in sample size or angular data. An extension of the existing Watson's test was presented in the paper. The proposed test can be applied when uncertainty is presented in circular data. From the radar data analysis and simulation study, it is concluded that the proposed outperforms the existing Watson's test. The proposed test has some limitations that it can be very practical if the software is available to perform it. The test can be applied to test neutrosophic random data. The proposed test for big angular data can be considered as future research.

Data Availability

The data is given in the paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This paper was funded by the Deanship of Scientific Research (DSR) at King Abdulaziz University, Saudi Arabia; therefore, the authors are thankful to the DSR for their financial and technical support. The authors are deeply thankful to the editor and reviewers for their valuable suggestions to improve the quality and presentation of the paper.

References

[1] J. L. Romeu, *Anderson-Darling: A Goodness of Fit Test for Small Samples Assumptions*, RAC START, 2003.

[2] T. W. Anderson, "Anderson-Darling tests of goodness-of-fit," *International Encyclopedia of Statistical Science*, vol. 1, pp. 52–54, 2011.

[3] M. Arshad, M. Rasool, and M. Ahmad, "Anderson Darling and modified Anderson Darling tests for generalized Pareto distribution," *Pakistan Journal of Applied Sciences*, vol. 3, no. 2, pp. 85–88, 2003.

[4] M. Formenti, L. Spadafora, M. Terraneo, and F. Ramponi, "The efficiency of the Anderson–Darling test with a limited sample size: an application to backtesting counterparty credit risk internal models," *Journal of Risk*, vol. 21, no. 6, 2019.

[5] T. U. Islam, "Ranking of normality tests: an appraisal through skewed alternative space," *Symmetry*, vol. 11, no. 7, p. 872, 2019.

[6] Y. Li, Y. Wei, B. Li, and G. Alterovitz, "Modified Anderson-Darling test-based target detector in non-homogenous environments," *Sensors*, vol. 14, no. 9, pp. 16046–16061, 2014.

[7] G. Marsaglia and J. Marsaglia, "Evaluating the Anderson-Darling distribution," *Journal of Statistical Software*, vol. 9, no. 2, pp. 1–5, 2004.

[8] M. Rahman, L. M. Pearson, and H. C. Heien, "A modified Anderson-Darling test for uniformity," *Bulletin of the Malaysian Mathematical Sciences Society*, vol. 29, no. 1, 2006.

[9] N. M. Razali and Y. B. Wah, "Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests," *Journal of statistical modeling and analytics*, vol. 2, no. 1, pp. 21–33, 2011.

[10] D. K. Wijekularathna, A. B. Manage, and S. M. Scariano, "Power analysis of several normality tests: a Monte Carlo simulation study," *Communications in Statistics-Simulation and Computation*, pp. 1–17, 2019.

[11] L. Landler, G. D. Ruxton, and E. P. Malkemper, "Circular data in biology: advice for effectively implementing statistical procedures," *Behavioral Ecology and Sociobiology*, vol. 72, no. 8, p. 128, 2018.

[12] J. Cremers and I. Klugkist, "One direction? A tutorial for circular data analysis using R with examples in cognitive psychology," *Frontiers in Psychology*, vol. 9, p. 2040, 2018.

[13] N. I. Fisher, *Statistical Analysis of Circular Data*, Cambridge University Press, 1995.

[14] G. K. Kanji, *100 Statistical Tests*, Sage, 2006.

[15] L. Landler, G. D. Ruxton, and E. P. Malkemper, "Circular statistics meets practical limitations: a simulation-based Rao's spacing test for non-continuous data," *Movement Ecology*, vol. 7, no. 1, p. 15, 2019.

[16] L. Landler, G. D. Ruxton, and E. P. Malkemper, "Model selection versus traditional hypothesis testing in circular statistics: a simulation study," *Biology open*, vol. 9, no. 6, 2020.

[17] K. V. Mardia, C. C. Taylor, and G. K. Subramaniam, "Protein bioinformatics and mixtures of bivariate von Mises distributions for angular data," *Biometrics*, vol. 63, no. 2, pp. 505–512, 2007.

[18] G. Puglisi, A. Leonetti, A. Landau, L. Fornia, G. Cerri, and P. Borroni, "The role of attention in human motor resonance," *PLoS One*, vol. 12, no. 5, article e0177457, 2017.

[19] L. P. Rivest, T. Duchesne, A. Nicosia, and D. Fortin, "A general angular regression model for the analysis of data on animal movement in ecology," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 3, no. 65, pp. 445–463, 2016.

[20] U. Rutishauser, I. B. Ross, A. N. Mamelak, and E. M. Schuman, "Human memory strength is predicted by theta-frequency phase-locking of single neurons," *Nature*, vol. 464, no. 7290, pp. 903–907, 2010.

[21] W. H. Warren, D. B. Rothman, B. H. Schnapp, and J. D. Ericson, "Wormholes in virtual space: from cognitive maps to cognitive graphs," *Cognition*, vol. 166, pp. 152–163, 2017.

[22] P. Yedlapalli, S. V. S. Girija, and A. V. D. Rao, "On stereographic circular Weibull distribution," *Journal of New Theory*, vol. 14, pp. 1–9, 2016.

[23] M.-S. Yang and J.-A. Pan, "On fuzzy clustering of directional data," *Fuzzy sets and systems*, vol. 91, no. 3, pp. 319–326, 1997.

[24] J. B. Benjamin, I. Hussain, and M.-S. Yang, "Possiblistic C-means clustering on directional data," in *2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, Suzhou, China, 2019.

[25] L. Chen, V. P. Singh, S. Guo, B. Fang, and P. Liu, "A new method for identification of flood seasons using directional statistics," *Hydrological Sciences Journal*, vol. 58, no. 1, pp. 28–40, 2013.

[26] O. Kesemen, Ö. Tezel, and E. Özkul, "Fuzzy c-means clustering algorithm for directional data (FCM4DD)," *Expert Systems with Applications*, vol. 58, pp. 76–82, 2016.

[27] M. A. Lubiano, M. Montenegro, B. Sinova, S. de la Rosa de Sáa, and M. Á. Gil, "Hypothesis testing for means in connection with fuzzy rating scale-based data: algorithms and applications," *European Journal of Operational Research*, vol. 251, no. 3, pp. 918–929, 2016.

[28] A. Pewsey and E. García-Portugués, "Recent advances in directional statistics," 2020, https://arxiv.org/abs/2005.06889.

[29] A. Pewsey, M. Neuhäuser, and G. D. Ruxton, *Circular Statistics in R*, Oxford University Press, 2013.

[30] F. Smarandache, *Neutrosophy. Neutrosophic Probability, Set, and Logic vol. 105*, ProQuest information & learning, Ann Arbor, Michigan, USA, 1998.

[31] F. Smarandache, *Introduction to Neutrosophic Measure, Neutrosophic Integral, and Neutrosophic Probability*, Infinite Study, 2013.

[32] S. Broumi, A. Bakali, M. Talea, and F. Smarandache, "Bipolar Neutrosophic Minimum Spanning Tree," *Infinite Study*, 2018.

[33] S. Broumi and F. Smarandache, "Correlation coefficient of interval neutrosophic set," *Applied Mechanics and Materials*, vol. 436, pp. 511–517, 2013.

[34] Y. Guo and A. Sengur, "NCM: neutrosophic c-means clustering algorithm," *Pattern Recognition*, vol. 48, no. 8, pp. 2710–2724, 2015.

[35] M. Abdel-Baset, V. Chang, and A. Gamal, "Evaluation of the green supply chain management practices: a novel neutrosophic approach," *Computers in Industry*, vol. 108, pp. 210–220, 2019.

[36] M. Abdel-Basset, M. Mohamed, M. Elhoseny, L. H. Son, F. Chiclana, and A. E. N. H. Zaied, "Cosine similarity measures of bipolar neutrosophic set for diagnosis of bipolar disorder diseases," *Artificial Intelligence in Medicine*, vol. 101, p. 101735, 2019.

[37] F. Smarandache, "Introduction to Neutrosophic Statistics," Infinite Study, 2014.

[38] M. Aslam, "Neutrosophic analysis of variance: application to university students," *Complex & intelligent systems*, vol. 5, no. 4, pp. 403–407, 2019.

[39] M. Aslam, "A new attribute sampling plan using neutrosophic statistical interval method," *Complex & Intelligent Systems*, vol. 5, no. 4, pp. 365–370, 2019.

[40] M. Aslam, "A new method to analyze rock joint roughness coefficient based on neutrosophic statistics," *Measurement*, vol. 146, pp. 65–71, 2019.

[41] M. Aslam and M. Albassam, "Application of neutrosophic logic to evaluate correlation between prostate cancer mortality and dietary fat assumption," *Symmetry*, vol. 11, no. 3, p. 330, 2019.

[42] J. Chen, J. Ye, and S. Du, "Scale effect and anisotropy analyzed for neutrosophic numbers of rock joint roughness coefficient based on neutrosophic statistics," *Symmetry*, vol. 9, no. 10, p. 208, 2017.

[43] J. Chen, J. Ye, S. Du, and R. Yong, "Expressions of rock joint roughness coefficient using neutrosophic interval statistical numbers," *Symmetry*, vol. 9, no. 7, p. 123, 2017.

[44] J. Pratihar, R. Kumar, S. Edalatpanah, and A. Dey, "Modified Vogel's approximation method for transportation problem under uncertain environment," *Complex & intelligent systems*, vol. 7, no. 1, pp. 29–40, 2020.

[45] M. Aslam, "Radar data analysis in the presence of uncertainty," *European Journal of Remote Sensing*, vol. 54, no. 1, pp. 140–144, 2021.

Hindawi
Journal of Sensors
Volume 2021, Article ID 9376711, 15 pages
https://doi.org/10.1155/2021/9376711

Research Article

Probability Integral Method Parameter Determination by SBAS-InSAR Technology and GWO Algorithm

Tieming Liu,[1] Tongkang Zhang,[2] Lichuan Chen,[3] Weiming Liao,[3] Yun Shi,[2,4] and Danchao Xu[5]

[1]*Xi'an Geological Environment Monitoring Station, Xi'an, Shaanxi 710100, China*
[2]*Aerial Survey and Remote Sensing Bureau of China Coal Geology Bureau, Xi'an 710199, China*
[3]*Chongqing Institute of Geology and Mineral Resources, Chongqing 401120, China*
[4]*Key Laboratory of Coal Resources Exploration and Comprehensive Utilization, Ministry of Natural Resources, Xi'an, Shaanxi 710021, China*
[5]*Wenzhou Polytechnic, Wenzhou 325000, China*

Correspondence should be addressed to Yun Shi; shiyun0908@hotmail.com

Received 25 April 2021; Accepted 9 August 2021; Published 31 August 2021

Academic Editor: Hyung-Sup Jung

This paper proposed a method based on the SBAS-InSAR and gray wolf optimization algorithm aiming at the time-consuming and laborious defects of the traditional method used to obtain the expected parameters of the probability integral method and the shortcomings of the InSAR technology in the field of large gradient deformation detection in the mining area. The fitness function of the algorithm was established based on the geometric relationship between the radar side imaging and the three-dimensional model of the probability integral method. The stable sinking point of the settlement boundary obtained by SBAS-InSAR was used as the input value for the calculation of the predicted parameters of the probability integral method. Firstly, the simulation experiment was employed for the simulation of the direction of the InSAR line of sight combined with the geological mining conditions of the assumed working face, thereby obtaining the probability integral prediction parameters of the working face. Consequently, the maximum relative error of q, b, $\tan\beta$, and θ_0 does not exceed 8%, and that of $S1$, $S2$, $S3$, and $S4$ does not exceed 35.5% (low parameter sensitivity). The error of the LOS-direction deformation fitting is 0.076 m, which meets the tolerance requirements, and the result is trustworthy. At last, the parameter finding method is applied to the engineering example, that is, the 112201 working face of Xiaobaodang Coal Mine in the northern Shaanxi mining area. The settlement value of the stable boundary point is obtained based on the SBAS-InSAR results, which is substituted into the fitness function. And the GWO optimization algorithm is used for optimization and parameter finding; the probability integral expected parameters of the working face are calculated as $q = 0.63$, $b = 0.37$, $\tan\beta = 2.76$, $\theta 0 = 83.94$, $S1 = -36.34$ m, $S2 = 26.69$ m, $S3 = -45.64$ m, and $S4 = 39.62$ m. Substitute the obtained parameters into the probability integral model for the prediction of the vertical and horizontal displacements of the working face, and verify its accuracy with the GPS measured data. The results showed that the maximum absolute error of vertical displacement reached 116 mm, the median error was 63 mm, and the maximum absolute error of north-south horizontal movement reached 56 mm; meanwhile, the median error was 23 mm, the maximum absolute error of east-west horizontal movement reached 61 mm, and the median error was 29 mm; all the above parameters are within the tolerance range, indicating that the method for the calculation of probability integral parameters proposed in this paper is applicable in actual engineering.

1. Introduction

As one of China's main energy sources, coal has always been a pillar industry of the lifeline of the national economy [1]. China is rich in coal resources, and the proven coal reserves rank the third in the world [2], and coal will continue to be China's main energy source for a certain period in the future. It is undeniable that the development and utilization of coal resources have played an important role in the development of China's economy. However, the environmental and geological problems accompanying the mining process are becoming more and more serious. The main environmental problems lie in geological disasters, including ground subsidence, ground fissures, landslides, and mudslides caused by mining [3]. In addition, the movement of rock formations caused by mining often results in the damages of roads, pipelines, villages, and any other infrastructure located in subsidence areas. In addition, it may also cause soil erosion on the ground [4], destroy the stratum structure, and seriously affect the local economic development. Among them, as the most serious problem, the ground subsidence will directly result in the destruction of surface houses, bridges, and other infrastructures, thereby posing a great threat to the safety of lives and property of the residents. In that case, the accurate prediction of these displacements is both necessary and helpful for the assessment of potential damage and the improvement of the design of coal mining. In recent decades [5], the techniques, such as empirical prediction methods [6] and influence function methods, have been developed, which often require a large number of repeated and accurate three-dimensional displacement measurement samples to build a reliable prediction model before performing forward displacement prediction [7]. However, to monitor ground displacement caused by mining. The traditional geodetic methods, such as leveling and global positioning system measurements, are usually based on the measurement at only a few sparse points, which makes it quite difficult to use these sparse point measurements to describe the entire subsidence area [8]. In addition, these methods are costly, time-consuming, and highly labor-intensive which often result in unreliable prediction models and inaccurate forward predictions. Interferometric Synthetic Aperture Radar (InSAR) technology can play a certain role and greatly improve the reliability of predicting mining displacement in an economically feasible way due to its low cost, high accuracy and efficiency, large spatial coverage area, high temporal and spatial resolution [9], and the availability to measure surface deformations related to geophysical and human activities. However, due to the fact that the mining area subsidence is often accompanied by a large number of large gradients, the detection capability of InSAR will be constrained. This is because that the impact of spatial miscorrelation is relatively large [10], resulting in hollow area appeared in InSAR interferograms. This is also a big limitation of the current mining area deformation monitoring by InSAR. However, because the subsidence basin is characterized by small boundary and large center in terms of magnitude, InSAR can detect the settlement information of stable boundary points [11].

To better protect the safety of surface buildings and facilities, it is necessary to predict and protect the surface movement caused by mining. At present, the probability integral method is the most widely used mining subsidence prediction method [12], whose prediction quality is mainly determined by the accuracy of the predicted parameters [13]. Because in the probability integral method, the predicted parameters are affected by the complex geological factors and mining conditions; the expression with simple mathematical methods is not available. Moreover, the obtaining of parameters through actual measurement based on the establishment of a mobile surface observation station is also easily restricted by a variety of conditions, and it is difficult to meet the requirements of safe production in mining areas. With the development of computer technology, various artificial intelligence algorithm methods are mastered and applied to the field of probabilistic integration to obtain parameters, which mainly include neural network model [14], support vector machine [15], genetic algorithm [16] (GA), and Particle Swarm Optimization (PSO) [17]. A large number of experiments have proved that the utilization of machine learning to obtain the parameters of the probability integral method shows the advantages of high accuracy, easy operation, and strong practical effect. Gray wolf optimization (GWO) [18] is a population intelligent optimization algorithm, which is mainly characterized by simple operation, few parameters, and easy implementation. Article [19] improved the convergence factor strategy and introduced a dynamic weight strategy to optimize the performance of the gray wolf algorithm; Article [20] combined the advantages of differential evolution and the gray wolf algorithm and proposed a hybrid algorithm for three-dimensional chip testing to further balance the algorithm's exploration and mining capabilities; Article [21] proposed a hybrid gray wolf optimization algorithm based on chaos principle and elite reverse learning strategy to solve the problem of high-dimensional function optimization; Article [22] embeds an optimal optimization algorithm. The difference reverse learning strategy and a dynamic random difference mutation operator are used to improve the gray wolf optimization algorithm to improve the clustering optimization performance; Article [23] proposed a hybrid algorithm of gray wolf optimization and locust optimization to solve the text mining problem. Article [24] uses the GWO algorithm to invert the parameters of the wind turbine power model. Therefore, this paper constructs the fitness function of the gray wolf optimization algorithm based on the probabilistic integral method model and the principle of InSAR side-view imaging and proposes a probabilistic integral parameter calculation method that combines SBAS-InSAR and GWO algorithms. This method combines the advantages of InSAR technology and machine learning algorithm and reverses the predicted parameters of the probability integral method. The method is novel and provides new method support for the parameter inversion of the probability integral method and the field of mining subsidence prediction.

2. Introduction to Gray Wolf Optimization Algorithm

Gray wolf optimization (GWO) [18] is a population intelligent optimization algorithm derived from the wolf pack search and hunting activities. The wolf pack has a strict hierarchical distribution relationship used to determine the optimal parameter value in the optimization process. Similarly, this paper uses the global search ability of the GWO algorithm in the process of searching for prey to perform global optimization, so as to obtain the predicted parameters with an optimal probability integral method. The specific establishment steps are as follows.

Firstly, it is assumed that the population size of the wolf pack consists of n individuals, and then, the position of the i th wolf is determined, where it refers to the d-dimensional position of the ith wolf in a multidimensional space, and the first three optimal solutions of the objective function value are provided. There are three wolves with the highest rank, namely, αwolf, βwolf, and δwolf. And the three wolves are mainly responsible for hunting. Besides, the hunting behavior can be described by the following formula:

$$D = |QX_P(t) - X(t)|, \tag{1}$$

$$Q = 2r_1. \tag{2}$$

Formula (1) represents the distance between the individual and the prey, and formula (2) is the position update formula of the gray wolf. In the above formula, t is the number of iterations, A and Q are coefficient vectors, and $X_P(t)$ and $X(t)$, respectively, represent the positions of the prey and gray wolf after t iterations. The calculation formula of A and Q is

$$X(t+1) = X_P(t) - AD, \tag{3}$$

$$A = 2ar_2 - a, \tag{4}$$

$$a = 2 - 2 \times \left(\frac{t}{T}\right). \tag{5}$$

In the above formula, a is the convergence factor, as the number of iterations decreases linearly from 2 to 0, t is the current iteration number, and T is the maximum iteration. r_1 and r_2 take random numbers between [0,1].

The most important thing for wolves to do in hunting is to be able to determine the location of their prey, so we use αwolves, βwolves, and δwolves to search for the potential location of the prey, and lead other search individuals to constantly update the potential location according to these three optimal solutions, thereby obtaining the position of the prey. The specific steps are as follows:

$$D_\alpha = |Q_1X_P(t) - X(t)|, \tag{6}$$

$$D_\beta = |Q_2X_P(t) - X(t)|, \tag{7}$$

$$D_\delta = |Q_3X_P(t) - X(t)|, \tag{8}$$

$$X_1(t+1) = X_\alpha(t) - A_1D_\alpha, \tag{9}$$

$$X_2(t+1) = X_\beta(t) - A_2D_\beta, \tag{10}$$

$$X_3(t+1) = X_\delta(t) - A_3D_\delta, \tag{11}$$

$$X_p(t+1) = \frac{X_1 + X_2 + X_3}{3}, \tag{12}$$

and represent the distance between the position of the candidate wolf and that of the best three wolves, respectively. When $|A| > 1$, gray wolves will spread out to search for prey, while when $|A| < 1$, gray wolves will gather together to search for prey.

3. Probability Integral Model and InSAR Geometric Relationship to Establish Fitness Function

SBAS-InSAR technology is available to monitor and quantify the ground displacement that occurred in adjacent time intervals. Since the radar is based on side-view imaging, it can only measure the path length difference in the line of sight (LOS) direction [25], instead of the true three-dimensional deformation of the surface. To quantify the relationship between the actual three-dimensional ground displacement and the LOS deformation measured by SBAS-InSAR, it is assumed that the east-west, north-south, and up-and-down displacement components of the actual ground are U_E, U_N, and W, respectively, and then, we project U_E, U_N, and W to radar LOS direction [26]. The ascending satellite orbit with the azimuth angle αh refers to the angle between the north and the orbit direction (clockwise).

As shown in Figure 1, the LOS deformation is jointly contributed by the real three-dimensional ground displacement. In the case that the three-dimensional ground displacement in mining engineering can be expressed by the model parameters of the mechanical model, certain functional relationship exists between the model parameters based on equation (15) and the LOS deformation. Developed from the random medium theory, and derived from the probability density function in the prediction equation, PIM presents one of the most commonly used influence function methods in underground mining [27]. According to PIM theory, the vertical settlement $W(x, y)$ and horizontal displacement $U(x, y, \phi)$ caused by underground mining can be expressed as

$$\begin{aligned} W_{(x,y)} &= \frac{W^0(x)W^0(y)}{W_0}, \\ U_N &= \left[U^0(x)W^0(y)\cos\varphi_N + U^0(y)W^0(x)\sin\varphi_N\right], \\ U_E &= \left[U^0(x)W^0(y)\cos\varphi_E + U^0(y)W^0(x)\sin\varphi_E\right], \end{aligned} \tag{13}$$

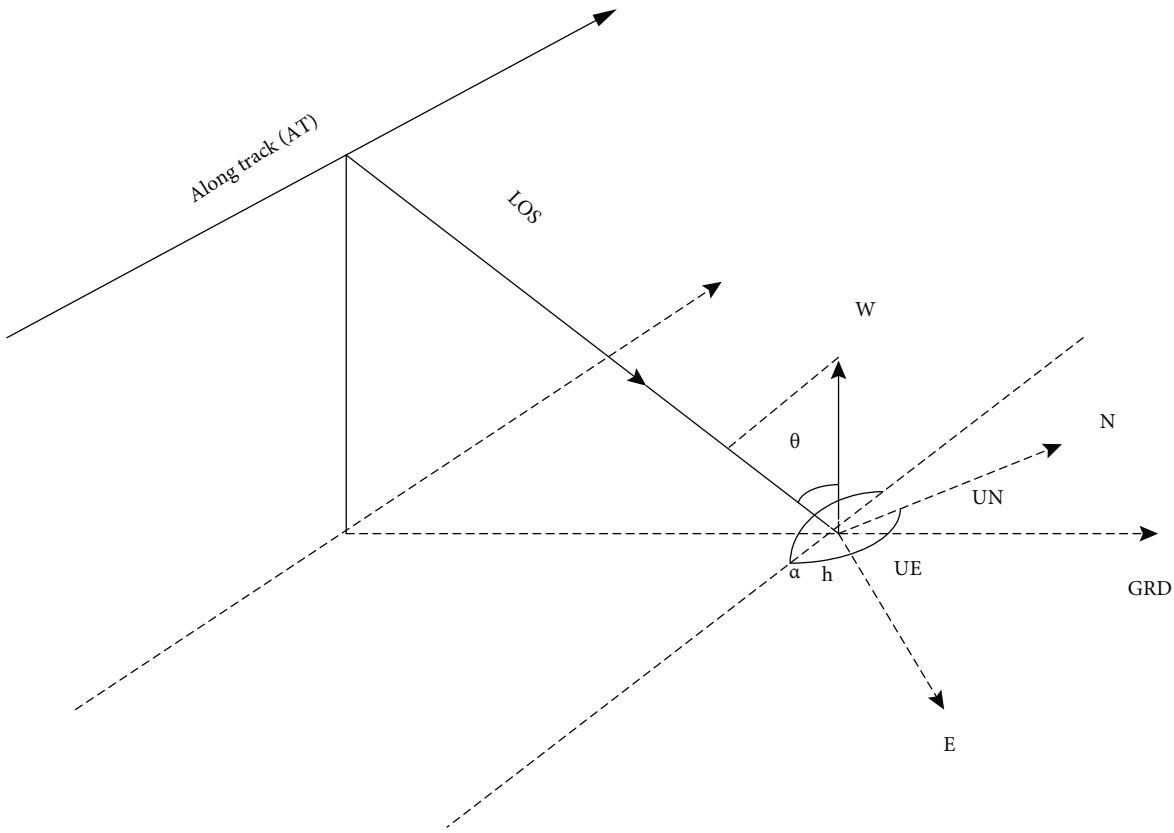


Figure 1: Schematic diagram of radar side-view imaging.

among them,

$$\begin{gathered} W_0 = mq\cos\alpha, \\ \begin{cases} \mathrm{W}^0(x) = \mathrm{W}'(x,r) - \mathrm{W}'(x-l,r), \\ \mathrm{W}'(x,r) = \dfrac{\mathrm{W}^0}{2}\left[\mathrm{erf}\left(\dfrac{\sqrt{\pi}}{r}x\right)+1\right], \end{cases} \\ W^0(y) = W'(y,r_1) - W'(y-L,r_2), \\ U'(x,b,r) = bW_0\exp\left(-\pi\dfrac{x^2}{r^2}\right), \\ \begin{cases} U^0(x) = U'(x,b,r) - U'(x-l,b,r), \\ U^0(y) + U'(y,b,r) + W'(y,r)\cot\theta_0 - U'(y-L,b,r) - W'(y-L,r)\cot\theta_0. \end{cases} \end{gathered} \tag{14}$$

Then, project to the LOS direction through the projection formula:

$$\mathrm{LOS} = [\cos(\theta) - \sin(\theta)\cos(\alpha)\sin(\theta)\sin(\alpha)]\begin{bmatrix} W \\ U_E \\ U_N \end{bmatrix}, \tag{15}$$

where θ refers to the incident angle of the radar, and LOS represents the deformation derived from SBAS-InSAR, which is defined as the direction of the line of sight of the satellite. W_0 represents the maximum possible settlement; q denotes settlement factor; erf represents the probability density.

$$\mathrm{erf}\left(\frac{\sqrt{\pi x}}{r}\right) = \frac{2}{\sqrt{\pi}}\int_0^{\left(\sqrt{\pi}/r\right)x} e^{-u^2du}. \tag{16}$$

$r = H/\tan\beta$ is the main influence radius of the mining area, where $\tan\beta$ stands for the tangent of the main influence angle; b denotes the horizontal displacement constant.

$$\begin{gathered} l = D3 - S3 - S4, \\ L = \frac{[(D1 - S1 - S2)\sin(\theta_0+\alpha)]}{\sin\theta_0}. \end{gathered} \tag{17}$$

$D1$ and $D3$ represent the length of the working face's inclination and strike direction, respectively, where $S1$, $S2$, $S3$, and $S4$ represent the deviation of the inflection point in the up, down, left, and right directions, respectively, θ_0 refers to the mining propagation angle; φ represents the clamp between the given direction and the forward direction of the work panel angle (counterclockwise) [28].

The fitness function of the GWO algorithm can be expressed as follows:

$$P = \min\left\|\mathrm{LOS}_0(x,y) - \mathrm{LOS}_p(x,y)\right\|, \tag{18}$$

where x and y refer to the vectors containing the pixel coordinates of the LOS displacement (indicated by LOS_0) measured by the SBAS-InSAR technology, and LOS_p represents

the predicted ground displacement in the LOS direction of the radar with a predetermined functional relationship. The 8 optimal parameter values of the probability integral method are used as the optimal position to be searched for by the gray wolf.

4. Simulation Experiment

4.1. Simulation of Geological Mining Conditions. Based on the coal measure strata in the northern Shaanxi mining area as the background, the overlying rock lithology of the simulated working face is set as medium hard, and the horizontal projections of $D1$ and $D3$ of the mining size along the strike and the inclination direction are 200 m and 800 m, respectively. Besides, the mining height H is set as 300 m, the coal seam inclination angle is 2°, and the mining thickness $M = 6$ m; the strike long-arm mining method is adopted for mining, and the all caving method is adopted for roof management. Because $D1/H \approx 0.66 < 1.2$ and $D3/H \approx 2.66 > 1.2$ in both the strike and inclination of the working face, insufficiency mining is used in the inclination, sufficiency mining is used in the strike of the working face, and in general, insufficiency mining is adopted [29]. The characteristics of the subsidence basin in the northern Shaanxi mining area conform to the probability integral method model. Therefore, in this simulation experiment, the main predictive parameters of the probability integral method are set as subsidence coefficient $q = 0.8$, horizontal movement coefficient $b = 0.25$, and tangent value of the main influence angle tan $\beta = 2$, maximum sinking angle $\theta_0 = 85°$, and inflection point offset $S1 = S2 = S3 = S4 = 20$ m.

4.2. InSAR Line-of-Sight Strain Simulation. The main SAR data used in this paper is the Sentinel-1A data of ESA. When the satellite is over northern Shaanxi for geometric imaging, the satellite's incident angle $\theta = 39.13°$, and the orbital inclination angle $\alpha h = 346.7°$. The pixel size is selected according to the spatial resolution of the SAR data during simulation. According to the set probability integral method model parameters, the subsidence value W, the horizontal movement value U_N in the north-south direction, and the horizontal movement value U_E in the east-west direction of the subsidence basin boundary point are predicted. Then, the orbit parameters of the SAR satellite and the projection relationship are used to calculate the LOS-direction deformation value of each point, with the simulation effect shown in Figure 2.

4.3. Analysis of Experimental Results. The boundary point of the LOS deformation value of the upper simulated working face was extracted as the input value of the algorithm fitness function, and the probability integral prediction parameter calculation model of the InSAR fusion gray wolf optimization algorithm was established for the calculation of the probability integral model parameters under the conditions of the mining area, thereby eliminating the influence of gross errors in the calculation process. Five experiments were carried out on the algorithm, and the average value was calculated to obtain a reliable result.

The specific steps for obtaining the parameters are as follows:

Step 1. Input the initial parameters of PIM. According to the geological mining conditions of the Loess mining area, the initial values of each parameter are set as $q = 0.8$, $b = 0.25$, tan $\beta = 2$, $\theta_0 = 85$, $S1 = 20$, $S2 = 20$, $S3 = 20$, and $S4 = 20$.

Step 2. Determine the initial parameters of the differential gray wolf optimization algorithm. The parameter values set in this paper are 30 populations, 500 maximum iterations, and 8 independent variable dimensions; besides, the boundaries are set.

Step 3. Determine the initial position of the population, calculate the value of the objective function in the parent population, and determine the three best individuals in the population as αwolf, βwolf, and δwolf which are denoted as X_α, X_β, and X_δ.

Step 4. Use equations (6)–(8) to find the distance between other individuals and the best three individuals, and then, update the position of each individual in the population according to equations (9)–(11).

Step 5. Iterate, calculate, and update the values of parameters Q and A according to formulas (2), (4), and (5), and update the position of X_α, X_β, and X_δ to prevent border crossing.

Step 6. Update the positions of the optimal αwolf, βwolf, and δwolf in the parent population.

Step 7. Determine the number of iterations; in the case that the number of iterations reaches the maximum value, select the optimal value of PIM parameter; otherwise, return to Step 4 to continue.

The results of parameter calculation are shown in Table 1.

The effect of parameter fitting is shown in Figure 3.

As shown in Table 1, the results of the probability integral predicted parameter obtained by the LOS deformation based on InSAR combined with the gray wolf optimization algorithm are more accurate. Among them, the maximum relative error of q, b, tanβ, and θ_0 fails to exceed 8%, and that of the inflection point offset $S1$-$S4$ does not exceed 35.5% (the parameter sensitivity is low and has little effect on the overall effect). Figure 3 illustrates the LOS fitting diagram in the process of parameter obtaining. According to the figure, the algorithm shows a better overall fitting effect on the simulated values of the InSAR boundary points, and the trends are more consistent. The residual error range fluctuates from -0.13 m to 0.109, the median error is 0.076 m, and the result is trustworthy. It suggests that the method to obtain the parameters of the probability integral combining the SBAS-InSAR and GWO algorithms is applicable in practice.

5. Case Analysis

5.1. Overview of the Study Area. Xiaobaodang Minefield is located in the southwest of Shenmu County, Yulin City, Shaanxi Province, bordering the Inner Mongolia Plain in the northwest, facing the Yellow River in the east and Xing

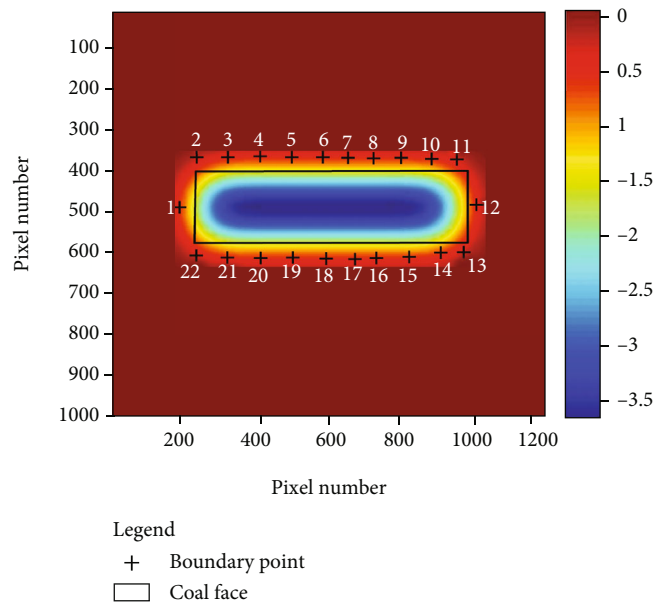


Figure 2: Simulated LOS to deformation value.

Table 1: GWO expected parameter inversion results.

Serial number	q	b	$\tan\beta$	θ_0	$S1$ (m)	$S2$ (m)	$S3$ (m)	$S4$ (m)
1	0.82	0.31	1.88	82.95	19.45	13.65	-14.31	8.48
2	0.86	0.25	2.13	86.12	17.29	16.22	-14.91	19.42
3	0.86	0.27	1.61	84.02	22.19	14.74	-12.81	16.26
4	0.83	0.26	2.06	83.22	23.8	12.81	-9.34	13.75
5	0.85	0.28	2.08	85.26	24.58	12.65	-13.15	17.47
Average	0.84	0.27	1.95	84.31	21.46	14.01	-12.90	15.08
Parameter true value	0.8	0.25	2	85	20	20	20	20
Absolute error	0.04	0.02	-0.05	-0.69	1.46	-5.99	7.1	-4.92
Relative error%	5	8	-2.5	-0.81	7.3	-29.95	35.5	-24.6

County in Shanxi Province across the river, and adjacent to Jiaxian County, Yulin City, Shaanxi Province in the south, which is under the jurisdiction of Dabaodang Town, Shenmu County, Yulin City. Its geographical coordinates are as follows: east longitude $109^\circ 49' 09.6553'' \sim 109^\circ 59' 41.9675''$; north latitude $38^\circ 39' 15.8090'' \sim 38^\circ 47' 32.0113''$. The Xiaobaodang Minefield is delineated by the connecting lines of four inflection points, with a width of 7.7 km from north to south and a length of 15.2 km from east to west, covering an area of 97.18 km^2. The traffic around Xiaobaodang Minefield is well developed, specifically, the Yushen Class II Highway runs in the southeast of the minefield, and the Baomao Expressway and the 210 National Highway pass through the west. Xiaobaodang Mine Field is located at the northern end of the Loess Plateau in northern Shaanxi, on the southeastern edge of the Mu Us Desert. The area is eroded by the Quaternary aeolian sand all year round, with most of the area covered by fixed dunes and semifixed dunes. The landforms are mainly characterized by wind-eroded aeolian desert hilly landforms. From the topographical point of view, the overall trend is higher in the southwest and lower in the east. The highest point is 1346.60 m. Shibuliang is located in the south. The lowest point in the east is 1198 m east. The elevation range is between 1250 and 1300 m, and the relative maximum height difference is 148.60 m.

The 112201 working face of Xiaobaodang Coal Mine is shown as the black rectangle mark in Figure 4. The 112201 working face was in operation from September 2018 to November 2019 adopting long-arm mining, and the all-caving method for roof management. Besides, sufficient mining was used in the strike direction of the working face, and insufficient mining was utilized in the inclination direction; in general, insufficient mining was used. The geometric parameters are summarized as follows: average mining depth $H \approx 340$ m, average mining thickness $m \approx 6$ m, mining length $D3 \approx 4560$ m, mining width $D1 \approx 350$ m, and dip angle $\alpha \approx 1^\circ$. Two monitoring lines, i.e., Z and Q, are arranged along the direction of the working face and the main section of the inclination; besides, 37 monitoring

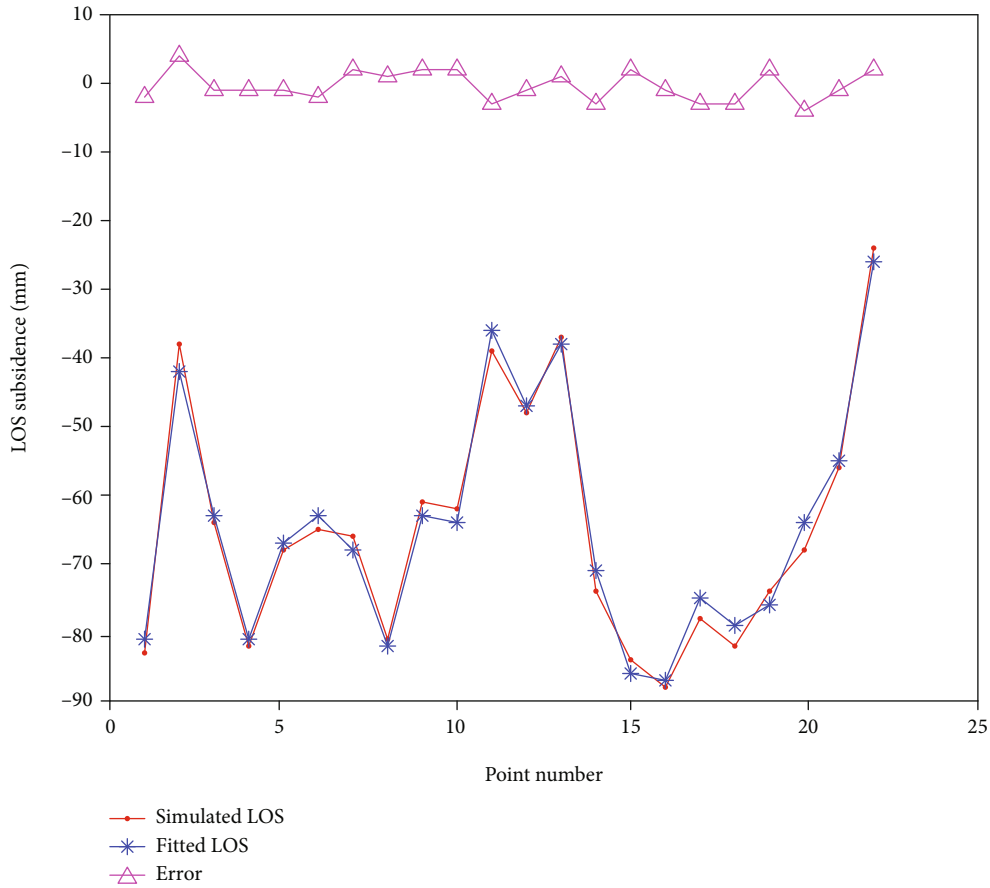


Figure 3: Comparison of simulated LOS deformation and fitted LOS deformation.

points are arranged in the strike direction, 40 monitoring points are arranged in the inclined direction, and 77 monitoring points are arranged in total. The monitoring line is shown in Figure 2. Based on the constructed model and the simulated mining subsidence prediction parameters (Table 1), the subsidence and horizontal movement after the mining subsidence and stabilization are predicted.

The location of the study area is shown in Figure 4.

5.2. SBAS-InSAR Subsidence Basin Boundary Information Extraction

5.2.1. Data Processing. The experimental data are mainly based on the data of Sentinel-1A of ESA. The Sentinel satellite is an earth observation satellite launched by ESA in 2014, which is equipped with a C-band synthetic aperture radar, and available for all-weather earth observation throughout the day. The revisit period of a single constellation is 12 days, in which the interference width can reach 250 km, and the ground resolution is $(5 * 20)$ m. In this paper, 41 scenes of the ascending orbit were extracted from the archived data between September 3, 2018, and January 8, 2020.

The data processing of SBAS-InSAR is mainly divided into two parts, and the first part is data preprocessing. The preprocessing part mainly includes the extraction of single-view complex image (SLC) data, and the selection of super master images for subpixel precision registration of the data. The second step is to remove the error phase. Firstly, set the space baseline and time baseline as 200 m and 33 d, respectively, according to the geological conditions of the study area. After phase unfolding, the maximum spatial baseline is 113.7 meters, and the minimum is 2.8 meters. As shown in Figure 5, select 20190113 as the main image. The generation of the interferogram is the most critical step; in this process, the GoIdstein filter is used to improve the phase expansion accuracy and measurement accuracy, thereby maximizing the signal-to-noise ratio of the interferogram. The minimum cost flow method is employed to unwrap the phase, and the coherence threshold is set at 0.2. With the aid of precision orbits and DEM, flat ground, phase, and noise are removed, and a high-coherence region unwrapping map is obtained. The coherence coefficient method is used for the extraction of high-coherence points, and 62 high-quality interference pairs are selected. The linear model equations of linear deformation rate and elevation error are established in the region of high coherence point, and the SVD algorithm is used to solve the model parameters, thereby obtaining the linear deformation phase. After the removal of the linear deformation phase, the residual error contains atmospheric phase, nonlinear deformation, and noise. The residual can be subjected to high-pass filtering in the time domain and low-pass filtering in the spatial

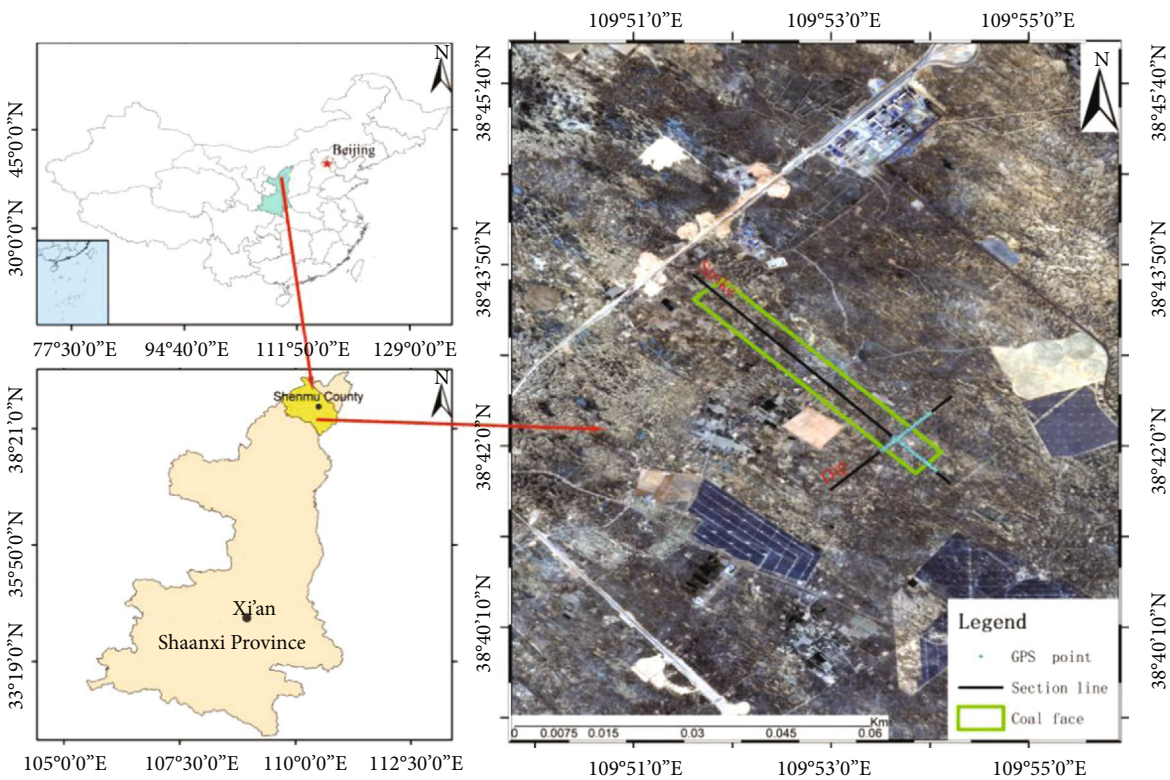


Figure 4: Location map of the study area.

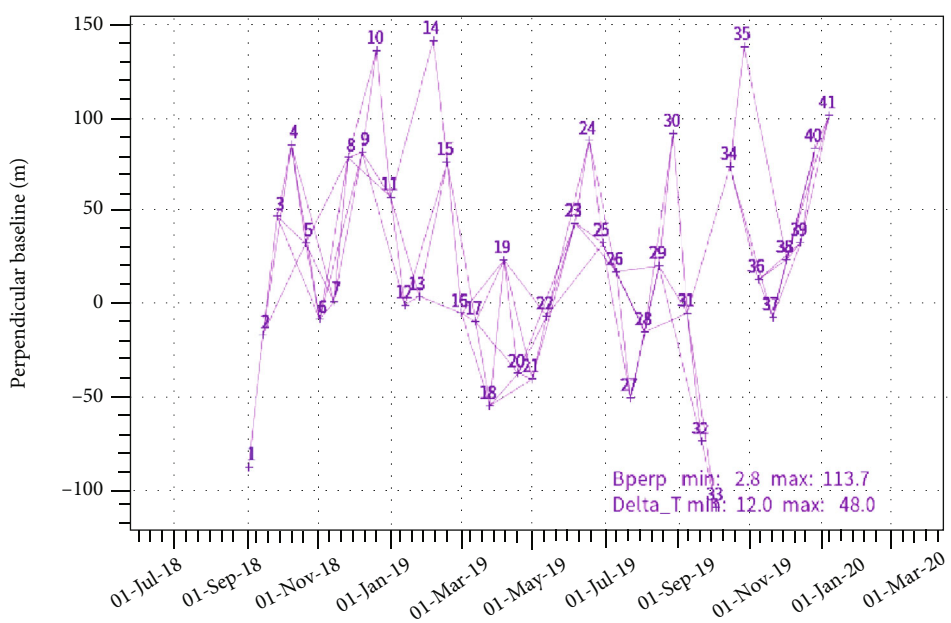


Figure 5: SBAS-InSAR time-space baseline diagram.

domain according to the characteristics of each component, thereby removing the atmospheric phase components and noise to obtain a nonlinear deformation phase. Finally, the least square method is utilized to obtain the deformation rate and cumulative deformation value of each coherent point.

5.2.2. Result Analysis. Figure 6 shows the annual average settlement rate of the study area. It can be seen from the figure that the center of the subsidence basin has been seriously out of coherence, resulting in the lack of most of the settlement information in the center of the working face, and only the marginal settlement with a smaller magnitude is still

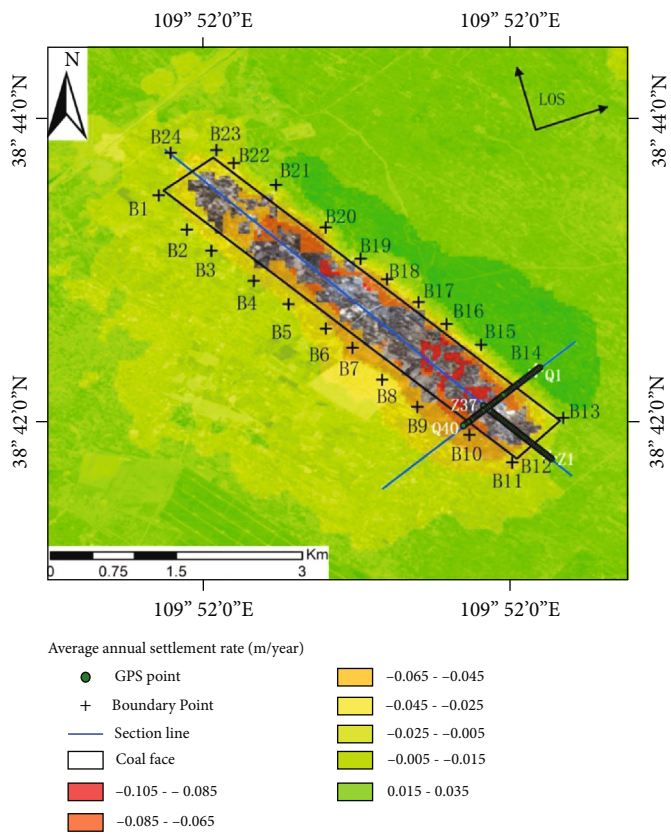


Figure 6: The annual average settlement rate of the study area.

preserved better. The settlement magnitude ranges from -0.105 m to 0.035 m, which is relatively small. The blue line is the two cross-sections of the direction and the inclination. The ground GPS observation points are arranged along the direction and the inclination, respectively, which are the green points in the figure. The direction point number is Z1-Z37, and the inclination point number is Q1-Q40. There are labels in the figure. In order to further study the time series settlement trend of the boundary points, the time series diagram obtained by SBAS-InSAR is derived as shown in Figure 7.

As shown in Figure 7, the black area refers to the working surface, the blue line represents the two section lines along the strike and the inclination, and the black mark points around the working surface are the extracted boundary points. The boundary points of B12 and B14 coincide with the points of Z2 and Q3 on the oblique observation line in the strike and inclination direction, which can be used to detect the observation accuracy of boundary points. During the study period, the 112201 working face was being mined, so the settlement area was mainly concentrated in the 112201 working face. It can be seen from the time sequence diagram that due to the unique geological conditions of the study area, the surface deformation due to coal mining in the Loess Plateau area is characterized by large gradients and large scale, which causes the spatial decorrelation of the SAR interferogram. In that case, high-coherence points are rarely selected above the working surface, which results in the absence of goaf center in the InSAR topographic map. However, based on the basin effect of the goaf surface subsidence and the law of surface deformation of mining subsidence, it can be found that the mining center of the mining area has the characteristics of large deformation speed. However, the basin boundary is a slow linear deformation point, which is in line with the deformation point that can be detected by InSAR. Therefore, only the time series settlement of the boundary point is realized in the timing diagram, and the mining center is not reflected in the InSAR results. Therefore, in this paper, we only discuss the deformation law of the boundary point of the subsidence basin. It can be seen that the settlement funnel extends from southeast to northwest, which is consistent with the mining direction. To verify the monitoring accuracy of SBAS-InSAR, B12 and B14 are utilized to verify their accuracy. Due to the fact that the deformation in the LOS direction is obtained by InSAR, while the GPS observation obtains the vertical and horizontal displacement, firstly, the displacements in three directions, i.e., the vertical direction, east-west, and north-south, obtained by the GPS observation are projected to the LOS upward, and the projection formula is shown in equation (15).

As shown in Figure 8, the GPS measurement results are projected to the LOS backward and compared with SBAS-InSAR. The two GPS observations at points B12 and B14 in the early stage of mining are two. Comparing the results under the conditions that the time and space references are unified, it can be seen that the GPS measured values at points B12 and B14 are slightly larger than the InSAR

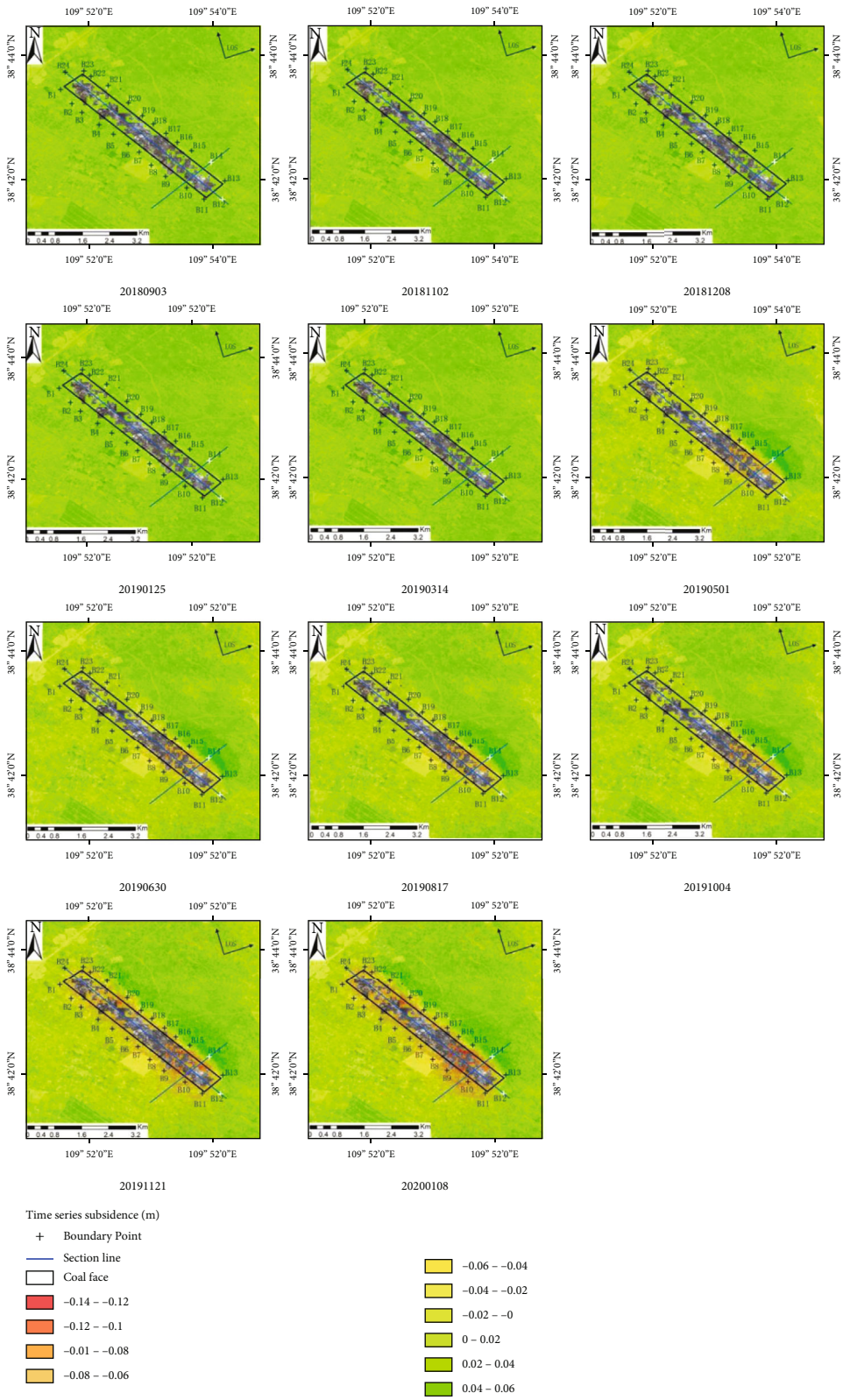


Figure 7: SBAS-InSAR monitoring time series deformation diagram.

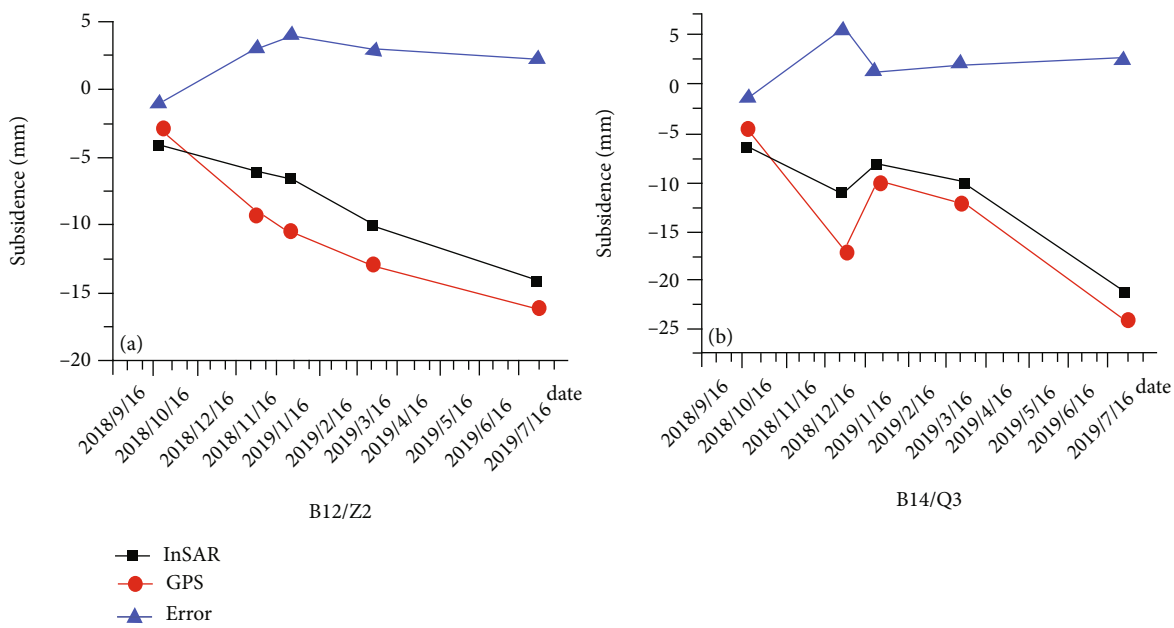


Figure 8: GPS/InSAR comparison results of boundary points.

monitoring value. And the maximum residual error is less than 5 mm, while the maximum residual error at point Q3 is 7 mm. It can be seen that the overall fit between the InSAR results and the GPS results is good, the correlation is high, and the residual line fluctuates in a small range near the 0 value; besides, the fluctuations are stable. It shows that the results of SBAS-InSAR at boundary points are trustworthy, which can be used to invert model parameters in mining subsidence.

5.2.3. Analysis of GWO Algorithm Parameter Inversion Results. Since the mining in 112201 working face was started in September 2018 and ended in November 2019, the subsidence basin was basically stable till the last scene selected in this paper during the imaging in January 2021. Therefore, the cumulative settlement in the last scene of the above SBAS-InSAR is used as the LOS input value of the fitness function, and the error fitting diagram of the fitness function is shown in Figure 9.

Due to the instability of the algorithm, the parameter value will be affected by the gross error. To mitigate the influence of these errors and the instability of the algorithm, we especially solve the parameters for 5 times and find the average value. The result of the solution is shown in Table 2.

As shown in the table above, the prediction results of the probability integral method are $q = 0.63$, $b = 0.37$, $\tan\beta = 2.76$, $\theta_0 = 83.94$, $S1 = -36.34$ m, $S2 = 26.69$ m, $S3 = -45.64$ m, and $S4 = 39.62$ m. To further verify the accuracy of the calculated parameters, we substituted the result parameters into the probability integral model, combined with the geological and mining conditions of the working face, to determine the vertical and horizontal displacements of the observation points on the observation line in the strike direction and the inclination direction of the working face. In addition, predictions and analyses of the accuracy of each observation point based on the GPS measured data were made, with the results shown as follows.

It is found in Figure 10 that the surface movement obtained by the inversion is in good agreement with the measured results. The point with the largest absolute error of sinking is Point Q24 on the oblique observation line, the absolute error of this point reaches 116 mm, and the median error is 63 mm. The point with the largest absolute error of north-south horizontal movement is Point Z15, with the error of 56 mm, and the median error of 23 mm. The point with the largest absolute error of east-west horizontal movement is Point Z34, with the maximum absolute error of 61 mm, and the median error of 29 mm, both of them are within the tolerance range. It illustrates that the inversion of the probabilistic integration method parameters based on the SBAS-InSAR and gray wolf optimization algorithm used in this paper is feasible, which is available to obtain more reliable surface movement probabilistic integration method parameters.

Since the prediction results of a small number of points on the edge of the subsidence basin with smaller settlement magnitudes are not representative, we only consider the points with settlement magnitudes greater than 100 mm in our accuracy evaluation and analyze the accuracy of the results, mainly from the absolute error (AE), relative error (RE), and average absolute percentage error (MAPE) evaluated. And a one-sample t test with a significance of 0.05 and a test value of 60 was done for the sample point error. The specific test results are shown in Table 3.

The error is tested and found to be in accordance with the normal distribution. A one-sample t test with a significance of 0.05 is performed. The test value is set to 50. The significance levels in the three directions of the test are 0.015, 0, and 0, indicating rejection null hypothesis. There is a 95% probability that the error value is less than 50,

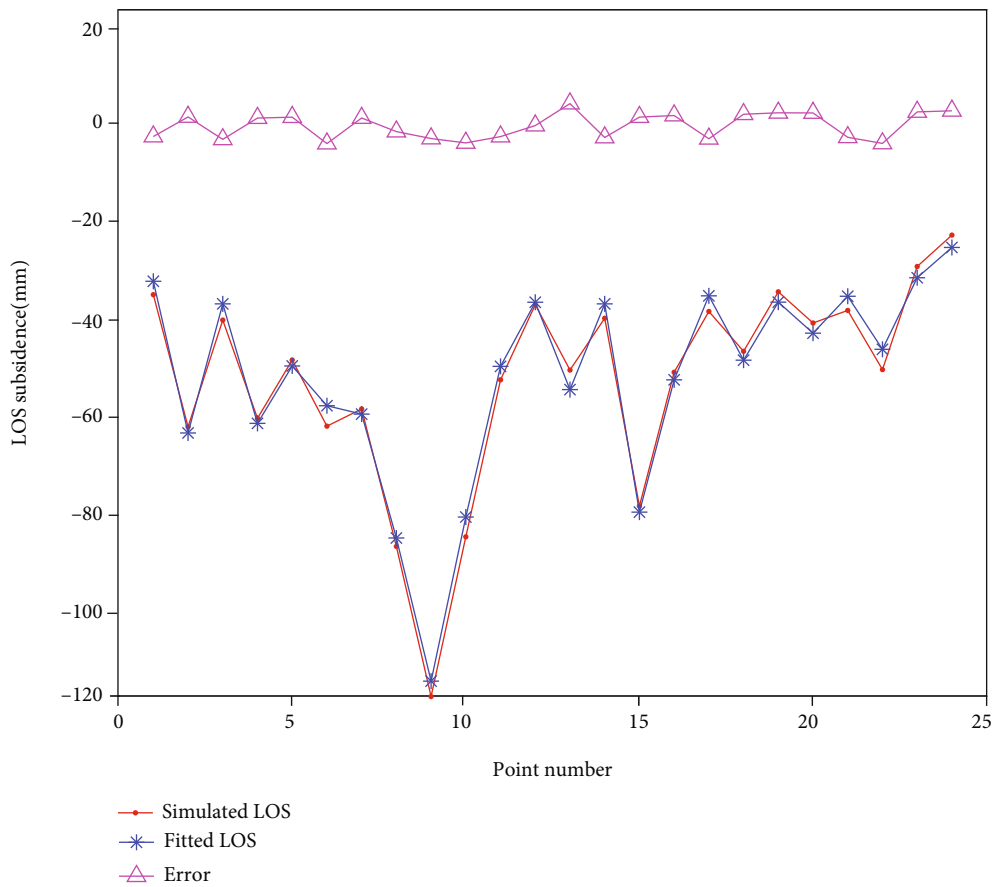


Figure 9: LOS fitting residual plot of fitness function.

Table 2: GWO expected parameter inversion results.

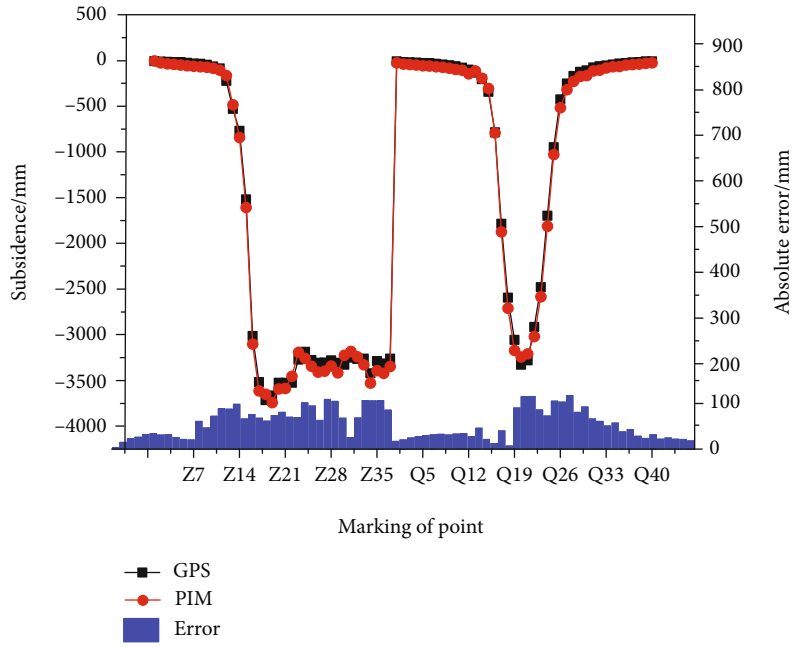
Serial number	q	b	$\tan\beta$	θ_0	$S1$ (m)	$S2$ (m)	$S3$ (m)	$S4$ (m)
1	0.62	0.35	2.36	85.12	-39.06	24.78	-46.62	41.25
2	0.58	0.39	2.93	83.19	-35.15	26.16	-44.89	37.54
3	0.65	0.37	3.05	83.24	-34.94	28.57	-49.18	43.16
4	0.67	0.37	2.74	83.22	-37.99	26.96	-38.52	33.76
5	0.64	0.35	2.71	84.92	-34.58	27.00	-48.99	42.39
Average	0.63	0.37	2.76	83.94	-36.34	26.69	-45.64	39.62

and the confidence interval is estimated. Further, verify the accuracy of the predicted value.

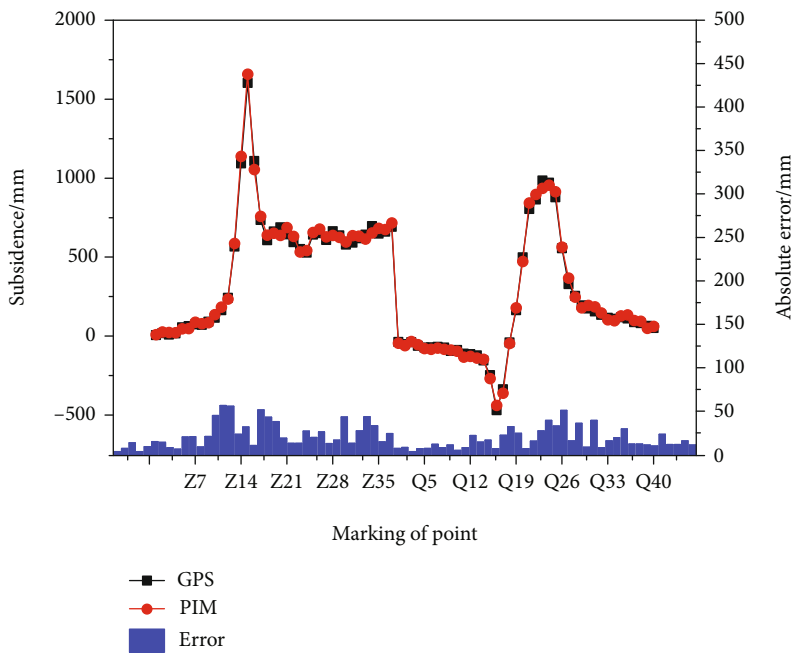
6. Discussion

The prediction of mining subsidence has always been a problem worthy of discussion. The InSAR-PIM prediction algorithm proposed in this paper can be widely used in the field of mining subsidence prediction and has the following advantages. First, this method is based on the most accurate prediction method in the field of mining subsidence, and it has certain reliability. The second method is an improvement on the traditional probability integration method, which is obtained through machine learning methods. The predicted parameters of the probability integral method avoid the cumbersomeness of traditional field observations. Third, this method combines InSAR technology, machine learning algorithms, and probability integration theory. It is a combination of multiple methods and advanced mining subsidence prediction algorithms, which can provide a basis for future research in the field of mining subsidence prediction.

This method is based on the SBAS-InSAR technology and the GWO algorithm to obtain the predicted parameters of the probability integral method, which avoids the cumbersomeness of obtaining the predicted parameters of the probability integral method based on setting the rock movement observation station of the working face, but it also has certain limitations. First, this method is restricted by the resolution and wavelength of the SAR data, as well as the unwrapping error and atmospheric delay error in the data processing process. These factors may cause the accuracy of the edge points of the subsidence basin extracted by SBAS-InSAR to be insufficient. This directly affects the accuracy of the later parameters. Secondly, the gray wolf optimization algorithm is not fast enough to converge and the global optimization ability is not strong enough, which may cause the parameter inversion effect to be not good enough. The next step is to improve the optimization ability of the algorithm to make it more accurate. The expected accuracy.

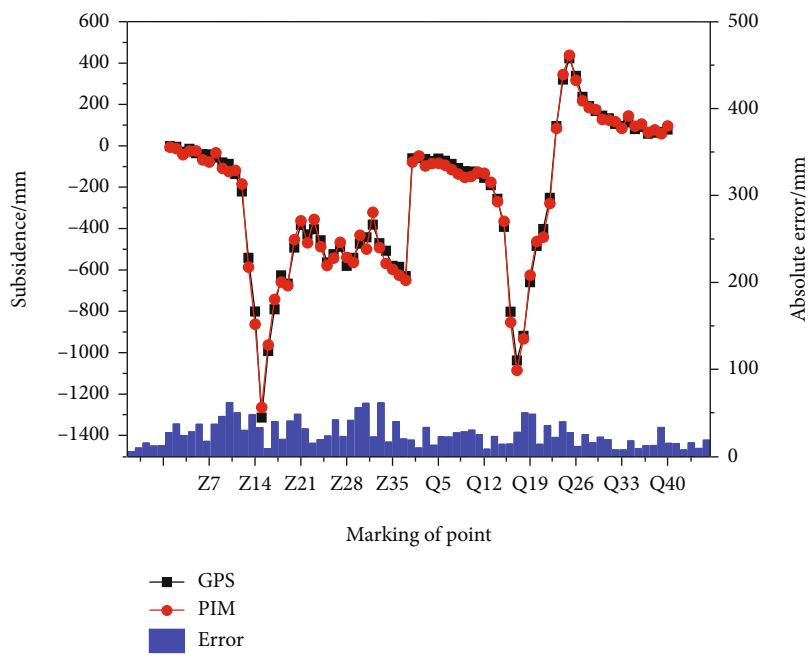


(a) Accuracy comparison chart in the vertical direction of the observation point



(b) Accuracy comparison chart for the north and south of the observation point

Figure 10: Continued.



(c) Accuracy comparison chart in the east-west direction of the observation point

Figure 10: Comparison between the predicted result of probability integral and the measured result of GPS.

Table 3: Statistical test table of forecast results.

Direction	MAX.AE (m)	MAX.RE (%)	Mape (%)	sig	Confidence interval
Vertical	0.116	35	9	0.015	[-27,-3]
East-west	0.056	25	7	0	[-33,-27]
North-south	0.061	21	8	0	[-27,-21]

7. Conclusion

This paper proposed a method to obtain the predicted parameters of the probability integral method based on the SBAS-InSAR and GWO algorithm, which makes up for the defect of obtaining the predicted parameters of the probability integral method by the traditional method, and the deficiency of InSAR technology in the large gradient deformation of the mining area. The following conclusions are drawn:

(1) The SBAS-InSAR technology is available for the extraction of the stable boundary point sinking information of the mining area, and the verification of the accuracy of the point that coincides with the trend observation line

(2) The maximum error of the strike coincidence point B12 is 4 mm, and that of the tendency coincidence point B14 is 7 mm, indicating that the boundary information of the subsidence basin extracted by this technology is trustworthy

(3) Establish the fitness function based on the geometric relationship between the InSAR lateral imaging and the probability integral model, and obtain the predicted parameters of the probability integral method for the 112201 working face of Xiaobaodang Coal Mine in the northern Shaanxi mining area through the GWO algorithm. Besides, the obtained results include that $q = 0.63$, $b = 0.37$, $\tan\beta = 2.76$, $\theta_0 = 83.94$, $S1 = -36.34$ m, $S2 = 26.69$ m, $S3 = -45.64$ m, and $S4 = 39.62$ m. To verify its accuracy, the calculated parameters were substituted into the parameters of the probability integral model for the prediction of the sinking and horizontal movement of the working face. In addition, the predicted results were compared with the GPS measured data. It showed that the sinking error was 63 mm. The horizontal movement errors of mm, north-south, and east-west are 61 mm and 29 mm, respectively, and the accuracy met the actual production requirements. Therefore, this paper proposed a method for obtaining probability integral prediction parameters that can be applied in actual engineering

Data Availability

This study uses the Sentinel-1A data from the European Space and Space Agency, the 30-meter DEM data released by the US SRTM, and the di rock movement observation data from the working face of Shaanxi Yubei Coal Industry Co., Ltd., Xiaobaodang Coal Mine.

Conflicts of Interest

The writing of the paper was assisted by relevant personnel from Xi'an Geological Environment Monitoring Station, China Coal Geology Bureau Aerial Survey and Remote Sensing Bureau, Xi'an University of Science and Technology, and Chongqing Institute of Geology and Mineral Resources, and there is no conflict of interest between them.

Acknowledgments

The research work is supported by the National Natural Science Foundation of China (41874012).

References

[1] National Development and Reform Commission, "Eleventh five-year plan for coal industry development," *2008 China Coal Enterprise Information Management Summit Forum*, vol. 1, no. 1, pp. 1–21, 2013.

[2] Q. Zhao, "Scientific and technological innovation and efficient and clean utilization of coal," *International Economic Cooperation*, vol. 6, pp. 33–39, 2012.

[3] W. Li, L. Chen, and J. Zhao, "The damage to the ecological environment caused by coal mining in my country and the countermeasures," *Coal Mining*, vol. 50, no. 5, pp. 35–37, 2011.

[4] X. Yang, G. Hu, N. Li, and X. Yang, "Pollution control and reclamation technology of coal gangue hills," *China Mining Industry*, vol. 3, pp. 34–36, 2008.

[5] D. Huayang, W. Jinzhuang, C. Meifeng, W. Lixin, and G. Zengzhang, "Seam dip angle based mining subsidence model and its application," *International Journal of Rock Mechanics and Mining Sciences*, vol. 39, no. 1, pp. 115–123, 2002.

[6] H. Kratzsch, *Mining Subsidence Engineering*, Springer-Verlag, New York, NY, USA, 1983.

[7] G. Ren, D. Reddish, and B. Whittaker, "Mining subsidence and displacement prediction using influence function methods," *Mining Science and Technology*, vol. 5, no. 1, pp. 89–104, 1987.

[8] J. Bao and Y. Wang, "The application of GPS-RTK in subsidence monitoring of mining area," *Coal Engineering*, vol. 2, pp. 121–123, 2012.

[9] F. di Traglia, C. de Luca, M. Manzo et al., "Joint exploitation of space-borne and ground-based multitemporal InSAR measurements for volcano monitoring: the Stromboli volcano case study," *Remote Sensing of Environment*, vol. 260, article 112441, 2021.

[10] Y. Kang, Z. Lu, C. Zhao, Y. Xu, J. W. Kim, and A. J. Gallegos, "InSAR monitoring of creeping landslides in mountainous regions: a case study in Eldorado National Forest, California," *Remote Sensing of Environment*, vol. 258, p. 112400, 2021.

[11] J. Yang, H. Fan, W. Zhao, and J. Feng, "Mining area settlement monitoring and prediction based on D-InSAR technology and grey Verhulst model," *Metal Mine*, vol. 3, pp. 143–147, 2015.

[12] P. Berardino, G. Fornaro, R. Lanari, and E. Sansosti, "A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 11, pp. 2375–2383, 2002.

[13] J. Chen, L. Wang, and Y. Guo, "Mine surface movement observation based on probability integration method," *Science of Surveying and Mapping*, vol. 39, no. 3, pp. 146–148, 2014.

[14] W. Lv, H. Huang, S. Chi, and B. Han, "Neural network optimization algorithm for probability integral prediction parameters," *Science of Surveying and Mapping*, vol. 44, no. 9, pp. 35–41, 2019.

[15] Z. Zhao, T. Shi, W. Dong, and Y. Liu, "Grey relational analysis and probability integral method parameter prediction of BP neural network," *Science of Surveying and Mapping*, vol. 40, no. 7, pp. 36–40, 2017.

[16] P. Li, Z. Tan, L. Yan, and K. Deng, "Parameter calculation method of probabilistic integration method based on support vector machine," *Journal of China Coal Society*, vol. 35, no. 8, pp. 1247–1251, 2010.

[17] J. Zha, W. Feng, and X. Zhu, "Probabilistic integration method based on genetic algorithm to predict parameter inversion," *Journal of Mining and Safety Engineering*, vol. 28, no. 4, pp. 655–659, 2011.

[18] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, no. 1, pp. 46–61, 2014.

[19] J. Zhou, S. Huang, M. Wang, and Y. Qiu, "Performance evaluation of hybrid GA–SVM and GWO–SVM models to predict earthquake-induced liquefaction potential of soil: a multi-dataset investigation," *Engineering with Computers*, vol. 15, no. 11, pp. 1324–1328, 2021.

[20] A.-L. Balogun, F. Rezaie, Q. B. Pham et al., "Spatial prediction of landslide susceptibility in western Serbia using hybrid support vector regression(SVR) with GWO, BAT and COA algorithms," *Geoscience Frontiers*, vol. 12, no. 3, pp. 390–404, 2021.

[21] R. Ma, M. Karimzadeh, A. Ghabussi et al., "Assessment of composite beam performance using GWO–ELM metaheuristic algorithm," *Engineering with Computers*, vol. 13, no. 6, pp. 1513–1516, 2021.

[22] S. Lakshminarayanan, M. Abdulgader, and D. Kaur, "Scheduling energy storage unit with GWO for smart home integrated with renewable energy," *International Journal of Artificial Intelligence and Soft Computing*, vol. 7, no. 2, 2021.

[23] P. K. Keserwani, M. C. Govil, E. S. Pilli, and P. Govil, "A smart anomaly-based intrusion detection system for the Internet of Things (IoT) network using GWO–PSO–RF model," *Journal of Reliable Intelligent Environments*, vol. 7, no. 1, pp. 3–21, 2021.

[24] L. Mingyu, Y. Fang, L. Zhigang, Z. Yushuo, L. Ning, and H. Kai, "Parameter identification of wind turbine power model and evaluation of wind resource utilization rate," *Acta Solar Energy*, vol. 41, no. 12, pp. 305–315, 2020.

[25] J. Zhu, Z. Yang, and Z. Li, "Research progress of 3D surface deformation monitoring and prediction in InSAR mining area," *Journal of Surveying and Mapping*, vol. 48, no. 4, pp. 135–144, 2019.

[26] Z. Wang, "Research on 3D surface deformation calculation method based on multi-source InSAR data," *Journal of Surveying and Mapping*, vol. 48, no. 9, p. 1206, 2019.

[27] H. Fan, L. Lu, and Y. Yao, "Method combining probability integration model and a small baseline subset for time series monitoring of mining subsidence," *Remote Sensing*, vol. 10, no. 9, p. 1444, 2018.

[28] Y. Xia and Y. Wang, "InSAR-and PIM-based inclined goaf determination for illegal mining detection," *Remote Sensing*, vol. 23, no. 12, pp. 3884–3884, 2012.

[29] L. Wang, X. Zhang, S. Chi, and J. Cha, "Research on the inversion model of mining subsidence prediction parameters combining InSAR and GA," *Journal of Wuhan University (Information Science Edition)*, vol. 43, no. 11, pp. 1635–1641, 2018.

Hindawi
Journal of Sensors
Volume 2021, Article ID 5599025, 13 pages
https://doi.org/10.1155/2021/5599025

Research Article

A Novel Rapid GNSS Network Solution in Mountainous Region Monitoring considering the Tropospheric Delay at Ground Points

Guangwei Jiang,[1,2] **Panlong Wang**,[2] **Bin Wang**,[2] **and Chuanlu Cheng**[2]

[1]*College of Geology Engineering and Geomatics, Chang'an University, 126 Yanta Road, Xi'an 710054, China*
[2]*Geodetic Data Processing Centre of Ministry of Natural Resources, PRC, 334 Youyi Road, Xi'an 710054, China*

Correspondence should be addressed to Guangwei Jiang; 2017026017@chd.edu.cn

Received 7 January 2021; Revised 18 May 2021; Accepted 21 May 2021; Published 3 August 2021

Academic Editor: Sang-Hoon Hong

Due to the short peak observation time of global navigation satellite systems (GNSS), the accuracy of the tropospheric delay estimation and the positioning are poor. In this study, a rapid GNSS network solution for mountainous regions is presented. The high-precision tropospheric delay at ground points is obtained from long-term ground observation data and used as *a priori* constraint in the double-difference equation of short-time synchronous peak observations to realize rapid and high-precision positioning. Chinese mountain survey networks with large elevation gradient (1000~2000 m) were selected for the experimental verification of the proposed method. The results show that the rapid peak positioning method weakened the effect of the residual tropospheric delay caused by the elevation difference, significantly improving the accuracy and reliability of the results. The positioning accuracy of the peak in upward direction was better than 1.1 cm, which meets the requirements of rapid short-span (~1 h) high-precision monitoring and achieves 24 h positioning accuracy. Compared with the traditional solution strategy, the precision of the method with respect to the north (N), east (E), upward (U), and zenith tropospheric delay (ZTD) significantly improved. The accuracy of U improved by more than 47%. Therefore, based on the high accuracy and reliability, information of ground stations can be fully utilized to significantly reduce the peak observation time and the operation costs of surveys in mountain regions.

1. Introduction

Tropospheric delay is a major error that occurs during the processing of global navigation satellite system (GNSS) data. The refined processing of the tropospheric delay correction is the premise of precise and rapid GNSS positioning and is of great significance for rapid and high-precision location services. Due to the complex climatic conditions in high-altitude regions and the lack of measured meteorological data, traditional empirical models are mainly based on atmospheric numerical models and the calculated zenith tropospheric delay (ZTD) exhibits a deviation [1, 2]. Therefore, the GNSS double-difference network solution for peaks requires longer observation times to improve the accuracy of the tropospheric delay estimation and determine the heights of mountains. However, the harsh observation environments at mountain peaks complicate long-term GNSS observations. High elevation gradients result in large differences in the tropospheric delays at both ends of the baseline, which cannot be reduced by traditional double-difference observations to meet the centimeter-level positioning requirements. Reducing the effects of the residual tropospheric delay is of vital importance for the improvement of the accuracy of short-term peak positioning.

Surveying practices showed that the observation time in alpine regions is generally less than 4 h. When short-time GNSS peak observation data are estimated by using the traditional delayed tropospheric estimation method, the results of tropospheric parameter estimation largely deviate. This will affect the ambiguity fixing results, thus leading to a low accuracy and poor reliability of the positioning results, especially in high elevation gradient areas [3, 4]. Therefore, in the case of a high elevation gradient in the regional GNSS observation network, there is a strong correlation between spatial position parameters and tropospheric estimation parameters. The accuracy and reliability of the position results mainly

depend on the spatial and temporal accuracy of the regional tropospheric model [5]. Regarding the accuracy of the tropospheric model, Wang et al. analyzed the applicability of the common tropospheric EGNOS/UNB3m/GPT/GPT2 model in different regions and seasons in China [6]. Jian et al. proposed the integration of a ZTD with an atmospheric numerical model to address the insufficient spatiotemporal resolutions of existing tropospheric models and evaluated the model accuracy [7]. Yibin et al. evaluated the global accuracy of the Global Geodetic Observing System (GGOS) total zenith delay product and applied the corrected GGOS-ZTD product to precise point positioning (PPP) to improve the convergence speed in the height direction [1]. To weaken the effect of the tropospheric delay on the geodetic height in high elevation gradient areas, Di et al. studied the influence of the tropospheric delay estimation strategy on the baseline solution using a baseline with a large height difference. The results showed that the effect of the tropospheric delay estimation on the plane can be neglected, but the impact exceeded the uncertainty limit in the upward direction [8]. Wei studied the influence of the tropospheric delay on short baseline time series with large height differences and showed that the height direction time series exhibit false seasonal signals if the tropospheric delay is not estimated [9]. Kaifeng et al. improved the precision of the postprocessing kinematic (PPK) solution as *a priori* value by using the tropospheric zenith delay obtained from the PPP solution. With the development of multimode and multifrequency GNSS rapid positioning technology, the real-time kinematic (RTK) and PPP technology based on the continuously operating reference station (CORS) network has been developed [10]. Several scholars have proposed the improvement of the accuracy and convergence speed of PPP based on enhanced products of the regional CORS atmosphere [11, 12]. The use of the noncombined PPP method, which is based on real-time atmospheric product enhancement, significantly improves the PPP convergence time [13, 14]. These studies provided a reference for the treatment of the residual convection layer in high-altitude areas.

In most of the existing studies, long-term (>24 h) observation data were obtained for experimental objects with large height differences (~100 m). In actual surveys of mountains, such as the elevation measurement of Mount Qomolangma in China in 2005, the GNSS observation time did not exceed 2 h. There is a lack of data processing strategies for short-time double-difference network solutions in regions with large height differences. The research results show that only the effect of the large height difference of the tropospheric delay on the positioning can be addressed and the applicability of the method has several limitations. Based on the abovementioned considerations, the adoption of long-term ground observation data and the tropospheric delay as *a priori* constraints is proposed in this study to substitute simultaneous solutions in the double-difference observation equation, reduce the unknown estimable parameters of the normal equation, and improve the effect of ambiguity fixing as well as the accuracy of tropospheric delay parameters at peak and point coordinates. To investigate the feasibility and reliability of the method proposed in this paper, network experiments were carried out in different mountainous areas in the east, middle, and south of China to compare and analyze the positioning accuracy and advantages and disadvantages of different tropospheric treatment strategies.

2. Tropospheric Delay in Differential Positioning

During the actual positioning process, the effect of the tropospheric delay is similar to that of the ionospheric delay. The tropospheric delay is caused by the neutral atmosphere below 10 km, and the ionospheric delay is due to the propagation medium at a height of ~400 km [15]. Therefore, the tropospheric delay is more susceptible to climate change. When using the double-difference model for positioning, the effect of the tropospheric delay is generally eliminated or weakened based on parameter estimation [16–18]. Currently, three main methods are used for the correction of the tropospheric delay: (1) direct estimation using an empirical model, such as Saastamoinen based on measured meteorological parameters, and the GPT series and IGGtrop series based on the atmospheric numerical model; (2) parameter estimation in which the tropospheric wet delay is used as a parameter; and (3) the high-precision tropospheric delay is directly obtained from sounding data and regarded to be the true value. When the positioning is based on the double-difference model, the effect of the tropospheric delay is generally eliminated or weakened based on parameter estimation [16–18].

Different positioning technologies can be used to reduce the effect of the tropospheric delay error, but when the baseline distance or the height difference are large, it is difficult to eliminate or weaken the tropospheric delay error. The differential residual tropospheric delay affects the accuracy of the baseline solution and thus the solution of the integer ambiguity [19]. In high-precision GNSS positioning, the tropospheric delay is the largest error, followed by the fixed ambiguity. The tropospheric delay can be generally divided into relative and absolute tropospheric delays [20]. The relative tropospheric delay mainly affects the accuracy of the station height estimate, whereas the absolute tropospheric delay mainly causes the deviation of the baseline scale [21].

2.1. Relative Tropospheric Delay. The relative tropospheric delay is the difference in the tropospheric delay between the base stations at both ends of the baseline. Relative tropospheric delay is one of the key factors affecting the height of a station in alpine surveying scenarios, but it is difficult to correct this error using a model. The height errors can be determined as follows:

$$\Delta h = \frac{\Delta \mathrm{ZPD}}{\sin E_{\min}}, \tag{1}$$

where ΔZPD (zenith path delay) is the relative tropospheric delay and $E_{\min}$ is the cutoff altitude angle of the satellite.

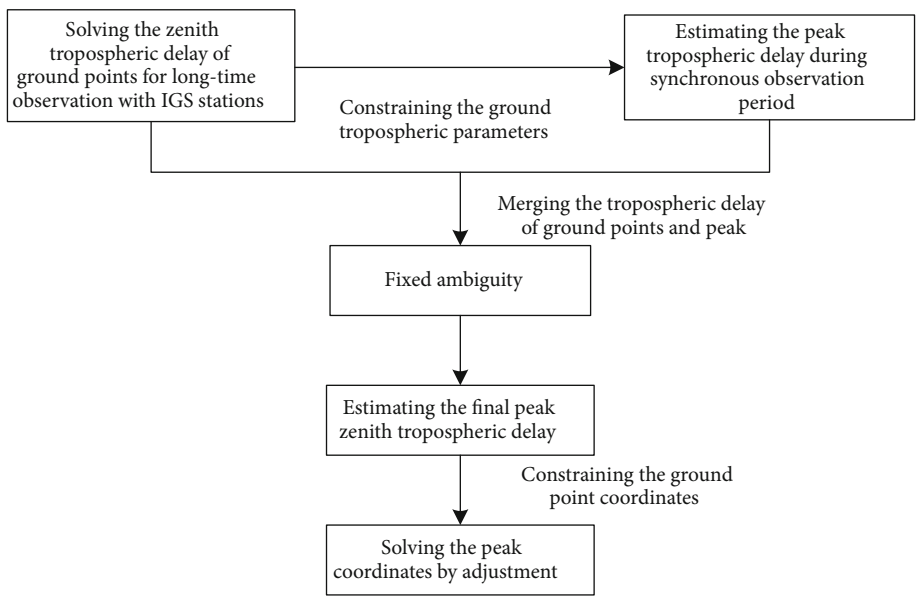


Figure 1: Flowchart of rapid peak positioning based on the *a priori* tropospheric delay at ground points.

Based on Equation (1), the effect of the relative zenith tropospheric delay ΔZTD on the height can be calculated as follows:

$$\Delta h = \frac{\Delta \mathrm{ZPD}}{\sin E_{\min}} = \frac{\Delta \mathrm{ZTD}/\sin E_{\min}}{\sin E_{\min}} = \frac{\Delta \mathrm{ZTD}}{\sin^2 E_{\min}}. \quad (2)$$

Based on this equation, a ZTD of 1 mm causes a height error of 3.5 cm when the cutoff altitude angle of the satellite is 10°. When the cutoff altitude angle of the satellite is 5°, a ZTD of 1 mm causes a height error of 13.2 cm.

2.2. Absolute Tropospheric Delay. The absolute tropospheric delay refers to the same value of tropospheric delay at both ends of the baseline, which mainly affects the scale factor of the baseline. The absolute tropospheric delay can be used to obtain a high-precision zenith delay value by nondifferential precise single-point positioning and a double-difference network solution. The following equation can be used to estimate the scale effect:

$$\frac{\Delta l}{l} = \frac{\mathrm{ZTD}}{R_e \sin E_{\min}}, \quad (3)$$

where l and Δl are the baseline length and deviation, respectively; ZTD is the absolute zenith tropospheric delay, and R_e is the Earth's radius. When the cutoff angle of the satellite altitude is 10°, the scale error caused by an absolute tropospheric delay of 2 m is ~2 ppm. When the cutoff angle of the satellite altitude is 5°, the scale error caused by a 2 m absolute tropospheric delay is ~4 ppm.

3. Double-Difference Function Model considering Delay Constraints in the Troposphere at Ground Points

When the GNSS synchronous observation network is solved with a double-difference baseline, the double-difference equation can be expressed as

$$\begin{cases} P_{1\mathrm{rs}}^{ij} = \rho_{rs}^{ij} + I_{rs}^{ij} + T_{rs}^{ij}, \\ P_{2\mathrm{rs}}^{ij} = \rho_{rs}^{ij} + \dfrac{f_1^2}{f_2^2} I_{rs}^{ij} + T_{rs}^{ij}, \\ \varphi_{1\mathrm{rs}}^{ij} = \rho_{rs}^{ij} - I_{rs}^{ij} + T_{rs}^{ij} + \lambda_1 n_{1rs}^{ij}, \\ \varphi_{2\mathrm{rs}}^{ij} = \rho_{rs}^{ij} - \dfrac{f_1^2}{f_2^2} I_{rs}^{ij} + T_{rs}^{ij} + \lambda_2 n_{2rs}^{ij}. \end{cases} \quad (4)$$

When a nonionospheric combination is used, the double-difference observation equation can be expressed as

$$\begin{cases} \varphi_{\mathrm{IF,rs}}^{ij} = \rho_{rs}^{ij} + \lambda_1 b_{\mathrm{IF,rs}}^{ij} + T_{rs}^{ij}, \\ P_{\mathrm{IF,rs}}^{ij} = \rho_{rs}^{ij} + T_{rs}^{ij}, \end{cases} \quad (5)$$

where the superscripts i and j and the subscripts r and s represent the given receiver and satellite, respectively; ρ_{rs}^{ij} is the geometric distance of the station; I_{rs}^{ij} is the double-difference ionospheric delay; T_{rs}^{ij} is the double-difference tropospheric delay; $\varphi_{\mathrm{IF,rs}}^{ij}$ and $P_{\mathrm{IF,rs}}^{ij}$ are between stations r and s of the observed satellite, which has a nonionospheric combination carrier and pseudorange double-difference observation values; $b_{\mathrm{IF,rs}}^{ij}$ is the double-difference nonionospheric combination ambiguity; and λ_1 is the wavelength at the frequency L_1.

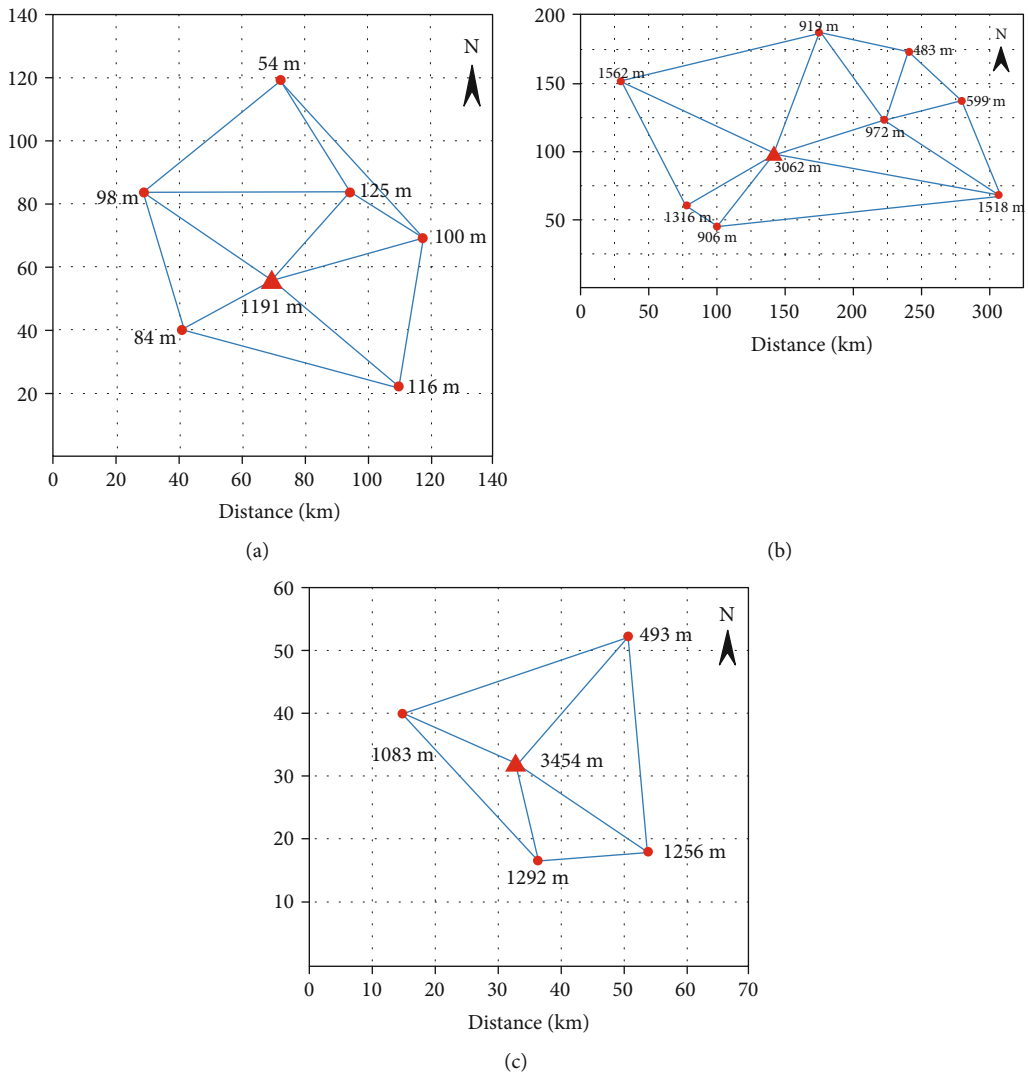


Figure 2: Distribution of the GNSS observation network in different areas. Values next to the red markers indicate elevation. (a) Zhejiang. (b) Yunnan. (c) Shaanxi.

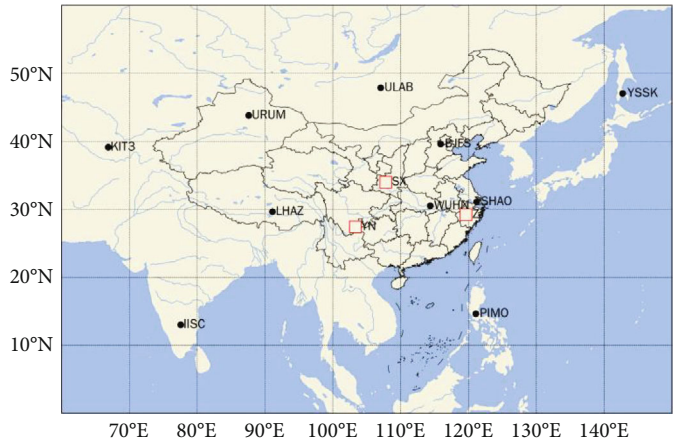


Figure 3: Distribution of IGS stations.

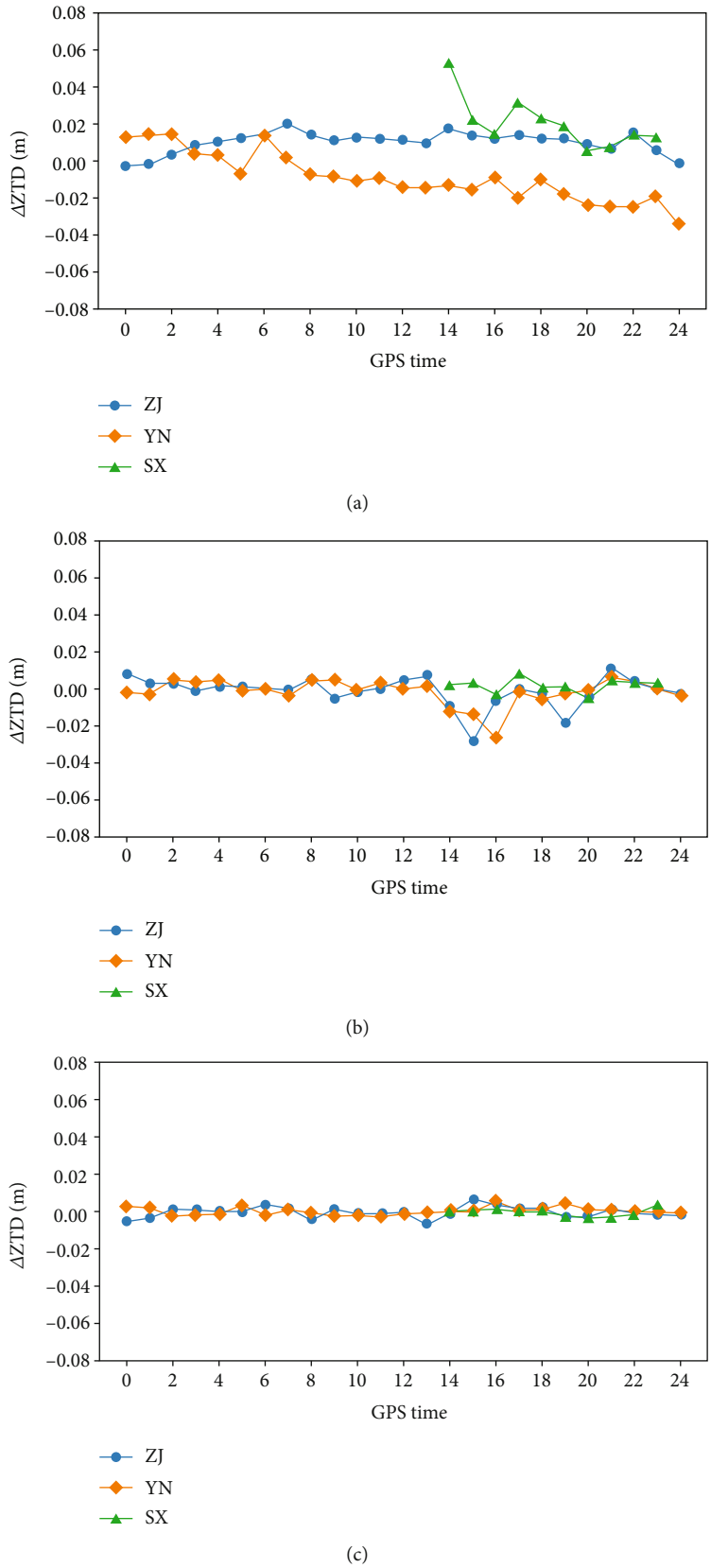


Figure 4: Deviation of the relative zenith tropospheric delay. (a) Scheme 1. (b) Scheme 2. (c) Scheme 3.

Index	N			E			U			ZTD		
	ZJ	YN	SX	ZJ	YN	SX	ZJ	YN	SX	ZJ	YN	SX
MAX	0.017	0.012	0.010	0.043	0.038	0.054	0.054	0.037	0.079	0.019	0.016	0.017
MEAN	0.004	0.002	0.005	0.009	0.005	0.010	0.020	0.019	0.053	0.008	0.006	0.006
RMS	0.006	0.004	0.006	0.016	0.010	0.020	0.025	0.022	0.059	0.007	0.008	0.008

Scheme 1: The Saastamoinen+Vienna mapping functions 1 (VMF1) model value is used for the correction of the tropospheric delay.

Index	N			E			U			ZTD		
	ZJ	YN	SX	ZJ	YN	SX	ZJ	YN	SX	ZJ	YN	SX
MAX	0.018	0.005	0.016	0.053	0.019	0.021	0.049	0.071	0.052	0.361	0.211	0.794
MEAN	0.003	0.002	0.005	0.007	0.005	0.005	0.013	0.014	0.013	0.076	0.056	0.213
RMS	0.005	0.003	0.007	0.015	0.007	0.008	0.018	0.020	0.021	0.118	0.086	0.318

Scheme 2: All stations carry out tropospheric delay estimation. The dry component of the Saastamoinen+Vienna mapping functions 1 (VMF1) model is used as the priori value. The wet component of the tropospheric delay is estimated at all stations. The tropospheric parameters are linearly modelled, so that each period includes the number of the station parameters for the tropospheric delay.

Index	N			E			U			ZTD		
	ZJ	YN	SX	ZJ	YN	SX	ZJ	YN	SX	ZJ	YN	SX
MAX	0.021	0.010	0.006	0.017	0.013	0.006	0.045	0.021	0.018	0.007	0.005	0.004
MEAN	0.002	0.002	0.003	0.003	0.003	0.003	0.013	0.009	0.009	0.002	0.002	0.002
RMS	0.004	0.003	0.003	0.004	0.004	0.004	0.010	0.011	0.010	0.002	0.002	0.002

Scheme 3: The corrections of the tropospheric delay and positioning results obtained from long-term observations of ground points are used as *a priori* constraints, and 1 h sectional solution is obtained for the synchronous peak observation period to estimate the peak tropospheric delay and position of each period.

Based on the assumption of the parameters of the ZTD at the ground points, the stochastic model uses the long-term observed ZTD of the ground points as the virtual observation value and the observation equation can be expressed as

$$\mathrm{VZTD}_r = \mathrm{ZTD}_r + \varepsilon_{r,\mathrm{VZTD}}, \tag{6}$$

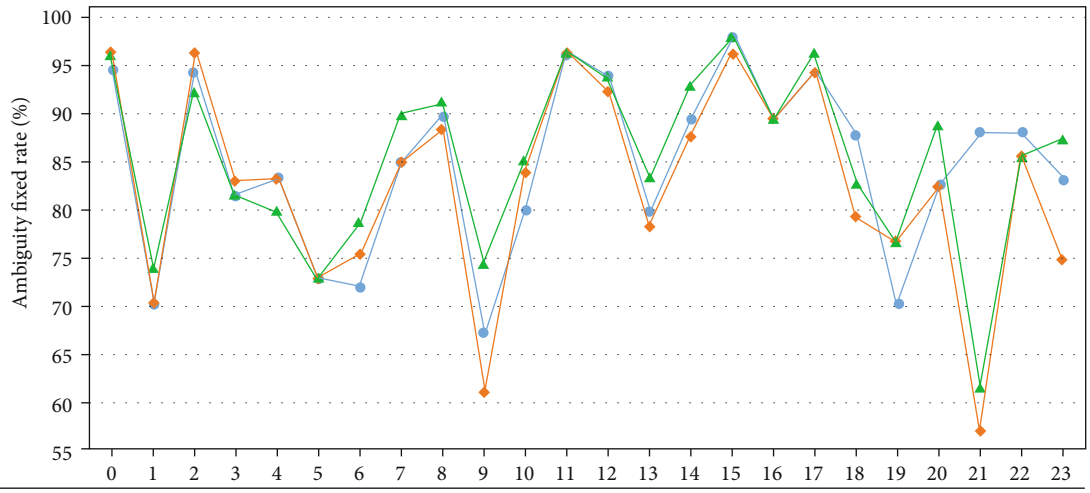
where $\varepsilon_{r,\mathrm{VZTD}}$ is the noise of the virtual ZTD. High-precision ground tropospheric delay correction can be achieved due to the ease of long-term GNSS observations in the terrestrial point environment. Equations (5) and (6) can be used to obtain the peak-to-peak tropospheric parameters based on high-precision *a priori* constraints of the troposphere at the ground points.

In this study, the high-precision positioning software BERNESE was used to calculate GNSS observation data at ground points to obtain high-precision positioning results and the tropospheric delay. Based on the use of high-precision *a priori* tropospheric information and ground coordinates as constraints, the double-difference equation can be solved and the number of evaluation parameters can be effectively reduced. The ill-conditioned normal equation can be improved, which enhances the accuracy and reliability of the tropospheric delay, and a peak positioning result is obtained. The solution flowchart is shown in Figure 1.

4. Methods

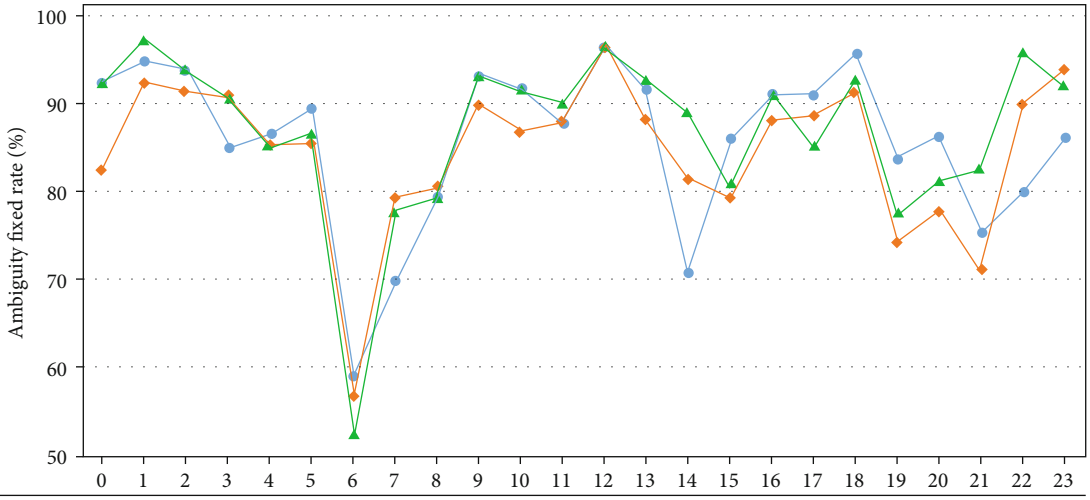
Surveying practice has shown that a short baseline must consider the effect of tropospheric delay error in high elevation gradient positioning solutions in different regions. In this study, high-precision ground tropospheric correction values were combined with the tropospheric delay of the peak estimation. First, the amount and precision of the tropospheric delay correction calculated from long-tern GNSS observation data for ground points were used as *a priori* constraints in the double-difference observation equation to estimate unknown parameters. Because of the high accuracy of the troposphere parameters solved over a long period of time, as a priori tropospheric constraint, the number of parameters to be determined can be reduced and the morbidity of the normal equation can be improved, thus improving the accuracy and reliability of the tropospheric delay and peak positioning results.

4.1. Experiments in Mountainous Areas. Considering that the GNSS double-difference calculation method is affected by the baseline length and geographical location of the observation network, three mountains in China were selected for experiments to verify the proposed method. The mountains are located in the Zhejiang (ZJ), Yunnan (YN), and Shaanxi (SX) provinces, respectively. The distribution of the stations is shown in Figure 2. The maximum height difference



Scheme 1	94.5	70.4	94.4	81.5	83.3	73.1	72.1	85.0	89.9	67.2	80.0	96.4	93.9	80.0	89.5	98.1	89.4	94.4	87.9	70.2	82.6	88.1	88.1	83.3
Scheme 2	96.4	70.4	96.3	83.1	83.3	73.1	75.4	85.0	88.4	61.2	84.0	96.4	92.4	78.3	87.7	96.3	89.4	94.4	79.3	76.6	82.6	57.1	85.7	75.0
Scheme 3	96.4	74.1	92.6	81.5	80.0	73.1	78.7	90.0	91.3	74.6	85.3	96.4	93.9	83.3	93.0	98.1	89.4	96.3	82.8	76.6	89.1	61.9	85.7	87.5

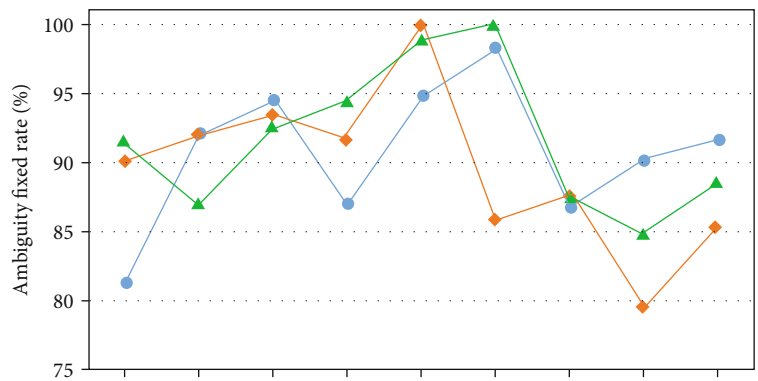
(a)



Scheme 1	92.5	95.0	93.9	85.1	86.6	89.5	59.1	69.8	79.3	93.3	91.7	87.9	96.6	91.7	70.8	86.1	91.2	91.2	95.8	83.9	86.4	75.4	80.0	86.3
Scheme 2	82.5	92.5	91.5	90.8	85.4	85.6	56.8	79.4	80.5	90.0	86.9	87.9	96.6	88.1	81.5	79.2	88.2	88.8	91.5	74.2	78.0	71.0	90.0	94.1
Scheme 3	92.5	97.5	93.9	90.8	85.4	86.8	52.3	77.8	79.3	93.3	91.7	90.1	96.6	92.9	89.2	80.6	91.2	85.0	93.0	77.4	81.4	82.6	96.0	92.2

(b)

Figure 5: Continued.



	14	15	16	17	18	19	20	21	22
Scheme 1	81.3	92.1	94.5	86.9	94.8	98.2	86.5	90.2	91.7
Scheme 2	90.1	92.1	93.5	91.7	100.0	85.7	87.6	79.5	85.4
Scheme 3	91.5	86.9	92.6	94.5	98.9	100.0	87.5	84.8	88.5

(c)

Figure 5: Success rate of ambiguity fixing: (a) ZJ, (b) YN, and (c) SX.

Table 1: Average ambiguity fixing rate (%).

Mean	ZJ	YN	SX
Scheme 1	84.7	85.8	90.7
Scheme 2	82.8	84.6	89.5
Scheme 3	85.5	87.1	91.7

between the peak and ground point is 1137, 2156, and 2961 m, respectively. Observations were carried out in the Zhejiang and Yunnan provinces for 24 h on March 7 and 1, 2019, respectively. The observation in the Shaanxi Province was carried out on July 31, 2006, for 9 h from 14:00 to 23:00 GPS time. The GNSS dual-frequency receivers were set up between the peak and ground point for continuous observations. The ground points were evenly distributed around the peak. The cutoff angle of the satellite was set to 10°, and the data sampling interval was 30 s.

4.2. Tropospheric Delay Correction and Peak Positioning Results. First, Chinese International GNSS Service (IGS) stations were selected to calculate the ground coordinates under the ITRF2014 framework and the 1 h ZTD using Bernese software. The distribution of the stations is shown in Figure 3. Subsequently, the coordinates of the ground points were used as starting points and positioning results were obtained for the long-term peak observations and ZTD and used as reference values. Finally, the peak data were divided based on 1 h intervals and the results of the short-term mountain network solution were compared and analyzed using the following three tropospheric delay parameter processing schemes.

To obtain a high-precision prior tropospheric delay of ground points, IGS stations around China were used as starting points to calculate the ground coordinates under the ITRF2014 framework and the 1 h ZTD using the Bernese software. The initial tropospheric delay was the dry component of the Saastamoinen model, and the wet component of the tropospheric delay was estimated at a 1 h interval via a piecewise linear method. The tropospheric mapping function adopted VMF1, while the tropospheric delay gradient parameter was not estimated. According to statistics, the ZTD error of ground points was less than 1.2 mm in any period.

To compare the spatiotemporal characteristics of the deviation of the relative ZTD based on the three schemes, three baselines with the largest height differences in different regions were selected and the calculated long-term relative ZTD of the baseline and peak positioning results were used as reference values. The values estimated using the three schemes were compared with the reference values:

$$\begin{cases} \text{bias} = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N}\left(X_i - X_i^{\text{ref}}\right), \\ \text{RMS} = \sqrt{\dfrac{1}{N}\displaystyle\sum_{i=1}^{N}\left(X_i - X_i^{\text{ref}}\right)^2}, \end{cases} \tag{7}$$

where X_i are the positioning and tropospheric zenith delay parameters estimated for each period of different schemes; X_i^{ref} are the reference values for the positioning and tropospheric zenith delay parameters estimated for each period of time; N is the total number of time periods used for the statistics; bias is the average deviation, which reflects the average deviation of different schemes; and RMS is the root mean square error, which reflects the accuracy of different schemes relative to the reference value and the stability of the parameters.

Based on the use of the abovementioned three schemes to locate hilltops with different elevation gradient, the differences between the hilltop positioning results and reference values of the ZTD were compared.

Figure 4 shows that the zenith obtained with Scheme 1 has the largest relative tropospheric delay and a notable

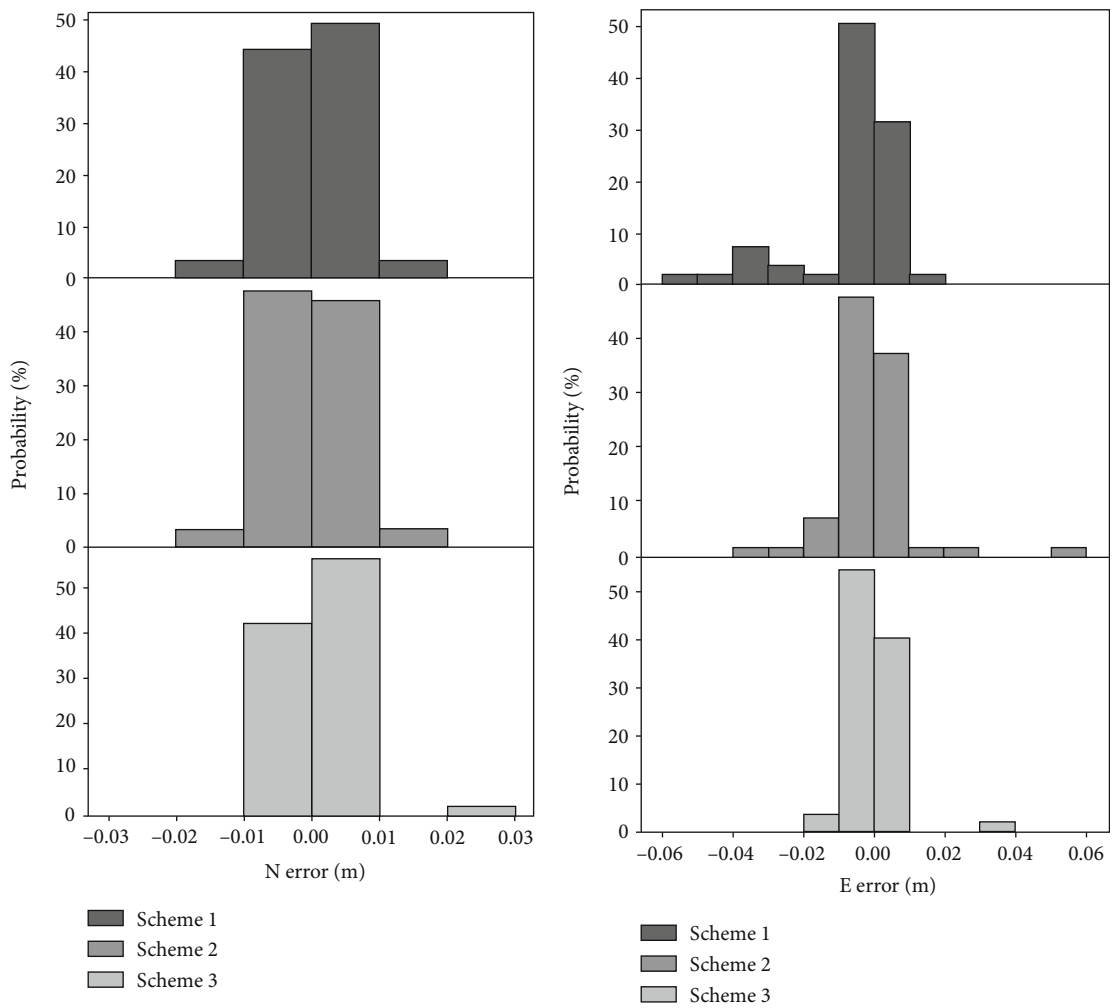
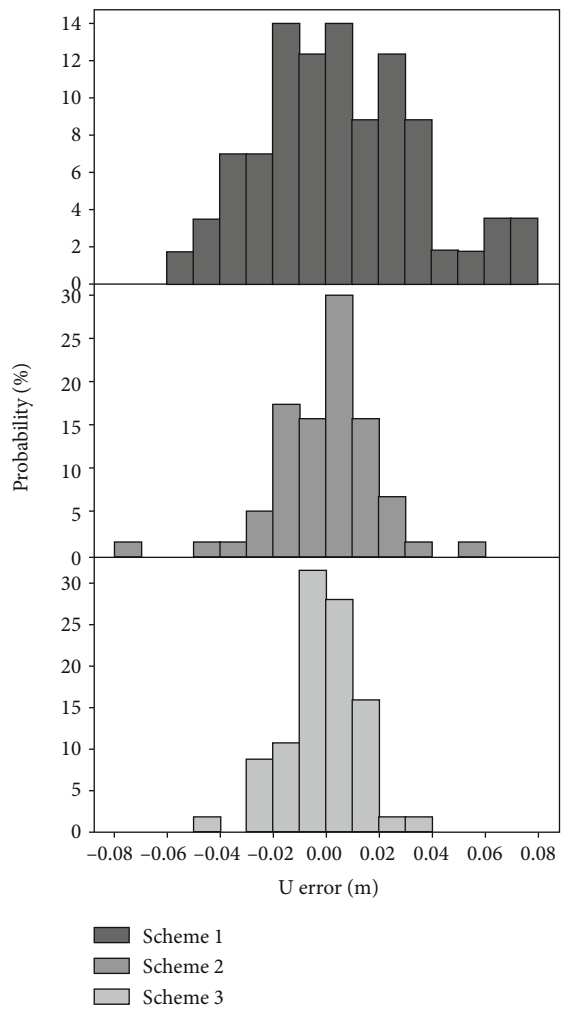


Figure 6: Continued.



(c)

Figure 6: Probability statistics for the positioning error. (a) North. (b) East. (c) Upward.

systematic deviation. The deviation increases with the height difference, and the maximum deviation is 5.38 cm. This indicates that the direct use of the tropospheric model cannot effectively weaken the effect of the tropospheric delay at baselines with high elevation gradient. The zenith troposphere delay based on Scheme 2 is large in several periods, with a maximum deviation of 2.85 cm. On the one hand, uncertainty exists in short-term tropospheric estimations. On the other hand, the correlation between tropospheric parameters and the height difference of the baseline is strong and reliable, and the estimated parameters cannot be correctly calculated in a short time. In Scheme 3, the correlation between the deviation of the zenith relative tropospheric delay and the height difference between the baseline sections is weak. The deviation of the tropospheric delay has the smallest error and the highest stability, and the deviation of the zenith relative troposphere delay is less than 1 cm, which further illustrates that Scheme 3 effectively weakens the effect of the residual tropospheric delay error of baselines with high elevation gradient.

According to the results of the tropospheric delay by the three processing strategies, the baseline ambiguity parameters were fixed. The average success rate of ambiguity fixing was calculated in each period (Figure 5), while the average success rate of ambiguity fixing of all periods is shown in Table 1.

Figure 5 shows that the success rate of ambiguity fixing was influenced by the tropospheric delay, while the impact was small. Comparing the three schemes, the success rate of ambiguity fixing in Scheme 3 was the highest, whereas that in Scheme 2 was the lowest.

The probability statistics for the error of the mountain peak positioning results of the three schemes and the effects of the three schemes on the positioning results were compared. The statistical results are shown in Figure 6.

Figure 6 shows that the error of the planar orientation is smaller than that of the U direction, and the error in the N direction is the smallest. Scheme 1 exhibits the worst stability with respect to the positioning results and has little influence on the positioning results in the N and E directions. The probability of a period with large error is small. The probability of a period with a U direction error below 2 cm is 49.1%, and the reliability of the positioning results is poor. Scheme 2 exhibits a higher stability than Scheme 1 and a

Table 2: Error of the solution and zenith tropospheric delay (m).

(a) Scheme 1

Index	N			E			U			ZTD		
	ZJ	YN	SX	ZJ	YN	SX	ZJ	YN	SX	ZJ	YN	SX
Max	0.017	0.012	0.010	0.043	0.038	0.054	0.054	0.037	0.079	0.019	0.016	0.017
Mean	0.004	0.002	0.005	0.009	0.005	0.010	0.020	0.019	0.053	0.008	0.006	0.006
RMS	0.006	0.004	0.006	0.016	0.010	0.020	0.025	0.022	0.059	0.007	0.008	0.008

(b) Scheme 2

Index	N			E			U			ZTD		
	ZJ	YN	SX	ZJ	YN	SX	ZJ	YN	SX	ZJ	YN	SX
Max	0.018	0.005	0.016	0.053	0.019	0.021	0.049	0.071	0.052	0.361	0.211	0.794
Mean	0.003	0.002	0.005	0.007	0.005	0.005	0.013	0.014	0.013	0.076	0.056	0.213
RMS	0.005	0.003	0.007	0.015	0.007	0.008	0.018	0.020	0.021	0.118	0.086	0.318

(c) Scheme 3

Index	N			E			U			ZTD		
	ZJ	YN	SX	ZJ	YN	SX	ZJ	YN	SX	ZJ	YN	SX
Max	0.021	0.010	0.006	0.017	0.013	0.006	0.045	0.021	0.018	0.007	0.005	0.004
Mean	0.002	0.002	0.003	0.003	0.003	0.003	0.013	0.009	0.009	0.002	0.002	0.002
RMS	0.004	0.003	0.003	0.004	0.004	0.004	0.010	0.011	0.010	0.002	0.002	0.002

Table 3: Statistical results for the average accuracy of the estimated parameters (m).

Index	Mean RMS (m)			Proportion of increase in Scheme 3/%	
Scheme	1	2	3	Compared with Scheme 1	Compared with Scheme 2
N	0.005	0.005	0.003	38	33
E	0.015	0.010	0.004	74	60
U	0.035	0.020	0.010	71	47
ZTD	0.008	0.174	0.002	74	99

larger positioning error. Scheme 3 yields the most stable positioning results. The probability of a large-error period is small, the N direction error of all periods is less than 1 cm, the E direction error in more than 96.4% of the period is less than 1 cm, and the U direction error in more than 86.0% of the period is less than 2 cm.

To further analyze the effects of different schemes on the positioning results, the peak positioning results and errors of the ZTD correction were compared. The statistical results are shown in Table 2.

Based on the statistical tables, the following conclusions can be drawn:

(1) Based on the comparison of Tables 2(a)–2(c), the plane solutions of the three schemes are better than those of the U direction. Scheme 3 exhibits the highest positioning accuracy and reliability, which are less affected by the elevation difference. The average accuracy of N and E is better than 0.4 cm, the accuracy of the U direction is better than 1.1 cm, and the average accuracy is 1.0 cm

(2) Scheme 1 directly reflects the accuracy of the tropospheric delay model. Based on the correction of the tropospheric zenith delay, it can be concluded that the accuracy of the Saastamoinen+VMF1 model reached the mm level in different regions. With increasing height difference, the accuracy of the U direction of the scheme decreases. When the height difference reaches 3000 m, the accuracy of U direction positioning is 5.9 cm, which indicates that the precision of the U direction positioning is 5.9 cm in the high-drop area and the height difference is the main factor affecting the accuracy of height positioning. The tropospheric model contains large errors and delays in the relative troposphere at both ends of the baseline

(3) The estimation of the tropospheric zenith delay parameters based on Scheme 2 exhibits a poor reliability and accuracy. However, the height results are significantly better than those derived from

Scheme 1 because the absolute tropospheric delay error affects the baseline scale and has little influence on the altitude positioning results. The maximum deviation of the positioning results is 1000–2000 m because the relative tropospheric delay error in the corresponding period is large

(4) Based on Scheme 3, the dimension of the observed equation is reduced, and the effective parameters are estimated using *a priori* tropospheric parameters as constraints. The estimated peak-to-peak ZTD accuracy reaches the mm level, which is better than that of Scheme 2. At different elevations in different areas, the average accuracy of elevation positioning is better than 1.0 cm. The weak correlation between the regional height difference and elevation positioning accuracy verifies the validity and reliability of this method. Based on the application of this method, the effect of the altitude difference on the accuracy of elevation positioning is effectively eliminated. Thus, this method is suitable for short-term, high-accuracy, and rapid positioning of GNSS control networks with large height differences

(5) Based on the statistical results obtained for the positioning accuracy of several schemes Table 3, the accuracy of Scheme 3 greatly improved with respect to N, E, U, and ZTD, especially that of U and ZTD. The accuracy in the U direction improved by more than 47%. The accuracies of Scheme 3 are 38%, 74%, 71%, and 74% higher than those in Schemes 1 and 33%, 60%, 47%, and 99% higher than those in Scheme 2. Overall, Scheme 3 is the most suitable for rapid positioning in special environments, such as mountainous areas

5. Conclusions

In this study, the methods and model accuracy of short-term measurements of the tropospheric delay correction with a high elevation gradient were analyzed. In view of the difficulty in estimating the tropospheric delay by using short-term double-difference networks in mountainous areas with large height differences, a high-precision correction of the tropospheric delay based on long-term ground observations is proposed, which can be used as *a priori* constraint and substituted into a double-difference solution model to estimate the peak zenith delay and coordinates. The proposed method was used for an experimental analysis using traditional tropospheric delay estimation and Saastamoinen +VMF1 tropospheric model correction. For both the tropospheric delay and the success rate of ambiguity fixing, the proposed method was most optimal among the three schemes. The results show that the positioning accuracy of this method is better in the N, E, and U directions compared with that of traditional methods. The accuracy in the U direction improved by more than 47%.

In mountainous areas, the rapid GNSS networking method based on tropospheric *a priori* information effectively solves problems regarding the poor stability and positioning accuracy of the geodetic peak height caused by inconsistent tropospheric delays at both ends of the short-term mountain baseline. The accuracy of the positioning results in extreme observation environments improved. This method can be applied to short-term high-accuracy and rapid positioning with large height differences in different areas. In the future, the method must be further improved. The application field should be expanded to realize rapid double-difference positioning based on regional CORS to reduce the peak observation time and to improve the efficiency of positioning operations in mountainous areas.

Data Availability

The datasets used or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare no conflict of interest.

Authors' Contributions

Conceptualization was done by G.J. and P.W. Methodology was prepared by P.W. Software was acquired by P.W. Validation was performed by G.J. and P.W. Formal analysis was performed by B.W. and C.C. Resources were acquired by P.W. Supervision was done by G.J. Writing and preparation of original draft was done by G.J. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (41574003, 41904040) and the Project for High-level-innovation Talents in Science and Technology, Ministry of Natural Resources (12110600000018003926).

References

[1] Y. Yibin, X. U. Xingyu, and H. Yufeng, "Precision analysis of GGOS tropospheric delay product and its application in PPP," *Acta Geodaetica et Cartographica Sinica*, vol. 46, no. 3, pp. 278–287, 2017.

[2] Y. Xiang, C. Mingjian, W. Jianguang, and C. Rui, "Adaptability analysis of GPT2w model in high latitudes," *Chinese Journal of Space Science*, vol. 40, no. 2, pp. 242–249, 2020.

[3] A. Xiangdong and Y. Dengke, "The impact of the height difference between stations on the baselines solution of short period GPS observations," *Journal of Geodesy and Geodynamics*, vol. 36, no. 6, pp. 534–538, 2016.

[4] Z. Jingyang and S. Shuangshuang, "Research progress of zenith tropospheric delay model and its accuracy analysis over China," *Progress in Geophysics*, vol. 33, no. 1, pp. 148–155, 2018.

[5] J. Junru, T. Tingye, and G. Fei, "A method for increasing precision of short baseline with large height difference by using semi-parametric model," *Journal of Geodesy and Geodynamics*, vol. 36, no. 4, pp. 319–322, 2016.

[6] J. G. Wang, J. P. Chen, J. X. Wang, J. J. Zhang, and L. Song, "Assessment of tropospheric delay correction models over China," *Geomatics and Information Science of Wuhan University*, vol. 41, no. 12, pp. 1656–1663, 2016.

[7] M. Jian, C. Tiejun, L. Xiaoli et al., "A high-accuracy method for tropospheric delay error correction by fusing atmospheric numerical models," *Acta Geodaetica et Cartographica Sinica*, vol. 48, no. 7, pp. 862–870, 2019.

[8] Z. Di, G. Jiming, C. Xuefeng, and Q. Fachao, "Influence of estimation of tropospheric delay on short baseline with big height difference," *Journal of Geodesy and Geodynamic*, vol. 34, no. 2, pp. 146–149, 2014.

[9] W. Wei, "Analysis on influence of tropospheric delay on short baseline time series from GPS stations with the large height difference," *Journal of Geodesy and Geodynamics*, vol. 38, no. 5, pp. 504–509, 2018.

[10] L. Kaifeng, O. Yongzhong, L. Xiuping, and W. Taiqi, "A differential estimation technique of troposphere delay for precise positioning in hydrographic surveying," *Geomatics and Information Science of Wuhan University*, vol. 38, no. 8, pp. 930–934, 2013.

[11] J. Geng, F. N. Teferle, X. Meng, and A. H. Dodson, "Towards PPP-RTK: ambiguity resolution in real-time precise point positioning," *Advances in Space Research*, vol. 47, no. 10, pp. 1664–1673, 2011.

[12] Y. Shoji, "A study of near real-time water vapor analysis using a nationwide dense GPS network of Japan," *Journal of the Meteorological Society of Japan*, vol. 87, no. 1, pp. 1–18, 2009.

[13] Y. Yibin, F. Xinying, P. Wenjie, and L. Lei, "Local atmosphere augmentation based on CORS for real-time PPP," *Geomatics and Information Science of Wuhan University*, vol. 44, no. 12, pp. 1739–1748, 2019.

[14] X. Li, X. Zhang, and M. Ge, "Regional reference network augmented precise point positioning for instantaneous ambiguity resolution," *Journal of Geodesy*, vol. 85, no. 3, pp. 151–158, 2011.

[15] D. Xiaoguang, *Application and Research of Tropospheric Delay Correction in GPS Data Processing*, Chang' an University, 2009.

[16] J. Junru, T. Tingye, and W. Guyue, "A method for increasing precision of single epoch short baseline solution with large height difference by using semi-parametric model," *Science of Surveying and Mapping*, vol. 42, no. 6, pp. 24–30, 2017.

[17] Z. Guoli, Y. Kaiwei, S. Xiaofei, and S. Chuanzhen, "Accuracy impact analysis of tropospheric correction model in double difference RTK solution," *Bulletin of Surveying and Mapping*, vol. 9, pp. 149-150, 2016.

[18] D. Wujiao, C. Zhaohua, and L. Ming, "Effect of height difference on GPS vertical accuracy," *Journal of Geodesy and Geodynamics*, vol. 29, no. 3, pp. 80–83 + 87, 2009.

[19] L. Ning, Z. Yongzhi, and X. Yongliang, "A three-step Kalman filter algorithm for near real-time estimating tropospheric wet delay on GPS reference stations," *Geomatics and Information Science of Wuhan University*, vol. 40, no. 7, pp. 918–923, 2015.

[20] G. Beutler, I. Bauersima, and W. Gurtner, *Atmospheric Refraction and Other Important Biases in GPS Carrier Phase Observations, Monograph12*, School of Surveying, University of New South Wales, 1988.

[21] M. Rothacher and G. Beutler, "The role of GPS in the study of global change," *Physics and Chemistry of the Earth*, vol. 23, no. 9–10, pp. 1029–1040, 1998.

Hindawi
Journal of Sensors
Volume 2021, Article ID 9986430, 13 pages
https://doi.org/10.1155/2021/9986430

Research Article

Hybrid Niche Immune Genetic Algorithm for Fault Detection Coverage in Industry Wireless Sensor Network

Jie Zhou, Hu Qin, Yang Liu, Chaoqun Li, and Mengying Xu

College of Information Science and Technology, Shihezi University, Shihezi 832000, China

Correspondence should be addressed to Jie Zhou; jiezhou@shzu.edu.cn

Received 20 March 2021; Revised 27 April 2021; Accepted 29 May 2021; Published 16 June 2021

Academic Editor: Hyung-Sup Jung

The industry wireless sensor network (IWSN) technology, which is used to monitor industrial equipment, has attracted more and more attention in recent years. Sensor nodes in IWSN can spontaneously complete distributed networking and carry out monitoring tasks under random deployment conditions. Therefore, a self-organized IWSN is particularly suitable for the fault detection and diagnosis of industrial equipment in complex environments. However, due to the detection, ability of a single sensor node is limited, and the monitoring distribution problem is a typical multidimensional discrete NP-hard combinatorial stochastic optimization problem, which is challenging to solve for the traditional mathematical methods. With the purpose of improving the target monitoring capability and prolonging lifetime of IWSN, a novel hybrid niche immune genetic algorithm (HNIGA) for optimizing the target coverage model of fault detection is proposed. It uses the genetic operation to evolve antibody groups and applies niche technology to maintain the diversity of antibody groups. As a result, HNIGA can effectively reduce the failure rate of detection targets. To verify the performance of HNIGA, a series of simulations under different simulation conditions are carried out. Specifically, HNIGA is compared with genetic algorithm (GA) and simulated annealing (SA). Simulation results show that HNIGA has a faster convergence speed and more robust global search capability than the other two algorithms.

1. Introduction

With the rapid development of the manufacturing industry, machine fault diagnosis plays an increasingly important role in ensuring the safe operation of equipment [1–3]. Unexpected equipment failures will cause severe damage to the equipment and cause more economic losses due to the interruption of plant operation. Therefore, a lot of fault diagnosis research works are proposed to focus on effective maintenance procedures.

Industrial wireless sensor network (IWSN) provides an effective scheme for machine fault diagnosis [4]. More specifically, IWSN is a distributed wireless network that transmits node information to users by monitoring, sensing, and processing node data through multiple sensor nodes [5–7]. Besides, various microsensors are integrated on each node, effectively collecting and monitoring different physical data such as temperature, humidity, acoustic parameters, and optical parameters. These pieces of information are transmitted to the sink node using wireless links in a multihop self-organizing manner. Then, they are sent to the monitoring center through GPRS, satellite, microwave communication, or the Internet. Subsequently, users can collect the information of interest or further control other nodes' actions on the network through the aggregation node. IWSN is usually deployed in harsh environments without infrastructure and unattended [8]. Due to their unique advantages of rapid deployment, survivability, high concealment, and low cost, IWSN is widely used in many fields.

An essential part of the IWSN applied to the fault detection and diagnosis of industrial machines is to improve the coverage rate of monitoring points [9]. Coverage control is a fundamental problem in IWSN, and it is also a measurement standard of quality of service (QoS) evaluation in IWSN. The IWSN coverage reflects the monitoring quality of the target area by the IWSN and provides the sensing services required by the system [10, 11]. An effective IWSN coverage control strategy can optimize the resource allocation of

the network, improve the energy efficiency of network nodes, and perceive the quality of service.

In engineering practice, to obtain data from the entire area, coverage is one of the most basic requirements. In many cases, wireless sensor networks require monitoring various areas of the network to fully cover the monitored targets. However, in some other wireless sensor network applications, a specific percentage of the network needs to be monitored [12]. Generally speaking, coverage problems can be classified into three categories according to the coverage object: target, area, and barrier coverage [13, 14]. Point coverage refers to the random deployment of sensor nodes near some discrete target points in a given area for data collection and monitoring. The point coverage problem is usually studied by selecting some of the sensor nodes randomly deployed in the area to monitor a specific point in anticipation of the location of the monitored point, optimizing the application of network resources, and ensuring the completion of the monitoring task. Given a fixed target area, the sensor node can monitor all the subareas under that target area is the area coverage, which is aimed at achieving the maximum coverage of the target area. Barrier coverage means that the sensing range of the working node should cover the whole moving trajectory of the moving object when it traverses the area along a specific track. Its purpose is to identify any intruder who may try to enter the network area. This identification helps prevent intruders from sneaking through the network [15].

In the industrial wireless sensor deployment process, two strategies are generally used: random deployment and planned deployment. Because IWSNs typically operate in complex environments and have many industrial wireless sensor nodes in the network, random deployment is often used. However, it is difficult to deploy all industrial wireless sensor nodes to the correct location at the same time by large-scale random deployment method, which is easy to generate unreasonable coverage structure, and form perceptual overlap and blind spot. In this case, the redundant deployment of sensor nodes compensates for the low sensor coverage. Besides, suppose there is an appropriate scheduling mechanism to schedule the activities of sensor nodes. In that case, it will significantly improve the coverage of sensors in IWSN and extend the lifetime of IWSN [16].

To obtain the maximum monitoring range and prolong the lifetime of IWSN with a limited number of nodes, this paper proposes a self-organizing IWSN target coverage method based on a hybrid niche immune genetic algorithm (HNIGA) and establishes a corresponding system model. HNIGA uses genetic operation to evolve antibody populations and niche technology to maintain the diversity of antibody populations, thereby increasing the diversity of the population and increasing the coverage of the target. In the simulations, HNIGA is compared with SA and GA. Simulation results show that the number of target coverage optimized by HNIGA has increased significantly. And considering the quality of service of the network, more rounds of monitoring can be carried out on the target points after the optimization of HNIGA, which means that the monitoring life of the whole network is effectively extended.

The structure of this paper is as follows. Section 2 introduces related research work on target coverage in IWSN, and Section 3 establishes the fault detection coverage model in detail. HNIGA is used to improve the performance of the monitoring capability of IWSN in Section 4. Section 5 shows and discusses the results of the simulation experiment of HNIGA. Then, Section 6 provides a comprehensive summary.

2. Related Work

To maximize capital productivity, reduce health, safety, and environmental (HSE) risks and minimize costs. Paper [17] proposes to collect, process, and analyze data from the device in real time with the help of IWSN through the Internet of Things technology. Paper [18] stated that if detailed machine health indicators are to be analyzed, the infrastructure should accommodate data from sources other than the machine controller itself. The paper proposes a digital architecture that meets these standards and demonstrates use cases in machine utilization and health monitoring to achieve this goal. Paper [19] proposes a hierarchical routing graph construction (HRGC) to meet the needs of real-time and reliable communication in IWSN. In IWSN, quality of service (QOS) is expressed as the accuracy with which the sensor covers or monitors the target set within its sensing range.

In paper [12], which is aimed at the problem of connected p-persistent coverage in wireless sensor networks (WSNs), they proposed a pDCDS algorithm, which is based on the connected dominating set of degree constraints, can significantly improve the network lifetime. In paper [15], a distributed boundary monitoring (DBS) algorithm is proposed, which can find the best node to ensure obstacle coverage. Paper [16] proposes an efficient scheduling method LAML based on learning automata. The algorithm makes nodes choose their appropriate state (active or sleeping) at any given time to extend the lifetime of the network better.

The key to using IWSN to monitor industrial equipment is to maximize the number of monitoring points under a limited number of industry wireless sensors while considering the energy limitation of the wireless sensor itself. When each target needs to be covered by a fixed number of sensors, the sensor nodes need to be dynamically adjusted to monitor deployment targets to improve detection efficiency. Researchers have done much work on the target coverage problem in IWSN and put forward many solutions.

Paper [20] designs a model based on hybrid optimization to solve the coverage and connection problems. An optimization method based on Hybrid Film Group Optimizer (TSO) is proposed to optimize the problem. Paper [21] points out that a critical issue in industrial wireless sensor networks is coverage. In that paper, an energy-efficient heuristic method is used to solve the k-coverage problem. The experimental results show that the performance of the heuristic algorithm is close to the best and shows an improvement in network lifetime. However, because the implementation of the solution is too complicated, it is not suitable for actual operation.

Paper [22] proposes an adaptive coverage and connectivity (ACC) scheme. It uses two basic methods, the first of which can provide the best coverage for all target objects,

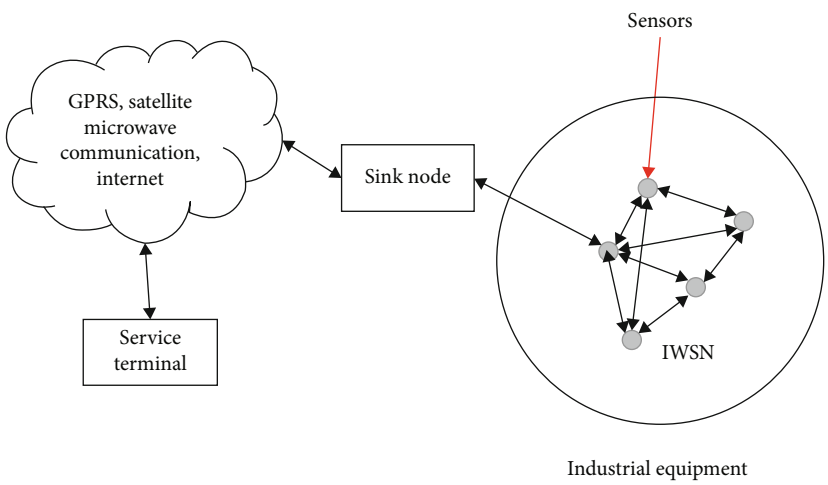


Figure 1: Fault diagnosis of industrial equipment in IWSN.

and its mathematical model can ensure coverage. The second method deals with network connectivity and energy consumption. Paper [23] divides the deployment set of all sensors into disjoint subsets or sensor coverings so that each sensor covering can monitor the entire target at any time. Then, a cuckoo search algorithm is proposed to optimize the sensor coverage problem to achieve the goal of maximizing the coverage of target points. However, the cuckoo algorithm has the problem of weak local search ability and slow convergence speed, and it cannot search for the optimal solution in a short time.

Paper [24] proposes an improved wireless sensor network scheduling based on genetic algorithm. An effective chromosome representation is given, and it is proved that it produces effective chromosomes after crossover and mutation operations. The derivation of the fitness function has four contradictory goals, choosing the least number of sensors, the full coverage of the selected sensor nodes, connectivity, and energy level. In that paper, a novel mutation operation is introduced to improve the performance of GA-based algorithms and speed up the convergence speed. However, GA has a specific dependence on selecting the initial population, which makes the algorithm prone to premature maturity and premature convergence.

Paper [25] uses SA and PSO to deploy sensor nodes effectively and determine the sensing range of sensor nodes in advance. Then, SA and Dempster-Shafer theory are used to effectively schedule the sensor nodes, thereby finding the optimal scheduling scheme and location of the sensor nodes. However, due to the requirement of higher initial temperature, slower cooling rate, lower final temperature, and enough sampling at each temperature, the optimization process is longer, and the convergence speed is slow. Thus, it is not easy to achieve the ideal optimization effect in a short time.

3. System Model

3.1. Problem Description. In industrial equipment monitoring, industry wireless sensors need to be used to monitor target points. Under the condition of limited resources, the sensor monitoring scheme optimized by HNIGA can maximize the number of monitored target points and improve the monitoring efficiency of IWSN. It can be better applied to the condition monitoring of industrial production equipment to avoid unnecessary losses. As in Figure 1, the industry wireless sensor nodes are arranged on industrial equipment. Each diagnostic node is an intelligent industry wireless sensor that can collect various signals generated by the operation equipment. According to the different data types that need to be collected, industry wireless sensors are divided into vibration sensors, flow sensors, and temperature sensors. These intelligent industry wireless sensors have specific data processing capabilities and can convert collected analog data into digital signals. Second, the industry wireless sensors send the collected data to the receiving device in a specific format through IWSN. The receiving device can convert the data collected from the sensors into computer-recognizable data and then submit it to the service terminal. The service terminal processes the collected data, and it can also actively send requests to the equipment for obtaining specified types of data. The service terminal records the collected or requested data, analyzes current or historical data according to specific diagnostic algorithms, evaluates the current equipment operation state or predicts the subsequent operating state of the equipment, and provides maintenance and diagnosis recommendations.

Lifetime analysis is an essential issue in IWSN design. The purpose of lifetime analysis is to extend network life as long as possible. Hence, network designers need to optimize network design from different perspectives to extend the lifetime of IWSN. There are three main methods commonly used to define the network lifetime, the definition based on the number of effective nodes, the lifetime definition based on the coverage, and the lifetime definition based on the connectivity.

However, in this paper, the definition of lifetime is based on the quality of service (QoS). The fundamental purpose of the network is to serve the fault diagnosis of industrial equipment in IWSN. If the network cannot meet the requirements, the network will be invalid. Thus, in the simulation, the network lifetime is defined as follows. The energy of sensor nodes shows a normal distribution. When the target coverage

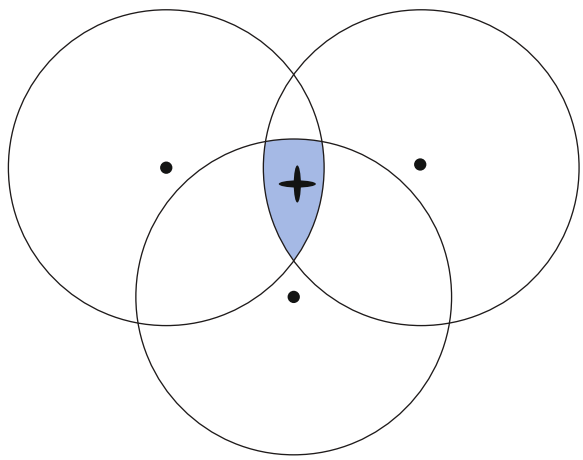


Figure 2: Coverage relationship between sensors and targets.

in the network is less than 50%, the whole monitoring network will stop working. The working life of the network is defined as the total working time.

To achieve monitoring goals, the IWSN must form a complete monitoring area coverage and meet the coverage requirements of certain applications.

3.2. Target Coverage Model. In the process of monitoring target points using IWSN, each industry wireless sensor node can monitor multiple target points within its sensing radius, and one target point can also be monitored by multiple industry wireless sensor nodes around it. Moreover, the cooperation of multiple industry sensor nodes can improve the accuracy of monitoring data.

To obtain accurate location information of the monitored target point, at least three industry wireless sensors are required to monitor the target point at the same time successfully. As shown in Figure 2, the black dots represent the industry wireless sensors, the circle represents the area that the sensor can monitor, and the four-pointed star is the target point to be monitored. Only when the target point is in the overlap area of the 3 industry wireless sensor monitoring ranges, the accurate position of the target point can be successfully obtained, which is beneficial to diagnosing and maintaining the target point.

Assuming that M industry wireless sensor nodes and N target points are placed in a monitoring area with a range of $L \times L$, in this paper, the value of L is 400 meters, and the coverage relationship between monitored target points and industry wireless sensor nodes is expressed by formula (1).

$$C = \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,N-1} & c_{1,N} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,N-1} & c_{2,N} \\ \vdots & & c_{m,n} & & \vdots \\ c_{M-1,1} & c_{M-1,2} & \cdots & c_{M-1,N-1} & c_{M-1,N} \\ c_{M,1} & c_{M,2} & \cdots & c_{M,N-1} & c_{M,N} \end{bmatrix} \cdot \left(c_{m,n} \in \{0, 1\}, 1 \le m \le M, 1 \le n \le N\right). \tag{1}$$

Due to the limitation of industry wireless sensor monitoring capabilities, the industry wireless sensor nodes can only select a limited number of monitored target points within their monitoring area for monitoring. The monitoring relationship between industry wireless sensor nodes and monitored target points in IWSN can be shown in formula (2).

$$S = \begin{bmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,\mathrm{N}-1} & s_{1,N} \\ s_{2,1} & s_{2,2} & \cdots & s_{2,N-1} & s_{2,N} \\ \vdots & \vdots & s_{m,n} & \vdots & \vdots \\ s_{M-1,1} & s_{M-1,2} & \cdots & s_{M-1,N-1} & s_{M-1,N} \\ s_{M,1} & s_{M,2} & \cdots & s_{M,N-1} & s_{M,N} \end{bmatrix} \cdot \left(s_{m,n} \in \{0, 1\}, 1 \le m \le M, 1 \le n \le N\right). \tag{2}$$

In the monitoring relationship matrix S, when $s_{m,n} = 1$, it shows that the nth sensor node is used to monitor m^{th} target point, and if $s_{m,n} = 0$, it means that although the mth target can be sensed by the nth node, but it is not monitored by m^{th} sensor node at this time. Considering the energy limitation of industry wireless sensor, each sensor can only monitor at most F target points, and the restrictions can be described as formula (3).

$$\sum_{m=1}^{M} s_{m,n} \le F, n = 1 \cdots N. \tag{3}$$

Each industry wireless sensor node can monitor at most F targets simultaneously, since the information of the monitored target point needs to be accurately obtained. To obtain the maximum number of targets that are detected by the sensor, the objective function is designed as formula (4).

$$\sum_{m=1}^{M} s_{m,n} \le F, n = 1 \cdots N. \tag{4}$$

Each industry wireless sensor node can monitor at most F targets simultaneously, since the information of the monitored target point needs to be accurately obtained. To obtain the maximum number of targets that are detected by the sensor, the objective function is designed as formula (5).

$$\max f(s_{11}, s_{12}, \cdots, s_{MN}) = \sum_{m=1}^{M} w_m, \tag{5}$$

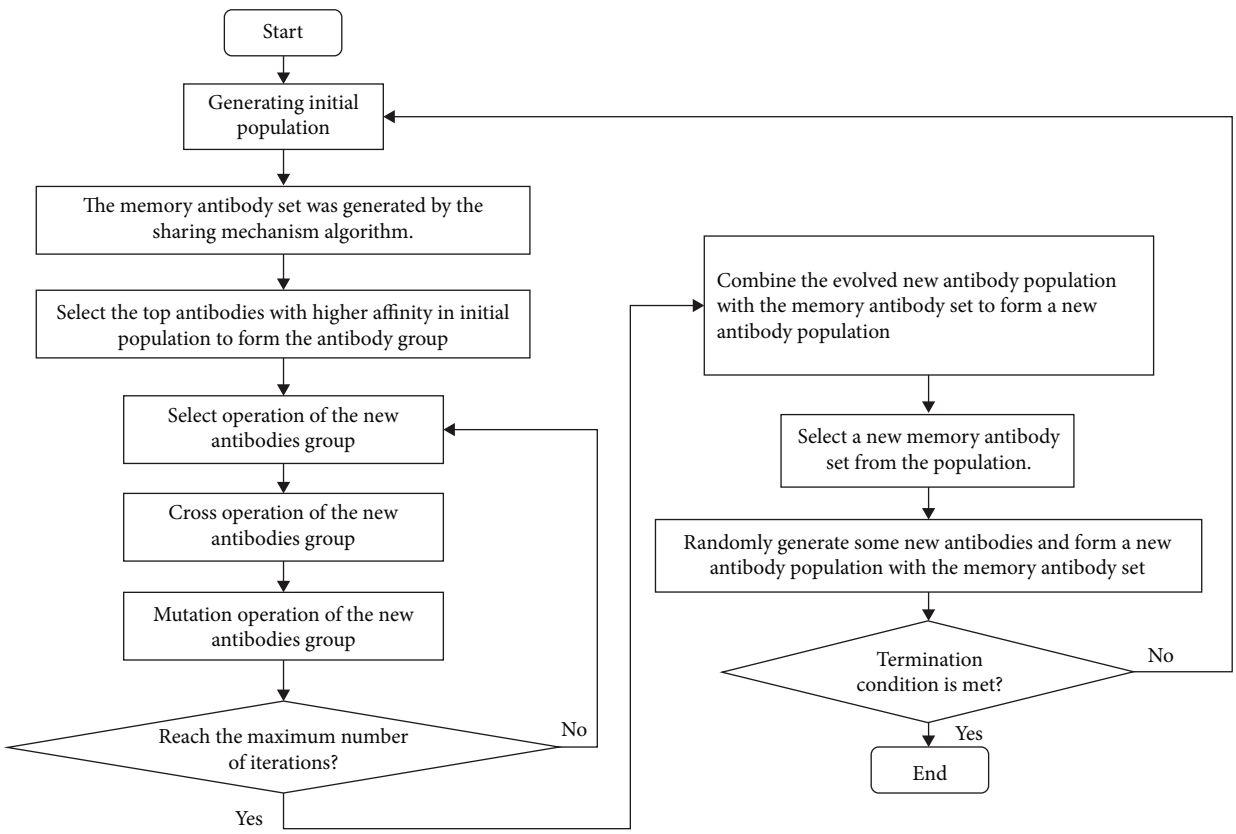


Figure 3: Algorithm flowchart.

where the objective function f denotes the number of targets successfully monitored when the monitoring matrix is R, and w_m is described as formula (6).

$$w_m = \begin{cases} 1 & \sum_{n=1}^{N} s_{m,n} \geq E, \\ 0 & \sum_{n=1}^{N} s_{m,n} < E. \end{cases} \tag{6}$$

The restrictions are described as formulas (7) and (8).

$$\sum_{m=1}^{M} s_{m,n} \leq F, n = 1 \cdots N, \tag{7}$$

$$s_{m,n} \leq c_{m,n}. \tag{8}$$

Equation (7) indicates that each node can simultaneously monitor F objects at most, and formula (8) represents that a target can be monitored by a sensor node only if it is within the node coverage.

In this paper, fault detection coverage is a kind of point coverage described as the industrial wireless sensors are randomly deployed in a designated monitoring area to continuously monitor the key points of industrial equipment in this area. The monitoring effect of the network is expressed by the target coverage, which refers to the proportion of the number of successfully detected target points in the total number of target points. To exclude the emergence of a single special case, the network coverage under the same condition is calculated 10 times, and then, the average value is taken as the final network coverage.

4. HNIGA-Based Fault Detection Coverage in IWSN

In this paper, a novel optimization algorithm HNIGA is used to solve the problem of fault detection coverage in IWSN, which follows the framework of conventional heuristic methods and is a randomized search algorithm that draws on natural selection and natural genetic mechanisms of the biological world. The algorithm uses niche technology to maintain the diversity of solutions. It has a high global optimization capability and convergence speed simultaneously. It overcomes two typical disadvantages of traditional genetic algorithms: GA is prone to premature, which makes the final search results in the local optimal solution. In the later stage of GA evolution, the search efficiency will decrease, and the search speed will slow down.

In HNIGA, the antigen is considered as a problem, and the antibody corresponds to the candidate solution of the problem. This paper introduces the niche technology based on the crowding-out mechanism, which can avoid the large number of antibodies with high affinity in the later stage of the algorithm evolution, filling the entire group. It can better maintain the diversity of solutions and has a high global optimization capability and convergence speed. The basic idea is

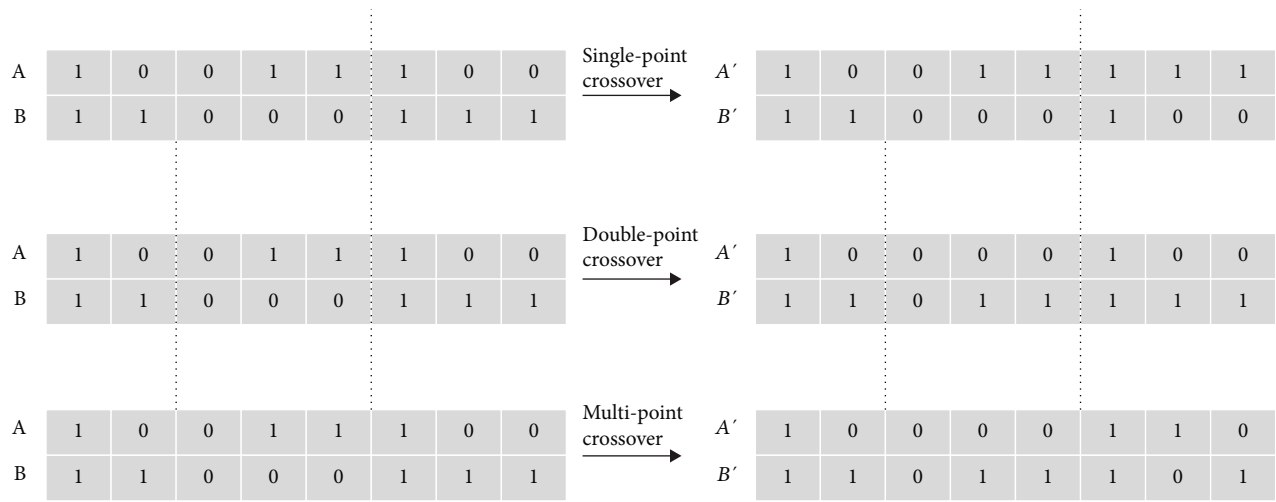


Figure 4: Cross operation diagram.

to set a crowding factor CF in the algorithm and randomly select 1/CF individuals from the group to form exclusion members. Then, with the similarity between the newly generated antibodies and the excluded members, some antibodies similar to the excluded members are excluded. The similarity between antibodies can be measured by the Hamming distance between the antibody code strings. New individuals always replace the antibodies close to them, which can prevent the population from converging to an only antibody to a certain extent. As the crowding process proceeds, the antibodies in the group are gradually classified, thus forming a small generation environment and maintaining the diversity of the group.

HNIGA uses a two-dimensional matrix to represent antibodies, and each antibody denotes a solution for target coverage. HNIGA's algorithm flow is as follows.

Step 1. Determine the evolutionary algebra and then initialize an antibody population using chaos operator.

Step 2. Apply the sharing mechanism algorithm to the initial population to obtain the memory antibody set.

Step 3. A certain number of antibodies with higher affinity are selected from the initial population to form a new population to be evolved. The sum of the number of this population and the number of memory antibody sets is the number of the initial antibody population.

Step 4. Evolve the new antibody population generated in Step 3 by selecting cross and mutation operations.

Step 5. Combine the evolved new antibody population with the memory antibody set to form a new antibody population and select a new memory antibody set from the population.

Step 6. Randomly generate some new antibodies and form a new antibody population with the memory antibody set, the size of which is the same as the initial population.

Step 7. Judge whether the termination condition is met and output the result if it is met; otherwise, return to Step 2.

The algorithm flowchart of HNIGA is shown in Figure 3.

4.1. Solution Encoding. In the fault detection coverage problem in IWSN, one of the most important things is to encode the coverage relationship matrix C and monitoring relationship matrix S. In the coverage relationship matrix C, if the target t is within the coverage radius of the sensor s, then $c_{t,s} = 1$. The number of 1 in the coverage relationship matrix C is the code length of the population. The dimension D of the individuals in the population is the total number of targets successfully sensed by all sensors.

$$D = \sum_{t=1}^{T} \sum_{s=1}^{S} c_{t,s}. \tag{9}$$

The coverage relationship matrix between the node and the monitored object is as shown in equation (10). The underlined values in matrix C are randomly coded.

$$C = \begin{bmatrix} 0 & \underline{1} & \underline{1} & 0 & \underline{1} \\ \underline{1} & 0 & \underline{1} & \underline{1} & \underline{1} \\ \underline{1} & \underline{1} & \underline{1} & 0 & 0 \\ \underline{1} & 0 & \underline{1} & 0 & \underline{1} \\ 0 & \underline{1} & 0 & 0 & \underline{1} \end{bmatrix}. \tag{10}$$

The monitoring relationship matrix S is shown in equation (11). When the value of the underline in matrix C is 1, the value of the corresponding position in R is recoded, and the position of the value 0 remains unchanged. In the recoded matrix, 1 and 0, respectively, indicate whether the target is successfully monitored or not.

$$S = \begin{bmatrix} 0 & \underline{1} & \underline{0} & 0 & \underline{1} \\ \underline{0} & 0 & \underline{1} & \underline{1} & \underline{1} \\ \underline{0} & \underline{0} & \underline{1} & 0 & 0 \\ \underline{1} & 0 & \underline{1} & 0 & \underline{0} \\ 0 & \underline{1} & 0 & 0 & \underline{1} \end{bmatrix}. \quad (11)$$

4.2. Initialization of Antibody Population. In the fault detection coverage problem in IWSN, small changes in the initial conditions will lead to entirely different subsequent state changes in a chaotic phenomenon, which is called the randomness and ergodicity of the chaotic phenomenon. Due to these characteristics, combining chaotic methods and evolutionary algorithms will produce some more excellent performance, such as increasing population diversity and jumping out of locally optimal solutions.

In the initial antibody group, adding the chaos operator can improve the algorithm's global search ability and avoid premature convergence. The one-dimensional Logistic mapping is as follows.

$$b_{g+1} = \mu \times b_g \times (1 - b_g), \quad (12)$$

$$A = \begin{cases} 0 & b_g \leq 0.5, \\ 1 & b_g > 0.5. \end{cases} \quad (13)$$

In formulas (12) and (13), b_g is a random number between 0 and 1, and g is the number of iterations. The generated chaotic sequence consists of random decimals between 0 and 1. μ is the control parameter. When $\mu \in (3.5699, 4]$, the system is in a chaotic state; when $\mu = 4$, the system reaches a completely chaotic state.

4.3. Affinity Evaluation. In the optimization problem of the fault detection coverage problem in IWSN, increasing the number of targets monitored by the sensor is the main goal to be optimized by the HNIGA. Its affinity can be calculated by formula (4).

4.4. Selection. The function of the selection operator is to select antibodies from the antibody population to form a new antibody population. The basis of selection is the calculation of the objective function of the individual, and its main purpose is to avoid the loss of valuable information in the HNIGA. Antibodies with higher affinity are more likely to be selected. However, the selection operation is not solely based on the affinity from high to low. However, according to the affinity, each antibody's probability in the population being inherited to the next generation population is calculated through the roulette wheel selection strategy. Antibodies with higher affinity also have the possibility of being eliminated, and antibodies with low fitness also have the possibility to pass on to the next generation.

The roulette selection strategy is used to calculate the probability of each antibody in the population inheriting the next generation population based on the affinity. The probability of each antibody being inherited to the next generation is the proportion of its affinity in the total affinity of the entire population. Therefore, the antibody selection is made through the roulette strategy. The higher the antibody affinity, the greater the probability that the antibody will be selected. The probability of entering the next generation can be calculated by formula (14).

Table 1: Parameter settings of HNIGA.

	Evolutionary algebra	Antibody population	Number of memory antibody sets	Mutation probability
HNIGA	100	40	10	0.08

$$P = n \times \frac{f(x)}{\sum_{i=1}^{n} f(x_i)}, \quad (14)$$

where P is the expected value of being selected, n is the population number, x is the number of the individual in the population, and $f(x)$ is the affinity of the antibody. By selection operator, it is possible to make the offspring individuals keep approaching the optimal solution in the optimization process of the genetic algorithm until the optimization is completed.

4.5. Crossover. The role of the crossover operator is to generate new antibodies. The crossover operator makes the homologous chromosomes form new chromosomes after mating, which determines the global search ability of HNIGA. Thus, it makes the crossover operator an essential part of HNIGA. The offspring produced by the crossover operator can inherit the essential characteristics of the parent. The crossover operator is mainly divided into two steps. The first is to determine the position of the crossover point in the crossover process according to the calculation of the crossover probability, and the second is to determine the way of swapping genes according to the design of the crossover operator. As shown in Figure 4, at present, standard crossover methods include single-point crossover, double-point crossover, and multi-point crossover. The difference between them lies in the number of crossover points randomly set in the individual coding, and then, these points exchange part of the chromosomes of two paired individuals. The schematic diagram of the crossover operation is as follows.

4.6. Immune Mutation. After cross-recombination is antibody variation, the individual variables of the offspring mutate with a small probability or step length. The probability or step length of the variable transformation is inversely proportional to the dimension and has nothing to do with the population size. According to research, mutation itself is a kind of local random search, combined with selection and recombination operators, to ensure the effectiveness of HNIGA, at the same time, the diversity of the population is guaranteed to prevent premature convergence.

4.7. Niche Crowding Operation. Combine the memory antibody group containing D individuals with the new population

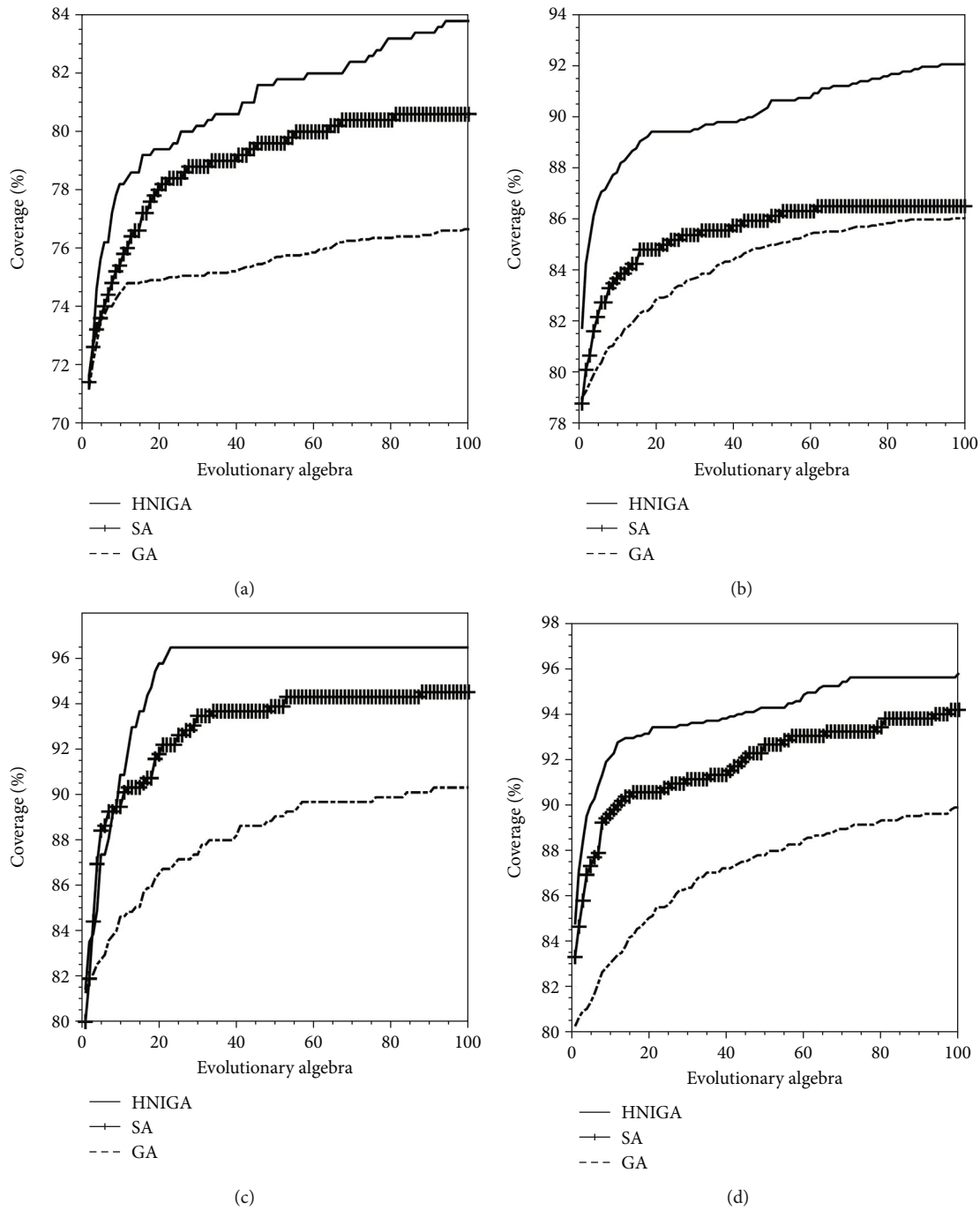


Figure 5: The coverage of successfully monitored targets varies with algorithm iteration under conditions of Table 2: (a) 50 m sensing radius, 100 sensors. (b) 60 m sensing radius, 100 sensors. (c) 70 m sensing radius, 100 sensors. (d) 80 m sensing radius, 100 sensors.

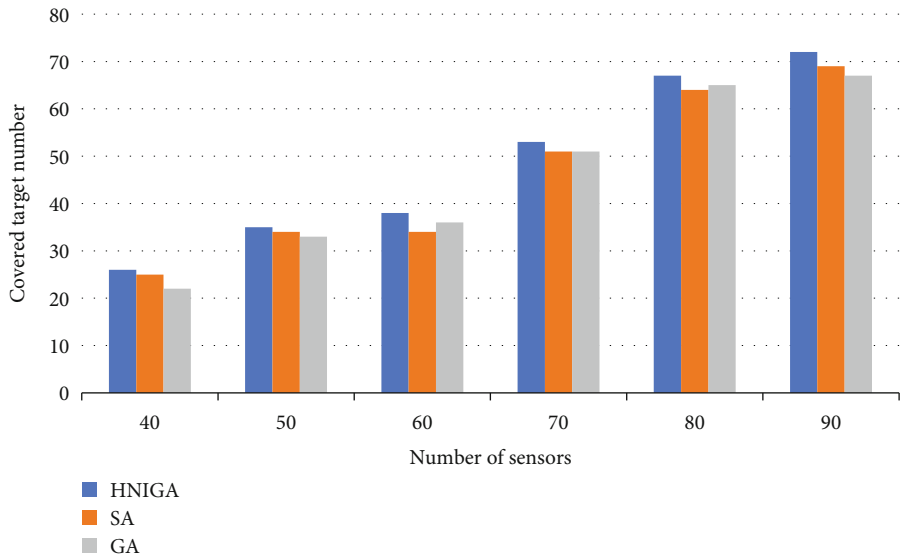
Table 2: The simulation conditions in Figure 5.

	Sensing range	Number of targets	Number of sensors
Figure 5(a)	50	100	100
Figure 5(b)	60	100	100
Figure 5(c)	70	100	100
Figure 5(d)	80	100	100

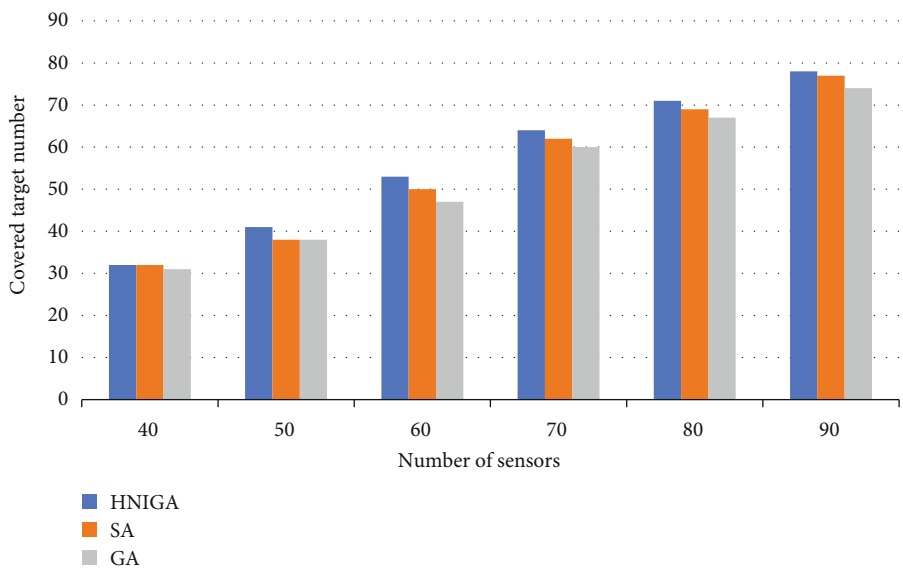
containing N antibodies acquired by immune operation evolution to generate a new population, which contains $D+N$ antibodies. Then, calculate the Hamming distance between antibodies X_i and X_j in this new population.

$$d_{i,j} = \mathrm{dist}\left(X_i, X_j\right) = \mathrm{sum}\left(X_i \oplus X_j\right), \tag{15}$$

where $i \in (1, 2, 3\cdots, M+N-1)$, $j \in (i+1, i+2, \cdots, M+N)$, and the function sum() is used to calculate the number of bits

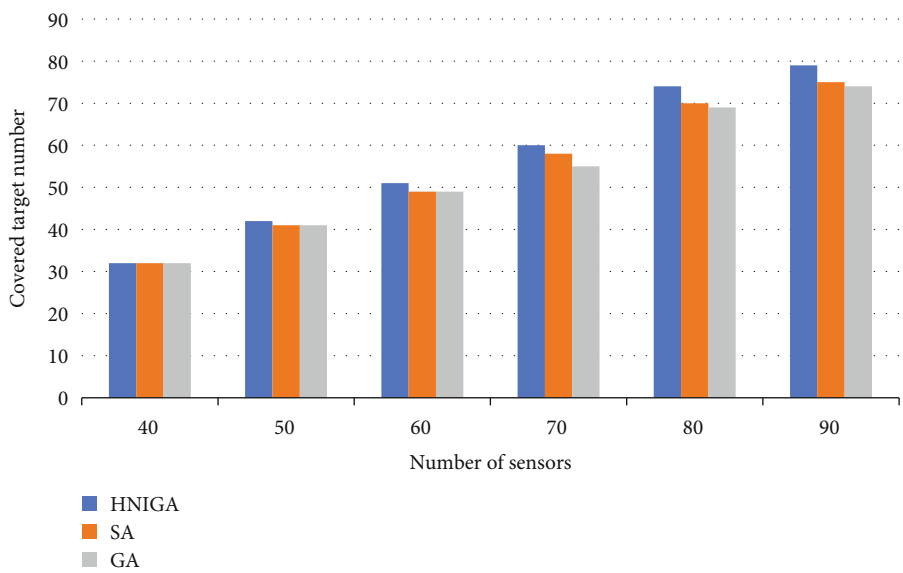


(a)

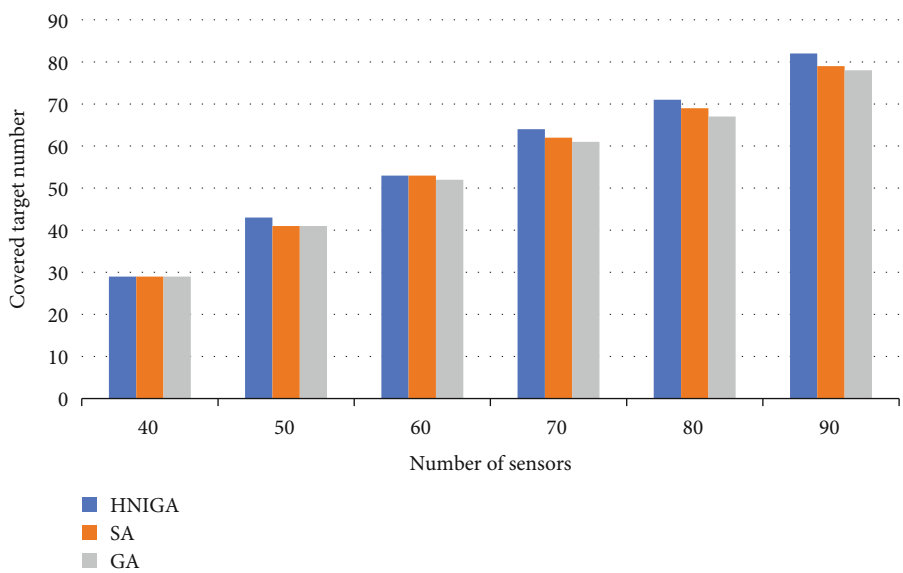


(b)

Figure 6: Continued.



(c)



(d)

Figure 6: Covered target number: (a) sensing radius of 55 meters; (b) sensing radius of 65 meters; (c) sensing radius of 75 meters; (d) sensing radius of 85 meters.

with the value of 1 in the *XOR* result. When $d_{i,j}$ is less than the average distance l, compare the fitness of the antibodies X_i and X_j; then, a penalty function is given to antibodies with higher affinity.

4.8. Find the Optimal Solution. In the fault detection coverage problem in IWSN, the purpose of this paper is to maximize the fault detection coverage. Continuous iterative evolution of the population usually produces antibodies with higher affinity. The iterative optimization process of HNIGA is repeated until the iteration reaches the termination condition and the algorithm ends. Finally, the algorithm will output the antibody with the highest affinity value as the best way to solve the problem.

Table 3: Number of sensor deaths during each round of monitoring.

Round	1	2	3	4	5	6	7	8	9	10
Dead	1	2	2	2	1	1	1	1	3	5
Round	11	12	13	14	15	16	17	18	19	20
Dead	1	1	1	2	1	2	2	3	2	2
Round	21	22	23	24	25	26	27	28	29	30
Dead	3	2	5	1	4	1	5	4	2	3
Round	31	32	33	34	35	36	37	38	39	40
Dead	3	1	2	3	2	1	3	5	1	1
Round	41	42	43	44	45	46	47	48	49	50
Dead	2	1	1	1	2	1	1	1	1	1

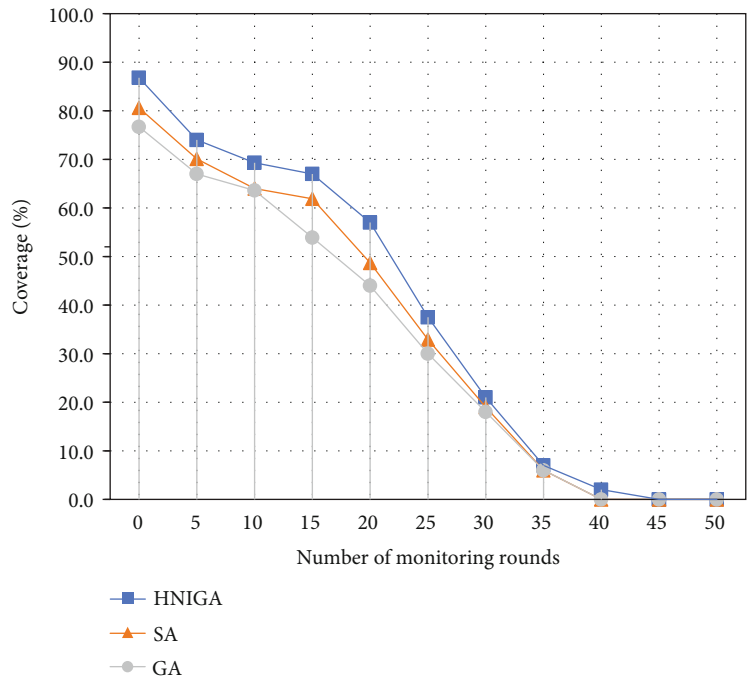


Figure 7: Target coverage under different periods.

5. Simulation

In this section, which is aimed at the problem of fault detection coverage and network lifetime in IWSN, this paper compares the optimization effect of HNIGA with traditional SA and GA through simulation. The hardware platform used in this experiment is a Ryzen5 3500x computer with 16 GB memory, and the software platform is MATLAB. In this case, the objective function (4) is used to calculate the number of targets successfully detected in the IWSN.

In the simulation, the sensing area is set to $400 \times 400\text{m}^2$, and 100 targets and 100 industry wireless sensor nodes are randomly distributed in this area. Assuming that each industry wireless sensor can monitor up to 4 targets, each target point must be monitored by 3 industry sensor nodes in the same period.

In HNIGA, the number of antibodies is 40, and the number of iterations is 100. Table 1 shows the detailed parameters of HNIGA, and the simulation results of the algorithm are shown in Figures 5(a)–5(d).

Figures 5(a)–5(d) are comparison diagrams of target coverage of HNIGA, SA, and GA when the industry sensor's sensing range is 50 m, 60 m, 70 m, and 80 m, respectively, and Table 2 shows the simulation conditions in Figure 5.

In (a), the monitoring radius of sensors is set to 50 m, and the number of sensors and target points are both 100. After optimization by HNIGA, the fault detection coverage can reach 83.8%, and the fault detection coverage optimized by SA can reach 80.6%, while fault detection coverage optimized by GA can reach 76.7%. In (b) except that the monitoring radius of the sensor is set to 60 meters, the other parameter variables do not change, and the maximum coverage of targets that can be monitored through the fault detection schemes optimized by HNIGA, SA, and GA are 92.9%, 87%, and 86.5%, respectively. In (c), monitoring radius of the sensor is set to 70 meters, and the maximum coverage of targets that can be monitored by the fault detection scheme optimized by the three algorithms are 95.6%, 93.8%, and 89.8%, respectively. In (d), monitoring radius of the sensor is set to 80 meters, and the fault detection coverage optimized by HNIGA, SA, and GA can reach 96.2%, 94.0%, and 90.2%, respectively.

As shown in Figure 6, it can be seen that with the increasing number of sensors in the network, the number of successfully monitored targets is also increasing. By comparison, it is obvious that HNIGA always has better optimization performance than the other two algorithms under different monitoring capabilities.

In IWSN, the network nodes are supplied with energy from the batteries they carry, which have a limited charge, and as time passes, the nodes' energy is continuously depleted. Since the energy of sensor nodes is normally distributed, it is unlikely that all nodes will stop working simultaneously due to energy depletion. However, all sensor nodes will stop working one after another. In this process, the monitoring sensors in the whole monitoring network will be reduced continuously, and the coverage will be reduced accordingly.

To verify the optimization effect of HNIGA on the network lifetime, we assume that 100 sensors with a radius of 50 m can carry out 50 rounds of monitoring at most, and a certain number of sensors will run out of energy and stop working in each round. At the beginning of each round of monitoring, we use HNIGA, SA, and GA to optimize the network and calculate the network coverage. The number of sensors dead in each round is shown in Table 3. The target

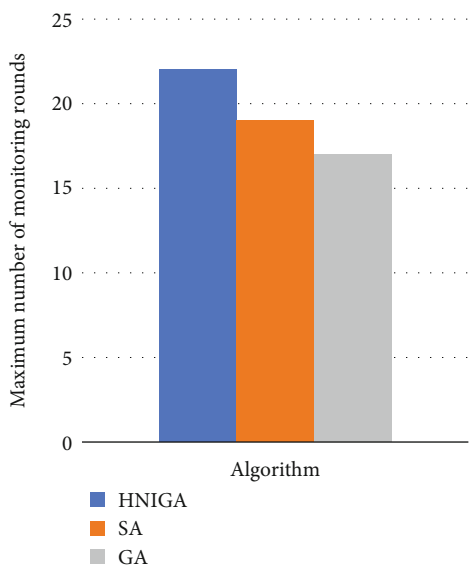


Figure 8: The maximum number of monitoring rounds of the three algorithms.

coverage under different periods is shown in Figure 7. The number of monitoring groups for different algorithms is shown in Figure 8.

From Figure 7, we can see that without considering that the network will stop working when the monitoring coverage is less than 50%, the IWSN optimized by HNIGA is always better than the other two algorithms in terms of target coverage. Considering the QoS of IWSN, the whole network will stop working when the target coverage rate is lower than 50%. From Figure 8, we can see that the IWSN optimized by HNIGA can complete 22 rounds of monitoring. However, the IWSN optimized by SA and GA can only complete 19 rounds and 17 rounds of monitoring, respectively. Since the coverage of the network optimized by the HNIGA is always higher than the other two algorithms, the IWSN optimized by HNIGA will be able to maintain a longer lifetime.

6. Conclusions

This paper proposes a novel hybrid niche immune genetic algorithm to optimize the fault detection coverage problem in IWSN. This algorithm has advantages in maximizing the number of monitored target points. Both HNIGA and traditional GA are based on population evolution. However, the difference is that HNIGA introduces chaos operator to overcome the problem that GA is easily affected by the initial population, and niche technology is introduced to increase the diversity of individuals at the same time. Besides, compared with SA, HNIGA is based on biological evolution as a prototype and has better convergence. When the calculation accuracy is required, the calculation time of HNIGA is less, and the robustness is high. Therefore, the proposed HNIGA has distinct advantages for fault detection in IWSN. To verify the algorithm's optimal performance, this paper designed a mathematical model of the target monitoring problem, and a large number of simulations are performed. Simulation results show that in the case of different perception radius, compared with SA and GA, HNIGA has a better optimization effect, avoids falling into the local optimum, effectively increases the fault detection coverage, and extends the lifetime of the whole network.

Data Availability

The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Disclosure

The funders had no role in the design of the study, in the collection, analyses, or interpretation of data, in the writing of the manuscript or in the decision to publish the results.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

This paper was funded by the Corps innovative talents plan, grant number 2020CB001, the project of Youth and Middleaged Scientific and Techno-logical Innovation Leading Talents Program of the Corps, grant number 2018CB006, the China Postdoctoral Science Foundation, grant number 220531, the Funding Project for High Level Talents Research in Shihezi University, grant number RCZK2018C38, the Project of Shihezi University, grant number ZZZC201915B, and the Postgraduate education innovation program of the Autonomous Region.

References

[1] B. Bengherbia, R. Kara, A. Toubal, M. O. Zmirli, S. Chadli, and P. Wira, "FPGA implementation of a wireless sensor node with a built-in ADALINE neural network coprocessor for vibration analysis and fault diagnosis in machine condition monitoring," *Measurement*, vol. 163, p. 13, 2020.

[2] H. F. Ling, T. Zhu, W. X. He, H. C. Luo, Q. Wang, and Y. Jiang, "Coverage optimization of sensors under multiple constraints using the improved PSO algorithm," *Mathematical Problems in Engineering*, vol. 2020, 10 pages, 2020.

[3] U. Saeed, S. U. Jan, Y. D. Lee, and I. Koo, "Fault diagnosis based on extremely randomized trees in wireless sensor networks," *Reliability Engineering & System Safety*, vol. 205, p. 11, 2021.

[4] L. M. Yang, L. M. Su, Y. X. Wang et al., "Metal roof fault diagnosis method based on RBF-SVM," *Complexity*, vol. 2020, 12 pages, 2020.

[5] V. Gnanasekar, S. S. Manohar, and M. S. Kumaran, "Priority-based channel scheduling and route discovery for IWSN applications," *Journal of Information Science and Engineering*, vol. 35, no. 4, pp. 769–786, 2019.

[6] S. Kim, C. Kim, H. Cho, and K. Jung, "A hierarchical routing graph for supporting mobile devices in industrial wireless sensor networks," *Sensors*, vol. 21, no. 2, p. 14, 2021.

[7] H. Wang, F. Yu, M. Li, and Y. Zhong, "Clock skew estimation for timestamp-free synchronization in industrial wireless sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 1, pp. 90–99, 2021.

[8] C. C. Du, S. Gao, N. S. Jia et al., "A high-accuracy least-time-domain mixture features machine-fault diagnosis based on wireless sensor network," *IEEE Systems Journal*, vol. 14, no. 3, pp. 4101–4109, 2020.

[9] S. Harizan and P. Kuila, "A novel NSGA-II for coverage and connectivity aware sensor node scheduling in industrial wireless sensor networks," *Digital Signal Processing*, vol. 105, p. 14, 2020.

[10] A. R. Aravind and R. Chakravarthi, "Fractional rider optimization algorithm for the optimal placement of the mobile sinks in wireless sensor networks," *International Journal of Communication Systems*, vol. 34, no. 4, p. 21, 2021.

[11] D. Dash, "Approximation algorithms for road coverage using wireless sensor networks for moving objects monitoring," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 11, pp. 4835–4844, 2020.

[12] H. Mostafaei, M. U. Chowdhury, R. Islam, and H. Gholizadeh, *Connected P-Percent Coverage in Wireless Sensor Networks Based on Degree Constraint Dominating Set Approach*, Assoc Computing Machinery, New York, 2015.

[13] L. Nguyen and H. T. Nguyen, "Mobility based network lifetime in wireless sensor networks: a review," *Computer Networks*, vol. 174, p. 24, 2020.

[14] K. Tarnaris, I. Preka, D. Kandris, and A. Alexandridis, "Coverage and k-coverage optimization in wireless sensor networks using computational intelligence methods: a comparative study," *Electronics*, vol. 9, no. 4, p. 24, 2020.

[15] H. Mostafaei, M. U. Chowdhury, and M. S. Obaidat, "Border surveillance with WSN systems in a distributed manner," *IEEE Systems Journal*, vol. 12, no. 4, pp. 3703–3712, 2018.

[16] H. Mostafaei, M. Esnaashari, and M. R. Meybodi, "A coverage monitoring algorithm based on learning automata for wireless sensor networks," *Applied Mathematics & Information Sciences*, vol. 9, no. 3, pp. 1317–1325, 2015.

[17] G. S. Wang, Y. Liu, X. W. Chen et al., "Power transformer fault diagnosis system based on Internet of Things," *Eurasip Journal on Wireless Communications and Networking*, vol. 2021, no. 1, p. 24, 2021.

[18] D. Newman, M. Parto, K. Saleeby, T. Kurfess, and A. Dugenske, "Development of a digital architecture for distributed CNC machine health monitoring," *Smart and Sustainable Manufacturing Systems*, vol. 3, no. 2, pp. 68–82, 2019.

[19] F. H. El-Fouly and R. A. Ramadan, "Real-time energy-efficient reliable traffic aware routing for industrial wireless sensor networks," *IEEE Access*, vol. 8, pp. 58130–58145, 2020.

[20] J. Chelliah and N. Kader, "Optimization for connectivity and coverage issue in target-based wireless sensor networks using an effective multiobjective hybrid tunicate and salp swarm optimizer," *International Journal of Communication Systems*, vol. 34, no. 3, 2021.

[21] M. Torshizi and M. J. Sheikhzadeh, "Optimum K-coverage in wireless sensor network with no redundant node by cellular learning automata," *Wireless Personal Communications*, vol. 110, no. 2, pp. 545–562, 2020.

[22] A. Prasanth and S. Jayachitra, "A novel multi-objective optimization strategy for enhancing quality of service in IoT-enabled WSN applications," *Peer-To-Peer Networking and Applications*, vol. 13, no. 6, pp. 1905–1920, 2020.

[23] T. L. Xiang, H. J. Wang, and Y. C. Shi, "Hybrid WSN node deployment optimization strategy based on CS algorithm," in *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, New York, 2019.

[24] S. Harizan and P. Kuila, "Coverage and connectivity aware energy efficient scheduling in target based wireless sensor networks: an improved genetic algorithm based approach," *Wireless Networks*, vol. 25, no. 4, pp. 1995–2011, 2019.

[25] M. Gupta and S. Varma, "Metaheuristic-based optimal 3D positioning of UAVs forming aerial mesh network to provide emergency communication services," *IET Communications*, vol. 2021, p. 18, 2021.

Hindawi
Journal of Sensors
Volume 2021, Article ID 4304548, 25 pages
https://doi.org/10.1155/2021/4304548

Research Article

True Orthoimage Generation Using Airborne LiDAR Data with Generative Adversarial Network-Based Deep Learning Model

Young Ha Shin and Dong-Cheon Lee

Department of Environment, Energy & Geoinformatics, Sejong University, Seoul 05006, Republic of Korea

Correspondence should be addressed to Dong-Cheon Lee; dclee@sejong.ac.kr

Received 22 April 2021; Accepted 20 May 2021; Published 14 June 2021

Academic Editor: Hyung-Sup Jung

Orthoimage, which is geometrically equivalent to a map, is one of the important geospatial products. Displacement and occlusion in optical images are caused by perspective projection, camera tilt, and object relief. A digital surface model (DSM) is essential data for generating true orthoimages to correct displacement and to recover occlusion areas. Light detection and ranging (LiDAR) data collected from an airborne laser scanner (ALS) system is a major source of DSM. The traditional methods require sophisticated procedures to produce a true orthoimage. Most methods utilize 3D coordinates of the DSM and multiview images with overlapping areas for orthorectifying displacement and detecting and recovering occlusion areas. LiDAR point cloud data provides not only 3D coordinates but also intensity information reflected from object surfaces in the georeferenced orthoprojected space. This paper proposes true orthoimage generation based on a generative adversarial network (GAN) deep learning (DL) with the Pix2Pix model using intensity and DSM of the LiDAR data. The major advantage of using LiDAR data is that the data is occlusion-free true orthoimage in terms of projection geometry except in the case of low image quality. Intensive experiments were performed using the benchmark datasets provided by the International Society for Photogrammetry and Remote Sensing (ISPRS). The results demonstrate that the proposed approach could have the capability of efficiently generating true orthoimages directly from LiDAR data. However, it is crucial to find appropriate preprocessing to improve the quality of the intensity of the LiDAR data to produce a higher quality of the true orthoimages.

1. Introduction

True orthoimages are vertical views of the Earth's surface, eliminating distortion of the objects and allowing a view of nearly any point on the ground with a uniform scale. Therefore, the true orthoimages are geometrically equivalent to topographic maps that show true geographic locations of the terrain features. The geometric distortions are caused by the relief displacement of the terrain features (i.e., height variation of the terrain and object surfaces) and perspective projection of the optical cameras that ultimately result in the occlusion areas. In particular, because of the occlusion caused by tall buildings, the surrounding objects are not photographed. Therefore, recovery or compensation of the occlusion areas is crucial in true orthoimage generation [1]. It is a challenging task to automate the entire procedure for generating true orthoimages. Most approaches have focused on detection and recovery of the occlusion areas. During the last decades, numerous studies for generating true orthoimage have been carried out ever since pixel-based differential rectification has been introduced [2]. Traditional photogrammetric methods require aerial triangulation to obtain exterior orientation parameters (i.e., exposure location and rotation angles) of each aerial image and precise 3D object model data such as the digital building model (DBM) with digital terrain model (DTM) to remove geometric distortion and to detect occlusion areas.

The most crucial tasks are identifying and recovering the occluded areas. In particular, in urban areas with dense high-rise buildings, it is quite difficult to recover occluded areas [3, 4]. As shown in Figures 1 and 2, distortions and occlusions occur in various situations, for example, occlusions on the ground surfaces, surrounding buildings, and even rooftops by superstructures. In addition, sloped or

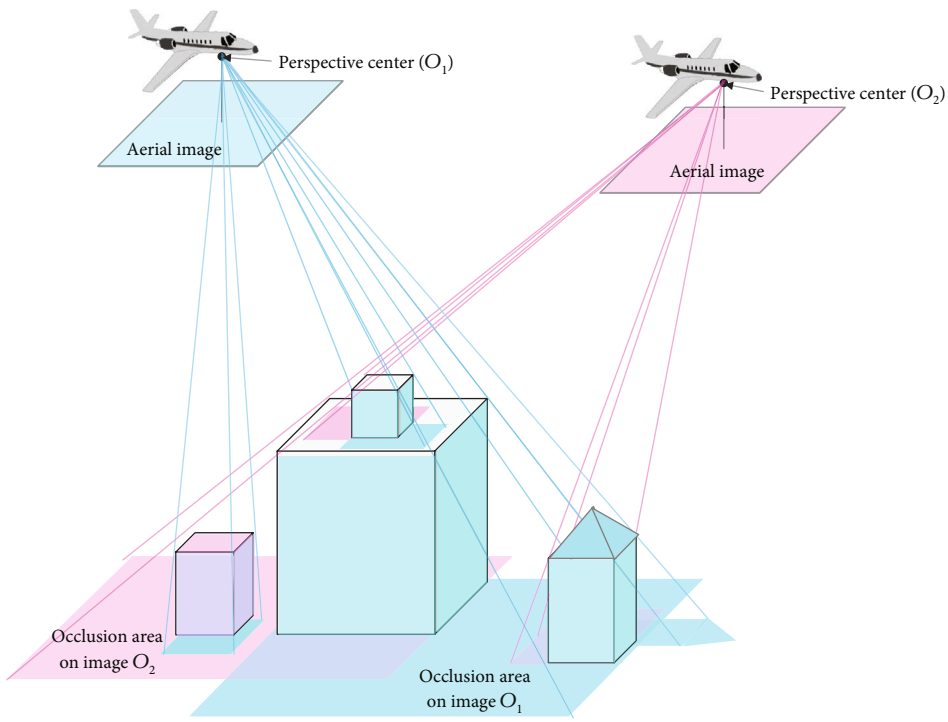


Figure 1: Complex nature of occlusion areas on optical aerial images. Different colors represent occlusions of corresponding images. White regions on rooftops depict nonocclusion areas on both images.

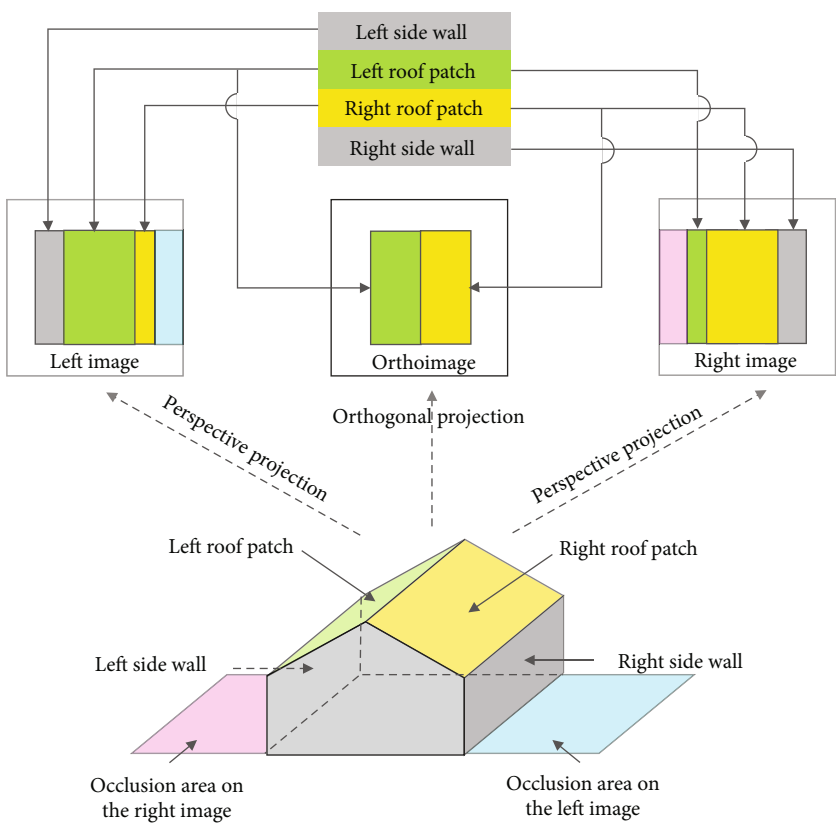


Figure 2: Distortion and occlusion in optical images due to relief displacement with central projection.

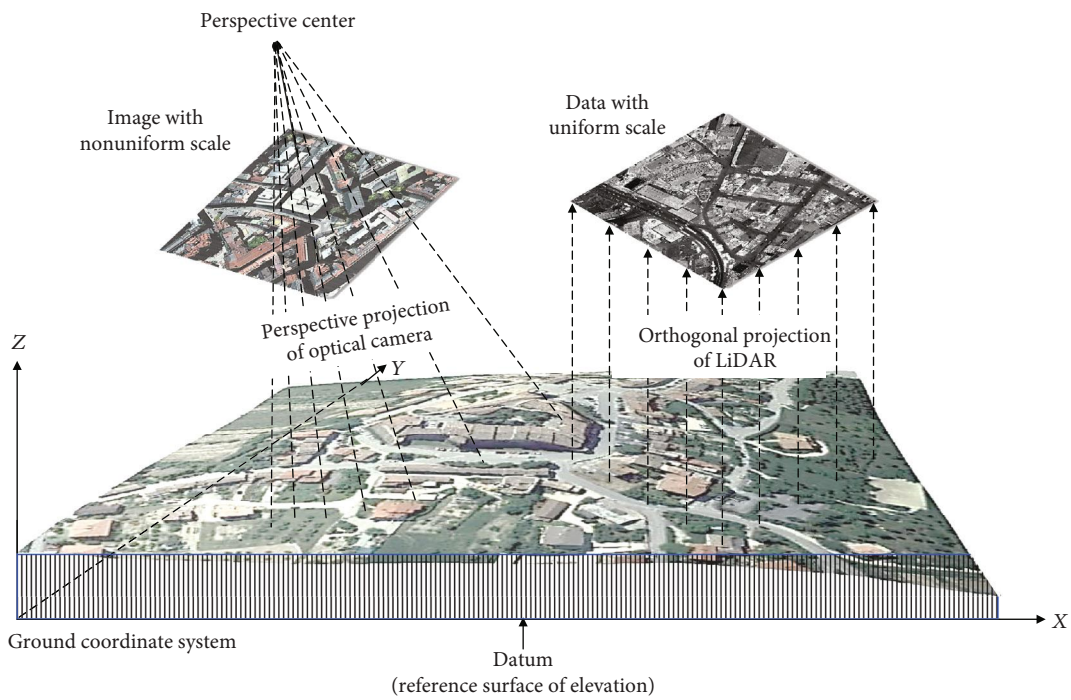


Figure 3: Comparison of orthogonal and central projection.

curved surfaces (e.g., gable, pyramid, hip, and dome roof) with varying heights cause various distortions in the images. Considering the correction of all possible details, generating "true" orthoimages automatically involves incredibly difficult processes. In order to identify and recover occluded areas completely, a high level-of-detail (LoD) building model as well as multiview images with accurate EOPs and high overlapping rates is required. Such traditional photogrammetric methods are complicated and costly.

LiDAR data is essentially composed of 3D coordinates and intensity information with a georeferenced orthogonal coordinate system. On the other hand, images are taken with perspective (or central) projection and need an additional process for georeferencing as shown in Figure 3. Therefore, the LiDAR data has a significant advantage in generating true orthoimage because it is not necessary to correct relief displacement and recover occlusions.

Most of the true orthoimage generation methods focus on how to find the areas occluded by the objects and to fill them using overlapped multiview images [5]. Regarding this matter, some of the remarkable approaches are as follows. One of the popular and effective methods is the Z-buffer algorithm for visibility analysis to identify the occlusion areas. The principle of occlusion detection in this method is quite straightforward. The Z-buffer technique identifies occlusions by keeping track of the number of DSM cells projected to a given image pixel. In this algorithm, the distance of each pixel between the projection center and the object surface is stored in a matrix, and the planimetric location of the pixel is recorded in another matrix. These two matrices are defined in an image that represents a visibility map [6, 7]. The Z-buffer method can be applied to the satellite images obtained from the line scan sensor by searching scan lines corresponding to the equivalent perspective centers [8]. Drawbacks of this technique are sensitivity to the image resolution and DSM cell size. Therefore, false occlusion or false visibility is generated when the image and DSM resolutions are not compatible [3, 9].

Kuzmin et al. [10] proposed a polygon-based method for detecting occlusion areas to generate true orthoimages. In this approach, the differential rectification is applied; then, occlusion areas are detected through the use of polygon surfaces from digital building models (DBMs). Habib et al. [9] and Gharibi and Habib [11] proposed an angle-based occlusion detection method. This method is aimed at improving the accuracy of the occlusion detection using angle and building height using LiDAR data and high-resolution aerial images. They introduced an adaptive radial sweep and spiral sweep approach in which the presence of occlusions can be discerned by sequential checking the off-nadir angles to the direction of light connecting the perspective center to the DSM cells along a radial direction. Yoo and Lee [12] developed a method of patch-based true orthoimage generation while most existing methods have applied a pixel-based approach. The concept of this method is similar to the polygon-based approach in detecting occlusion areas [10]. This method allows mutual recovery of all possible occlusion areas, even including occlusions caused by the superstructures on the rooftops, using multiple images and high level-of-detail digital building model data (i.e., DBM). All images involved in detecting and recovering occlusions can eventually become true orthoimages through compensating each other's occlusions. As a result, multiple true orthoimages are generated for the same scene. Besides the papers

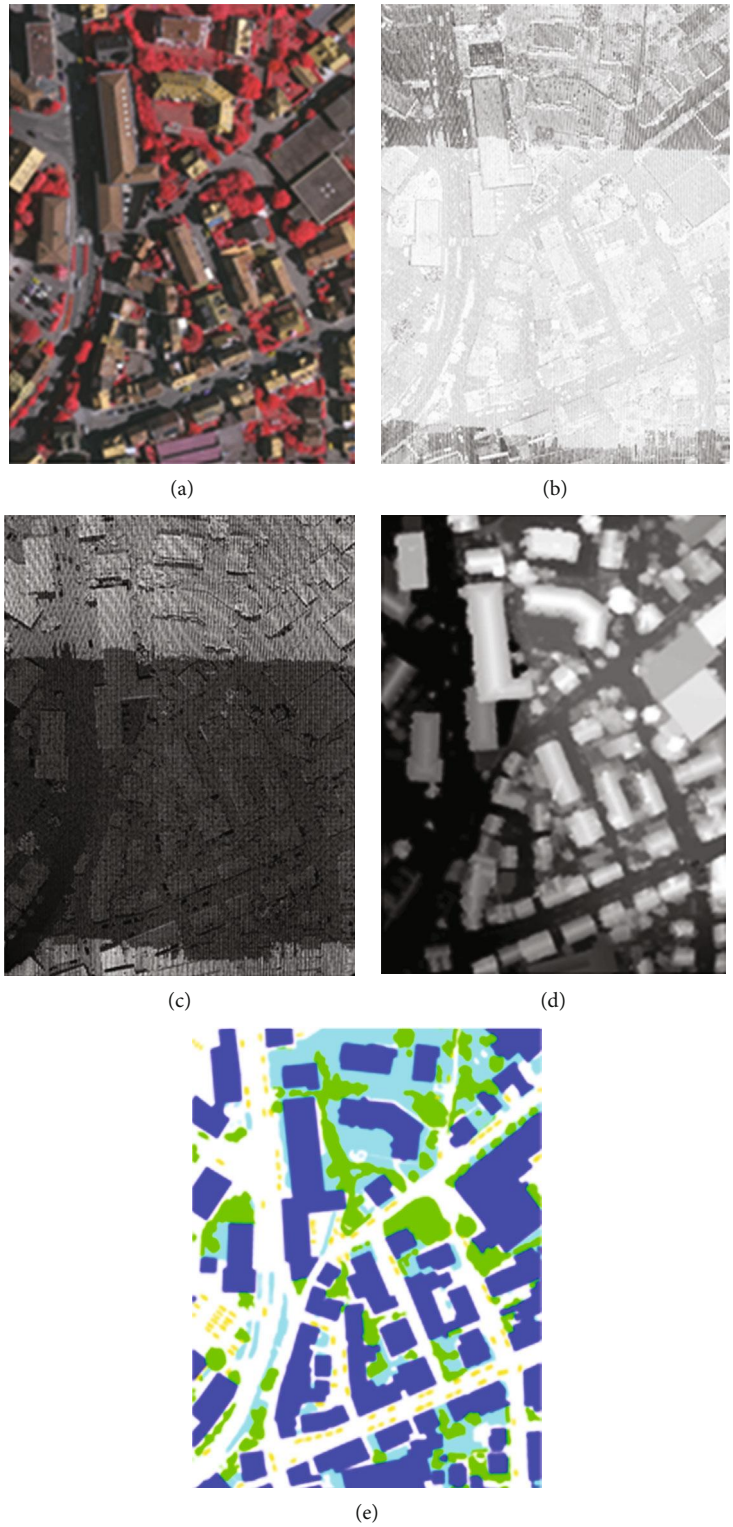


Figure 4: Example of Vaihingen datasets: (a) IR true orthoimage; (b) LiDAR intensity; (c) LiDAR DSM; (d) image matching DSM; (e) label data.

mentioned above, there are numerous numbers of literatures concerning true orthoimage generation approaches [13–19]. Generating true orthoimages is not trivial and requires complicated procedures. In consequence, a standard method is currently not available.

This paper is aimed at investigating the feasibility of utilizing DL to generate true orthoimage rather than traditional methods. Only a few studies related to the neural network or DL-based orthoimage generation are available so far. Bagheri and Sadeghian [20] briefly discussed the ability of the genetic

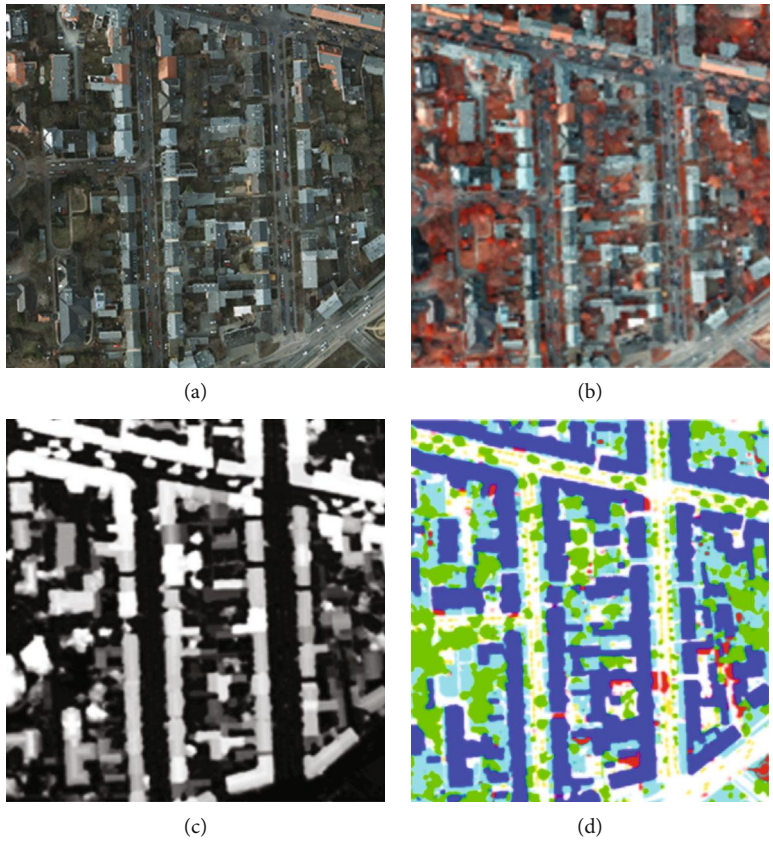


Figure 5: Example of Potsdam datasets: (a) natural color (RGB) true orthoimage; (b) IR true orthoimage; (c) image matching DSM; (d) label data.

algorithm and artificial neural network (ANN) for geospatial products such as geometric modeling of satellite images for orthoimage generation and interpolation of DTM. This paper is based on the previous preliminary study on DL-based true orthoimage generation [21]. In this paper, we have carried out intensive and systematic experiments to train DL under various conditions to improve the quality of the true orthoimages. We propose a method based on the Pix2Pix for image-to-image translation that is one of the generative adversarial network (GAN) models. GAN consists of two networks, the generator, which tries to produce results similar to real images (i.e., true orthoimages), and the discriminator, which judges fake and real images until the results are satisfied. Such a mutually adversarial mechanism improves the quality of the results [22]. The training, validation, and test datasets include infrared (IR) orthoimages, intensity and DSM derived from LiDAR data, and label data provided by the International Society for Photogrammetry and Remote Sensing (ISPRS). Finally, generated IR true orthoimages were converted to the pseudonatural color images using a neural network model with prior initial weights instead of random values. The initial weights were adjusted during the training of the model. Generating true orthoimages with DL that could circumvent the complex steps involved in the traditional method would be an attractive approach even though more improvements in various aspects would be necessary.

2. Datasets and Proposed Method

2.1. Description of Datasets. The Vaihingen and Potsdam datasets provided by ISPRS were used in the experiments [23]. ISPRS datasets are suitable for geoinformatics-related study including this paper since the datasets consist of color infrared (CIR) true orthoimages, airborne LiDAR data, DSM created by image matching, and label data that can be used for training, validation, and test of the DL models (see Figure 4) [24]. The Potsdam datasets consist of IR and natural color (i.e., RGB) true orthoimages, DSM created by image matching, and label data that were not involved in training the DL models but used as the "unseen" data to evaluate the trained DL models. It is noticed that the LiDAR data is not available in the Potsdam datasets (see Figure 5). The label contains object categories for semantic segmentation such as building, road, impervious surface, tree, low vegetation, and water body. The datasets are summarized in Table 1.

The LiDAR data includes 3D coordinates and intensity values of the point clouds. 3D coordinates of the LiDAR data can be used to create DSM. Two types of the DSM (i.e., image matching DSM and LiDAR DSM) were used in the

Table 1: Description of datasets.

Dataset	Image	LiDAR	DSM	Label	GSD (m)	Usage
Vaihingen	IR true orthoimage	Intensity and 3D coordinates	Image matching	Land cover classes	0.09	Training, validation, and test data
Potsdam	Natural color orthoimage and IR true orthoimage	N/A	Image matching	Land cover classes	0.05	Unseen data

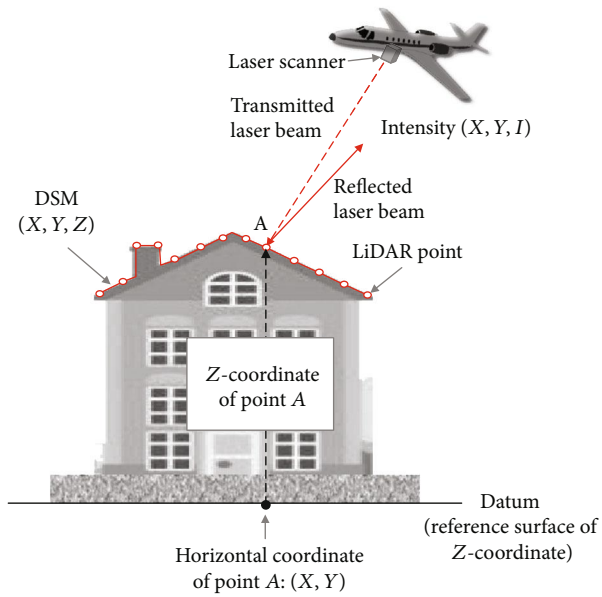


Figure 6: Intensity and DSM from LiDAR data.

experiments. The intensity values of the LiDAR data refer to the measure of the return strength of the laser pulses that are reflected from surfaces of the objects. The return intensity is based on the reflectivity of the surface and could be used as supplementary information for feature extraction and classification. DSM and intensity provide 3D geometric and topographic and physical properties, respectively. In these respects, we utilized both intensity (I) and height (Z) of the LiDAR data to train the DL models because intrinsic features of the data play an important role (see Figure 6).

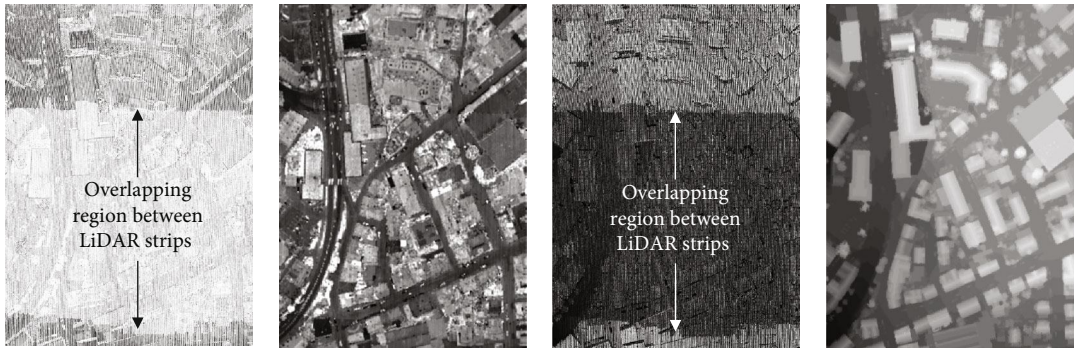
The intensity values can be visualized as black-and-white (B/W) imagery, which is referred to as an intensity map (see Figure 4(b)). However, the quality of the intensity map is poor in terms of spatial and radiometric resolution compared to the optical imagery. If the quality of the intensity maps is high and color-coded, the intensity map itself would be the true orthoimage. These facts gave us the insight to generate the true orthoimages using the LiDAR data using an image-to-image translation DL model.

2.2. LiDAR Dataset Resampling. Since each data has a different ground sample distance (GSD), further processing to make the same GSD of the data is required to train the convolutional neural network- (CNN-) based DL models. Irregularly distributed LiDAR point clouds were resampled to produce regularly gridded DSM. We used inverse distance weighted (IDW) interpolation for the resampling. Figure 7 shows the original LiDAR data of irregularly distributed point clouds, along with resampled data to the regular grid by the interpolation.

Since the point density of the Vaihingen LiDAR data is 4 pts/m^2, which is equivalent to 0.50 m GSD, the interpolation applied was inconsistent with the resolution of the other datasets (i.e., orthoimages, image matching DSM, and label data) that is 0.09 m. Overlapping is essential in collecting LiDAR data to avoid gaps between strips. It is evident that there are areas of overlap between the adjacent LiDAR strips as shown in Figures 7(a) and 7(c). The apparent visible stripe patterns in the original LiDAR data were removed after resampling.

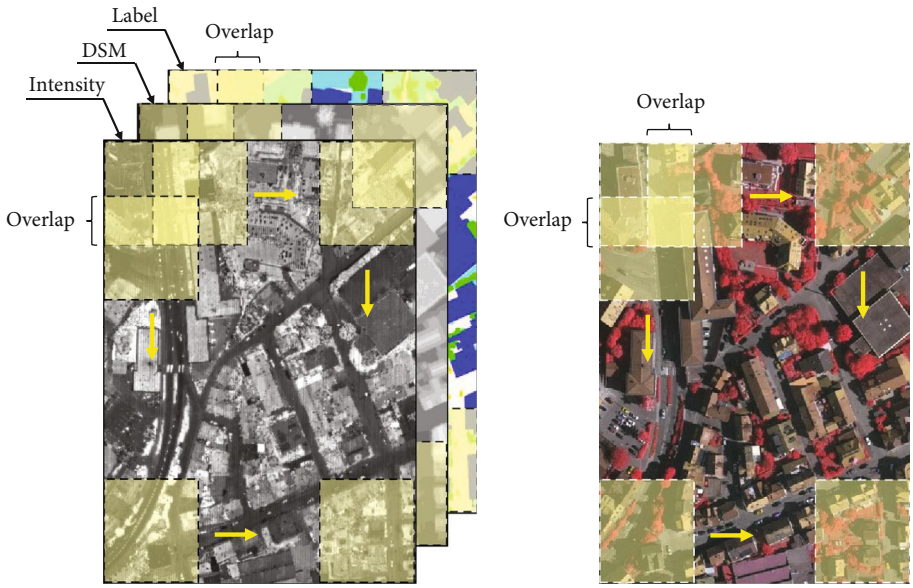
2.3. Training Dataset Partitioning. Each data was partitioned into tiles by overlapping along both X and Y directions. We created different tile sizes (256×256, 512×512, and 1024×1024 pixels) having different overlapping rates (50%, 60%, 70%, and 80%) for the various experiments (see Figure 8). Partitioning with overlapping provides substantial benefits as shown in Figure 9: (1) increasing the number of datasets (i.e., data augmentation) because the DL requires a numerous number of data and (2) improving performance of the DL models by training the same object (or region) under the various situations [25].

In many cases, the label data is used as the ground truth of the categorized land cover features for classification or



(a) (b) (c) (d)

Figure 7: Resampling LiDAR data using IDW interpolation for Vaihingen datasets: (a) original LiDAR intensity; (b) resampled LiDAR intensity; (c) original LiDAR 3D coordinates; (d) resampled LiDAR DSM.



(a) (b)

Figure 8: Partition of datasets with overlap for both length and width directions: (a) training datasets; (b) ground truth.

semantic segmentation. On the other hand, in this paper, the label data were used as the training data along with the LiDAR datasets while the orthoimages were used as ground truth since we aim to generate orthoimages by training LiDAR and label data using a GAN-based DL model.

2.4. Pix2Pix Model. GANs are algorithmic architectures that are composed of two neural networks, pitting one against the other to generate new synthetic instances of data that can pass for real data. One neural network that is the generator produces new data instances while the other that is the discriminator evaluates them for authenticity, i.e., the discriminator decides whether each instance of data that it reviews belongs to the actual training dataset or not [22]. The GAN models have been frequently used for image style transfer through the mapping from an input image to an output image. In particular, the Pix2Pix as a conditional GAN (cGAN) is well suited for translating an input image into an output image (i.e., image-to-image translation) [26]. The Pix2Pix is effective for various tasks such as semantic segmentation, generation of maps from aerial images, and colorization of B/W images [27]. As with most cGAN, Pix2Pix also requires pairs of images for training models (i.e., one for input and corresponding ground truth). The challenge of the Pix2Pix is to create paired training datasets while the CycleGANs utilize the unpaired datasets. However, if the distribution of the unpaired datasets of the CycleGANs is not stable, appropriate results might not be created. To improve the performance of the image-to-image transform, U-Net is used in the Pix2Pix. The generator is based on the U-Net architecture that has skip connection between each center of the symmetric layers [28]. Each skipped connection concatenates channels from layer i with channels from layer $n - i$, with n the total number of layers in the network.

In addition, Pix2Pix uses PatchGAN in the discriminator that works on small patches of the input image. Then,

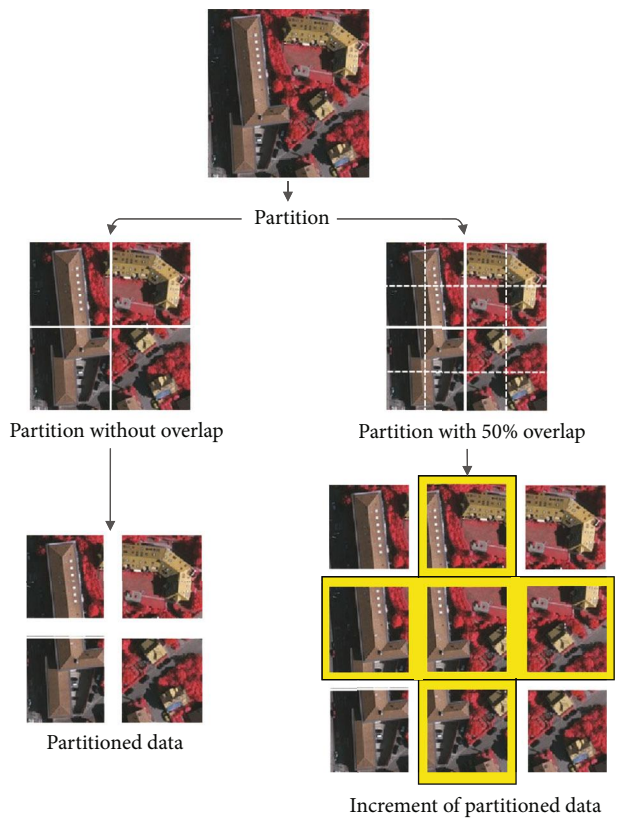


Figure 9: Demonstration of dataset increment by partitioning with 50% overlap.

averaging is performed to decide whether the entire image is real or fake. The advantage of using PatchGAN is that the discriminator captures high frequencies in the image by straining the model to focus on the local image patches. Therefore, the convolutional layer will only be receptive to the pattern at the patch size [26, 29]. Both generator and discriminator use modules of the form "convolution-batch normalization-activation function." The number of the convolution layers in the generator depends on the size of the input images because the convolutions are performed up to 1×1 pixel image. The image size is reduced by half since stride is 2 for pooling after convolution. For an example of a 512×512 image (i.e., $2^9 \times 2^9$ pixels), nine convolution layers are required. Figure 10 shows the Pix2Pix model and workflow of the proposed method. Figures 11 and 12 show the generator and discriminator of the Pix2Pix model, respectively.

ReLU is the most commonly used activation function in CNNs. Specifically, Pix2Pix utilizes the plane ReLU (see Figure 13(a) and Equation (1)) in the encoder and the leaky ReLU with slope 0.2 (see Figure 13(b) and Equation (2)) in the decoder of the generator, while leaky ReLUs also with slope 0.2 are used in the discriminator. It is known that using leaky ReLU is beneficial because of fixing the dying ReLU problem and speeding up the training [30, 31]:

$$\text{ReLU} : \phi(x) = \begin{cases} 0, & \text{for } x \leq 0, \\ x, & \text{for } x > 0, \end{cases} \tag{1}$$

$$\text{Leaky ReLU} : \phi(x) = \begin{cases} ax, & \text{for } x \leq 0, \\ x, & \text{for } x > 0, \end{cases} \tag{2}$$

where a is usually a small constant value (e.g., $a = 0.2$).

2.5. Pix2Pix Model Training Strategies. The Vaihingen datasets were used for training, validation, and testing of the DL models. The Potsdam datasets were used as unseen datasets to evaluate the models trained with the Vaihingen datasets. In order to analyze the performance of the DL models and eventually to find the optimal case, we carried out intensive experiments under various conditions in terms of the data type, data combination, partitioned tile size, and overlap rate. As for the label, each color of the original label images represents a specific land cover feature such as blue for buildings, green for trees, cyan for low vegetation, yellow for cars, and red for water bodies. The original color label images as well as converted to B/W label images were used for training to compare results from each label image (see Figure 14).

The training schemes were categorized based on the following cases (see Figure 15).

(1) Individual datasets: [intensity], [DSM], [color label], and [B/W label]

(2) Combined datasets: [intensity+DSM], [intensity+B/W label], [DSM+B/W label], and [intensity+DSM+B/W label]

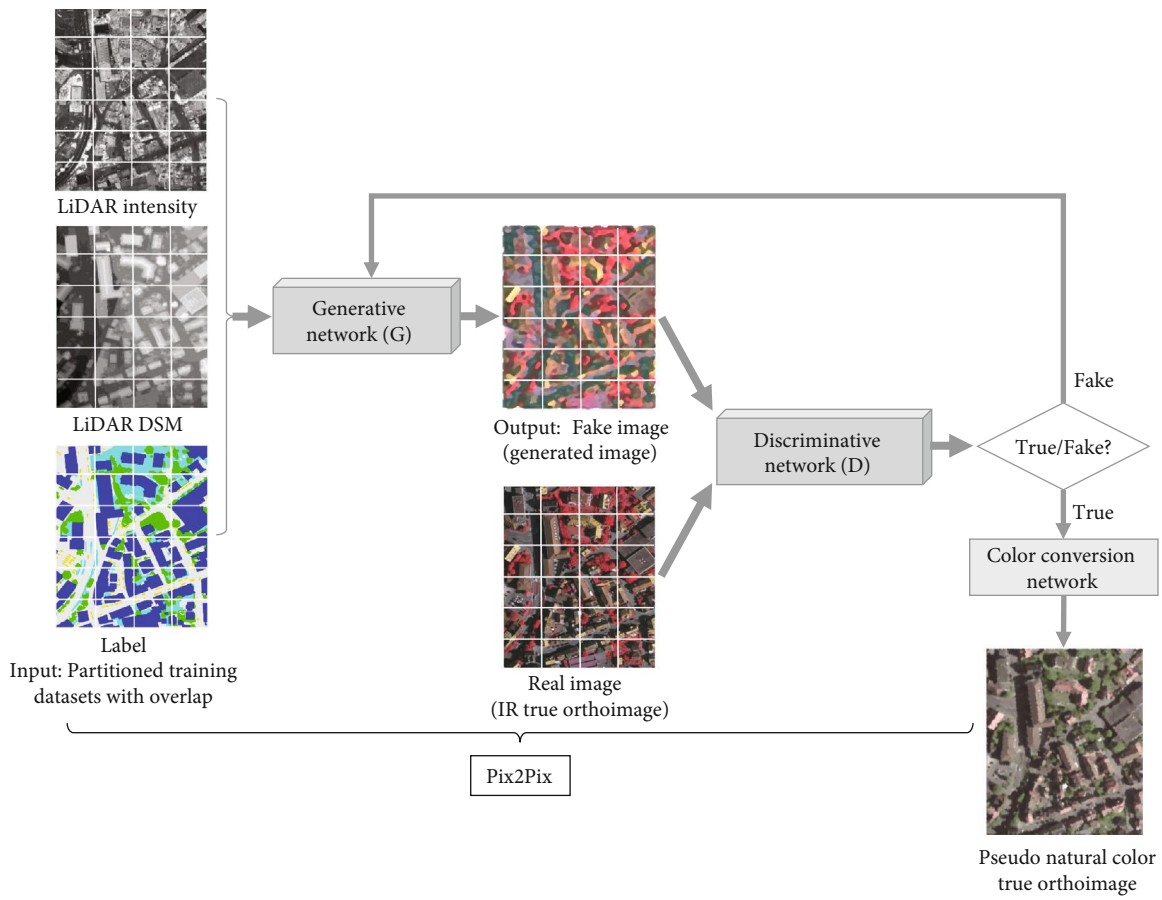


Figure 10: Workflow of the proposed method for true orthoimage generation using the DL model.

(3) Tile size: 256×256, 512×512, and 1024×1024

(4) Overlap rate: 50%, 60%, 70%, and 80%

Analyzing a variety of the results through the intensive experiments under the various conditions might be helpful to determine the optimal case. The test data and unseen data from the Vaihingen and Potsdam datasets, respectively, were inputted into the trained models. The optimal trained models would be used to the datasets from other areas (see Figure 16).

2.6. Infrared Image to Natural Color Image Conversion. A natural color (i.e., normal color) composite is an image displaying a combination of the primary visible red (R), green (G), and blue (B) components (or bands) to the corresponding channels that resemble what humans observe. On the other hand, an IR image is composed of the near-IR (NIR), R, and G bands. Since the IR images are false color composites, it is often necessary to convert the IR images to natural color images, so-called pseudonatural color images [32]. We performed the spectral conversion from IR to RGB color image using a simple neural network. Each IR true orthoimage from the test data of the Vaihingen area and unseen data of the Potsdam area was converted to the natural color images using a neural network model (see Figure 17) based on the spectral conversion equations:

$$\begin{bmatrix} \mathrm{R} \\ \mathrm{G} \\ \mathrm{B} \end{bmatrix} = \begin{bmatrix} w_{11} & w_{21} & w_{31} \\ w_{12} & w_{22} & w_{32} \\ w_{13} & w_{23} & w_{33} \end{bmatrix} \begin{bmatrix} \mathrm{NIR} \\ \mathrm{R} \\ \mathrm{G} \end{bmatrix}, \tag{3}$$

where $[\mathrm{R}, \mathrm{G}, \mathrm{B}]^T$ is output bands of a pseudonatural color image and $[\mathrm{NIR}, \mathrm{R}, \mathrm{G}]^T$ is input bands of an IR image. w_{ij} are weights of the spectral conversion matrix to be determined by training the neural network.

The target images (i.e., natural color images) for the color conversion network training were taken from the aerial imagery of Microsoft Bing Maps. In general, training of the neural network model starts with random initial weights. However, performance of the training would be improved if reasonable initial weights are available. We adopted the pseudo color conversion coefficients used in Satellite Pour l'Observation de la Terre (SPOT) multispectral satellite imagery as initial weights that are $w_{11} = 0.1$, $w_{21} = 0.9$, $w_{31} = 0.0$, $w_{12} = 0.3$, $w_{22} = 0.0$, $w_{32} = 0.7$, $w_{13} = 0.1$, $w_{23} = 0.1$, and $w_{33} = 0.8$. The sum of the weights for each raw should be

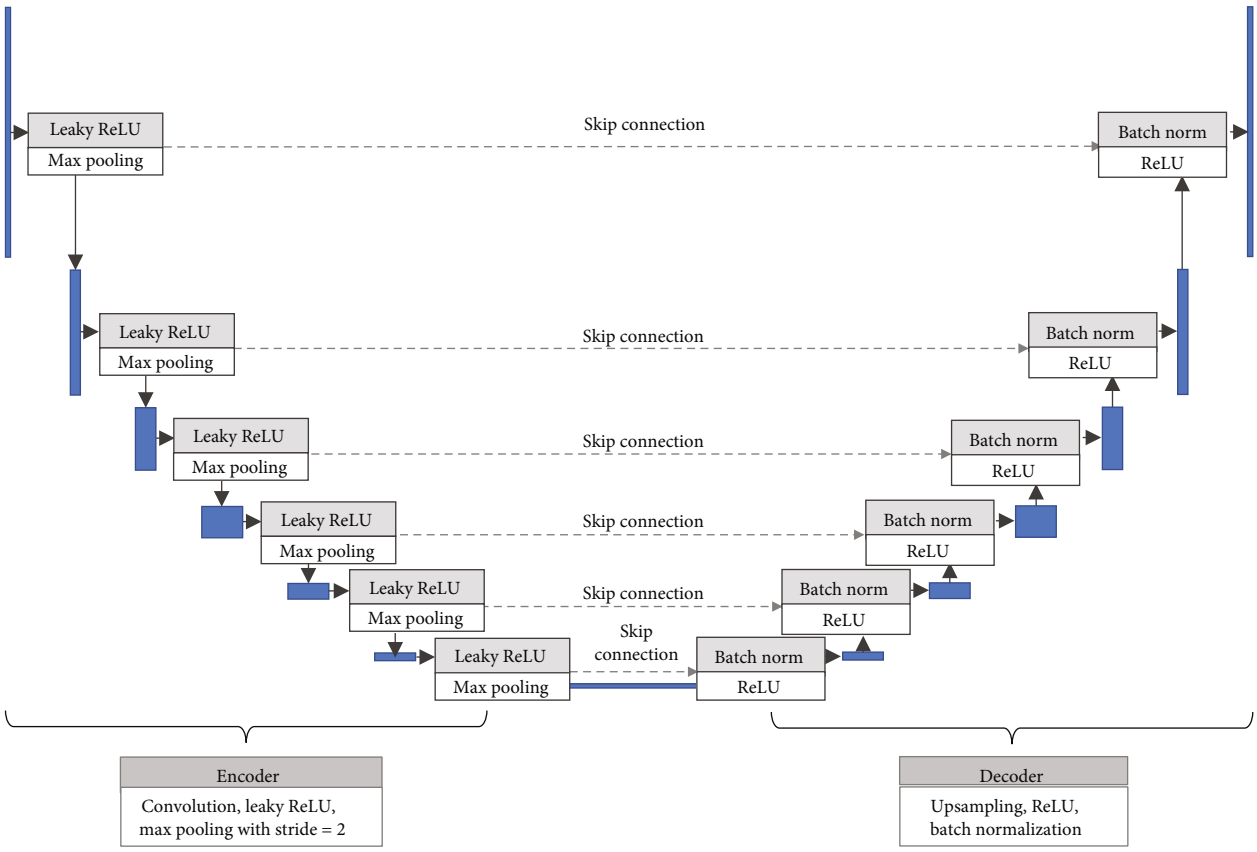


Figure 11: Generative network of the Pix2Pix model.

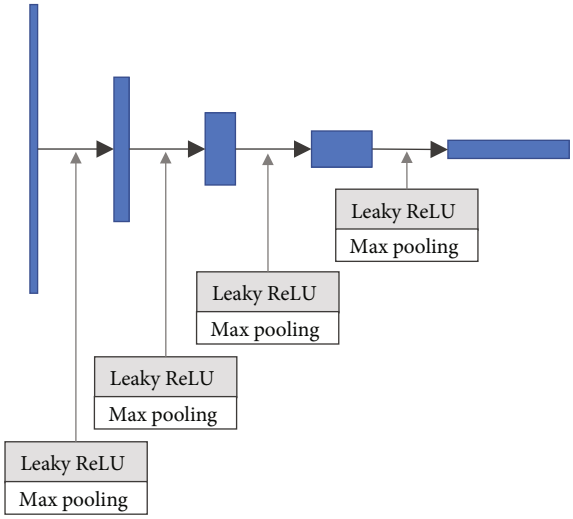


Figure 12: Discriminative network of the Pix2Pix model.

1 because the brightness of the images should not change. The weights were updated through the backpropagation.

3. Results and Analysis

As described in the training strategy, experiments using different types of training datasets under various conditions were conducted to generate true orthoimages and to analyze performance of the DL models. The performance was evaluated with plots of the loss for epoch and the Fréchet inception distance (FID) and structural similarity index measure (SSIM) that have been frequently used as evaluation measures of the GAN-based models [33–36]. FID is a metric to assess the quality of the images created by the generator. It

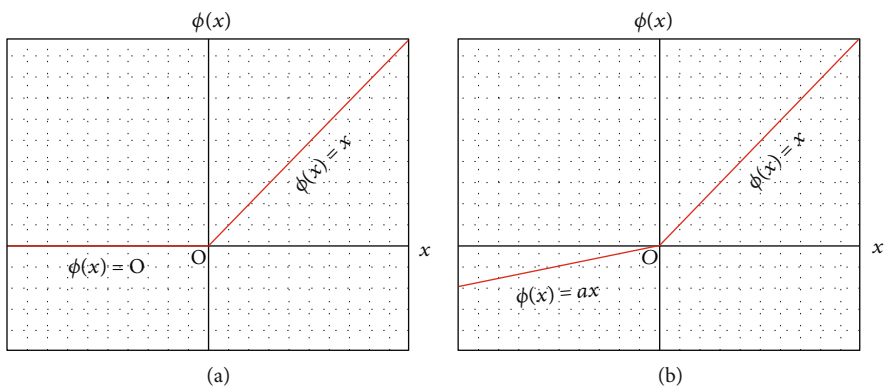


Figure 13: Activation functions used in Pix2Pix: (a) ReLU; (b) leaky ReLU.

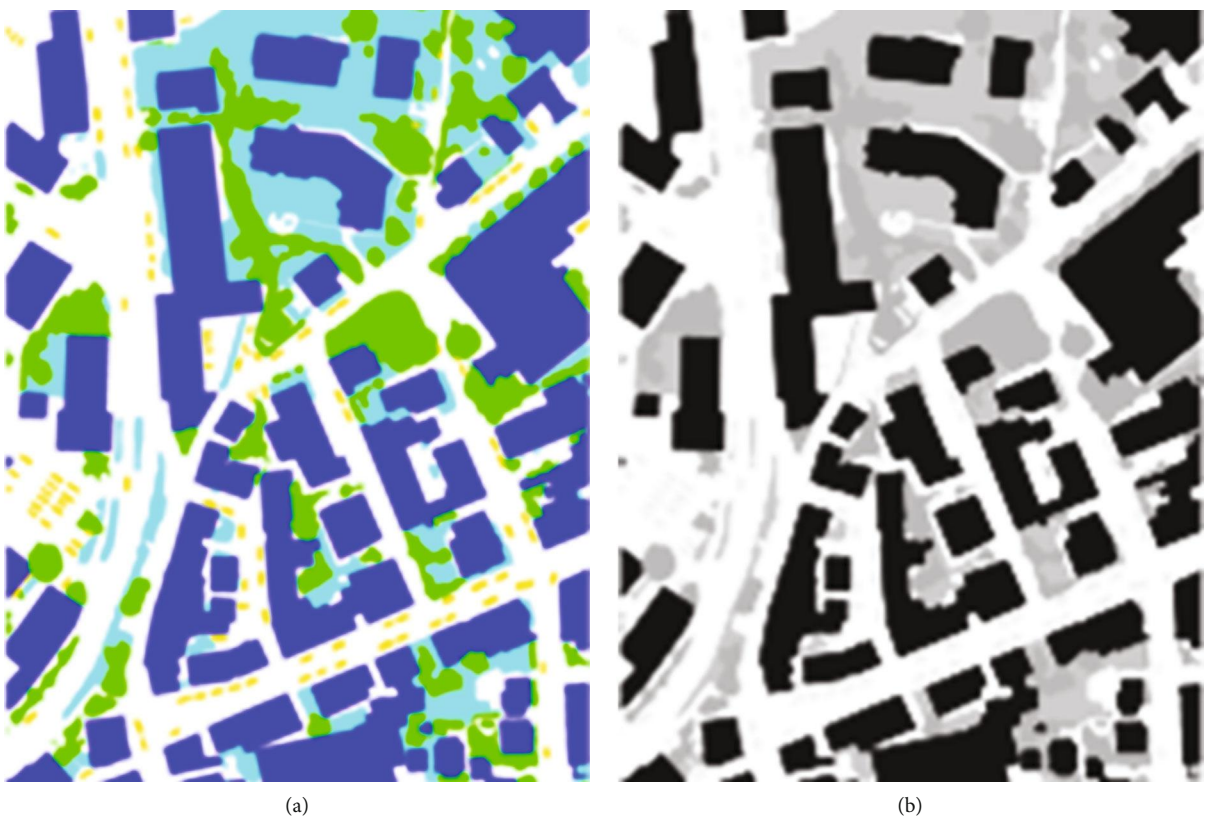


Figure 14: Label datasets: (a) color label; (b) B/W label.

measures the distance between feature vectors calculated for the generated images (i.e., fake images) and real images (i.e., ground truth) based on the mean and covariance. A lower FID indicates better quality of the generated image. The SSIM is a perception-based method and has been used for measuring the similarity between two images based on an initial uncompressed or distortion-free image as reference. The SSIM is also calculated using the mean and covariance as FID. The maximum value of the SSIM is 1, which indicates that the two images are perfectly structurally similar, while a value of 0 indicates no structural similarity. However, analyzing or assessing the quality of images based on the statistical or quantitative metric might be insufficient since such metrics represent overall quality without local details. Therefore, we carried out visual inspection for each case.

Training of the Pix2Pix models was performed with a maximum of 200 epochs. The learning rates were set to 0.0002 for the 1st to 100th epoch and then linearly decreased up to 0.000002 after the 101st to 200th epoch. The loss graphs for LiDAR intensity, LiDAR DSM, and label image are shown in Figure 18. The loss for each data rapidly decreased in the early epochs. In our experiments, the overall losses showed a similar pattern, i.e., the losses were getting stable and converged after the 60th epoch.

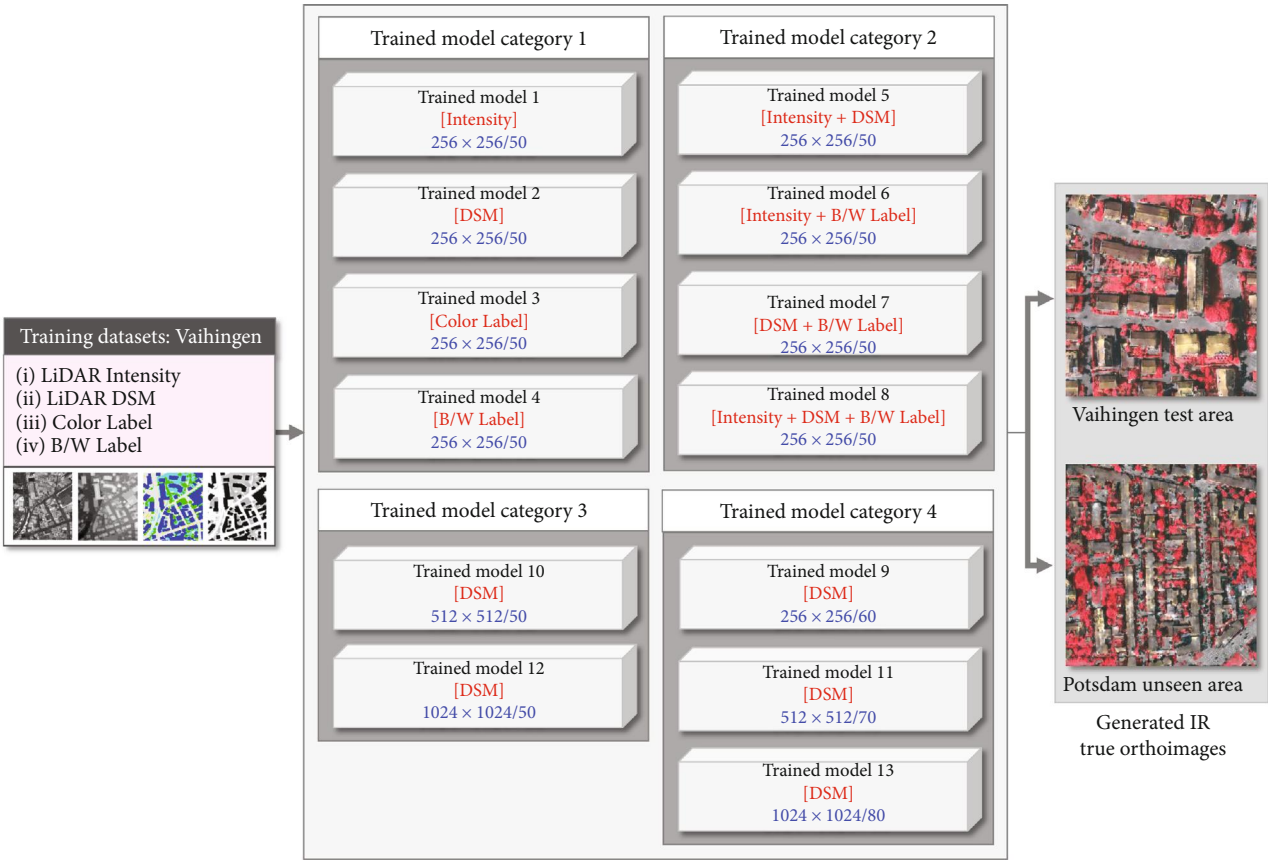


Figure 15: True orthoimage generation for test and unseen datasets using trained models. The numbers in each trained model represent tile size and overlap rate.

The loss graph in terms of the root mean squared error (MSE) for the IR to natural color conversion is presented in Figure 19.

The main purpose of the experiments is to determine the optimal training condition by analyzing the influences of the data characteristics, tile size and overlap of the data partitioning, and combinations of the different types of data for the training. The results and analyses of the generated orthoimages and pseudonatural color images for the test and unseen datasets are presented in various aspects.

3.1. Results of True Orthoimage Generation for Test Datasets. The first series of experiments is for examining the results from training each data independently with fixed partitioned tile size (256 × 256) and the overlap rate (50%) (see Figures 20–23). Since the LiDAR intensity has low radiometric resolution and noisy stripe pattern, the result is not satisfactory. Compared to the real true orthoimage (see Figure 20(a)), it is difficult to distinguish between buildings and trees (see Figure 20(b)). The result of the LiDAR DSM is better than the result of the intensity because DSM contains geometric information such as shape and surface slope of the objects (see Figure 20(c)). The label is an image of classified land cover features, and each class is represented with an arbitrary color. The experiments were conducted with color label (see Figure 21(b)) and B/W label (see Figure 21(c)) to investigate the effect of the label color scheme. The results show that there is no significant difference between color and B/W labels (see Experiments 1 to 4 in Table 2). The label has different characteristics from the LiDAR intensity and DSM. The advantage of using the label for training DL model is to produce clear and sharp boundaries of the buildings (see Figures 21(b) and 21(c)) while LiDAR data does not produce clear boundaries due to resampling (see Figures 20(b) and 20(c)).

The second series of experiments is for exploring the effect of combining individual data. We carried out experiments with all possible combinations of the datasets. Combining LiDAR intensity and LiDAR DSM (see Figure 22(b)) shows improved results compared with LiDAR intensity only (see Figure 20(b)). Combining LiDAR intensity and B/W label (see Figure 22(c)) provided better results than combining LiDAR intensity and LiDAR DSM. Moreover, combining LiDAR DSM and B/W label (see Figure 23(b)) resulted in improved true orthoimages. In this case, the shape of the roofs is identifiable, particularly the gable roofs.

Finally, all three data (i.e., LiDAR intensity, LiDAR DSM, and B/W label) were combined (see Figure 23(c)). The result was improved compared to combining LiDAR DSM and B/W label (see Figure 23(b)). Therefore, combining all datasets brought synergistic effect to improve the performance of the training. This fact is also shown in Table 2. Experiment 8,

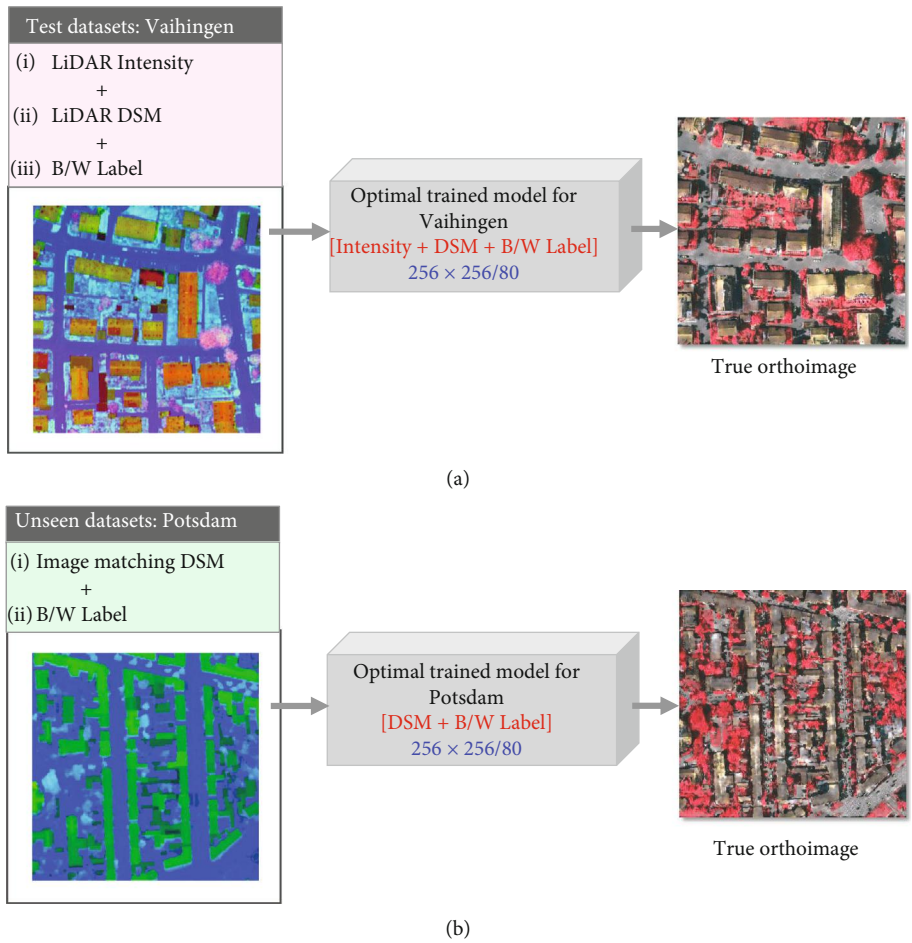


Figure 16: True orthoimages from optimal trained models: (a) Vaihingen test datasets; (b) Potsdam unseen datasets.

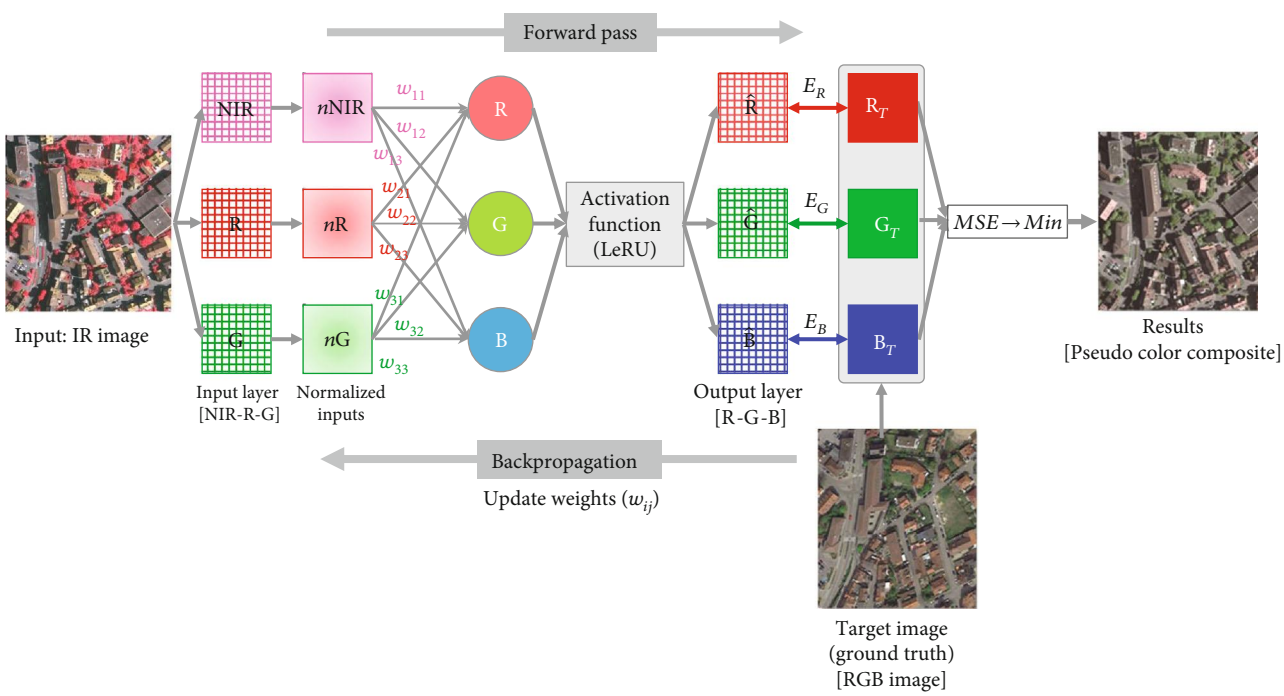


Figure 17: A neural network for IR to pseudonatural color image conversion.

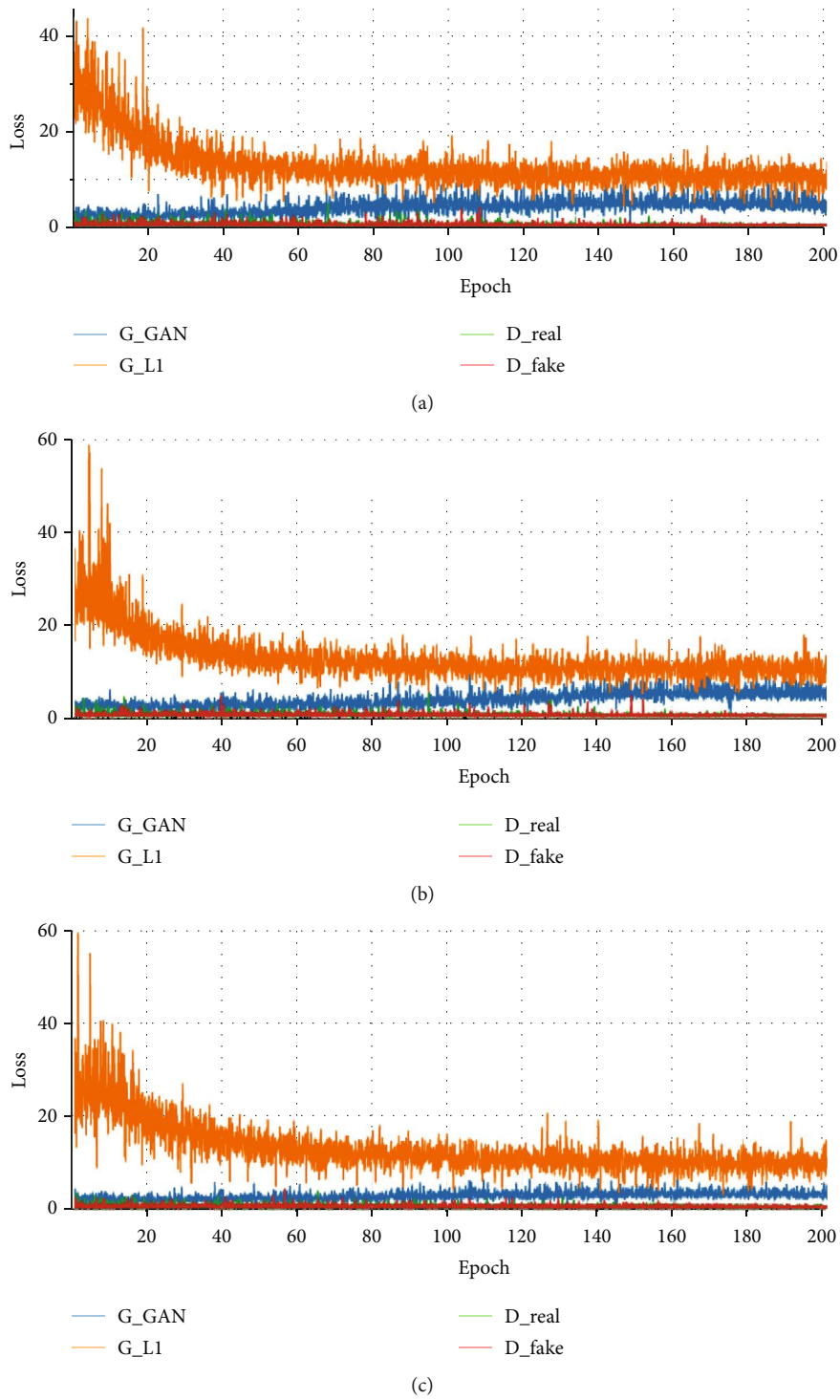


Figure 18: Loss graphs of Pix2Pix model training: (a) LiDAR intensity; (b) LiDAR DSM; (c) label.

which is combining all datasets, provides the smallest FID (i.e., 131.83) and the largest SSM (i.e., 0.0.434) among the individual dataset training and the combining training cases (see Experiments 1 to 8 in Table 2).

The third series of experiments is for comparing the influence of the overlap in data partitioning. This experiment was carried out using LiDAR DSM. The results were analyzed for 50% and 60% overlap with a fixed tile size of 256 × 256. For the visual comparison of the results, the real true orthoimage is provided in Figure 24(a). The larger overlap results in dataset augmentation that is required for training DL models. 60% overlap (see Figure 24(c)) provides better results than 50% overlap (see Figure 24(b)). The roof shapes are improved by increasing overlap. Experiment 2 (i.e., FID =

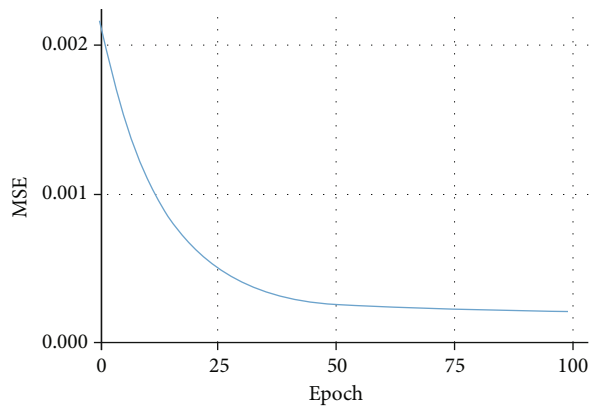


Figure 19: Loss graph of the color conversion network model.

231.74 and SSIM = 0.401) and Experiment 9 (i.e., FID = 197.63 and SSIM = 0.427) in Table 2 show that the larger overlap provides better results. More experiments for the tile size of 512 × 512 with 50% and 70% overlaps (see Figures 25(b) and 25(c)) and for the file size of 1024 × 1024 with 50% and 80% overlaps (see Figures 26(b) and 26(c) were performed. Similar results were obtained for all cases (see FID and SSIM of Experiments 10 and 11 and Experiments 12 and 13 in Table 2).

The fourth series of experiments are aimed at finding the optimal tile size. The real true orthoimages are shown in Figures 25(a) and 26(a) to compare the results from the experiments under different conditions. The results for different tile sizes are presented in Figures 20(c), 25(b), and 26(b) corresponding to the tile size of 256 × 256, 512 × 512, and 1024 × 1024, respectively. The results look similar to each other even though the tile size of 256 × 256 might be the best one according to both FID (i.e., 231.74) and SSIM (i.e., 0.401) in Table 2. On the other hand, FID and SSIM for tile size of 512 × 512 are 316.40 and 0.388, respectively. FID and SSIM for tile size of 1024 × 1024 are 400.89 and 0.363, respectively.

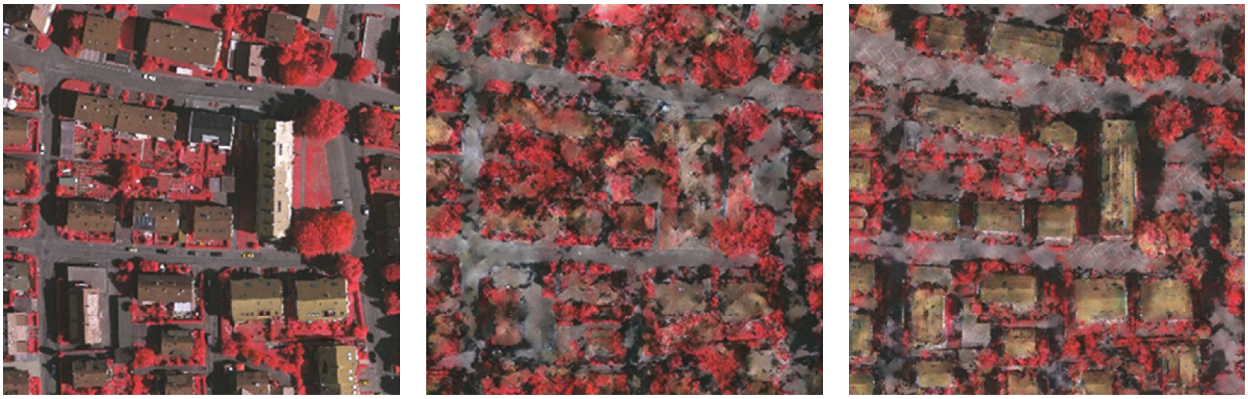
Based on the experiments considering various aspects including tile size and overlap rate, characteristics of the training datasets, quality evaluation metrics (i.e., FID and SSIM), and visual inspection, using all data (LiDAR intensity, LiDAR DSM, and B/W label) with 256 × 256 tile size and 80% overlap provided the best result with the smallest FID (i.e., 117.67) and the largest SSIM (i.e., 0.505) (see Figure 27(b) and Experiment 15 in Table 2). The second-best result was obtained by combining DSM and B/W label (see Experiment 14 in Table 2) with FID of 132.35 and SSIM of 0.436, and this trained model was applied to the unseen data of the Potsdam since the intensity data is not available in the Potsdam datasets. Figure 28 represents Table 2 with bar graphs for better analysis and interpretation of the results.

Image quality evaluation considers various elements such as color, tone, brightness, and contrast. Furthermore, evaluation of the true orthoimages is a complicated issue because assessment of the appropriate recovery of the occlusion areas is essential. For this reason, the difference between the real true and generated true orthoimages might be objective and both quantitative and qualitative ways of evaluating orthoimages. Figure 27 visualizes the difference between a real true orthoimage and a generated true orthoimage of the test datasets. The darker areas in Figure 27(c) represent regions of significant difference between the orthoimages, indicating that the image is incorrectly generated, such as the color and the tone of some buildings. In particular, the incorrectly generated regions are along the building boundaries due to the interpolation for resampling of the LiDAR data. However, there are limitations to evaluate the quality of the generated image based on the difference since the disparity might not reflect all factors of the image quality.

3.2. Results of True Orthoimage Generation for Unseen Datasets. The results of the generated true orthoimages for the Potsdam unseen datasets are presented in Figures 29–34. The intensity data could not be used, and the image matching DSM was used in lieu of the LiDAR DSM since the LiDAR data is not available in Potsdam datasets. The results of the unseen datasets were obtained by reusing the pretrained models that were trained with Vaihingen training datasets. In other words, neither training nor fine tuning of the hyperparameters was performed to generate the true orthoimage. Therefore, the quality of the results from unseen datasets is not expected to be better than the case of the test datasets. Analyses of the results from the unseen datasets might not be consistent with the results from the test datasets because of the respective datasets belonging to different places with different characteristics such as image tone, image scale, topology, and object shape. The series of experiments for the unseen datasets are the same order as the test datasets, i.e., analysis of the results from (1) the use of each data independently, (2) combination of individual data, (3) different overlap rates, and (4) different tile sizes.

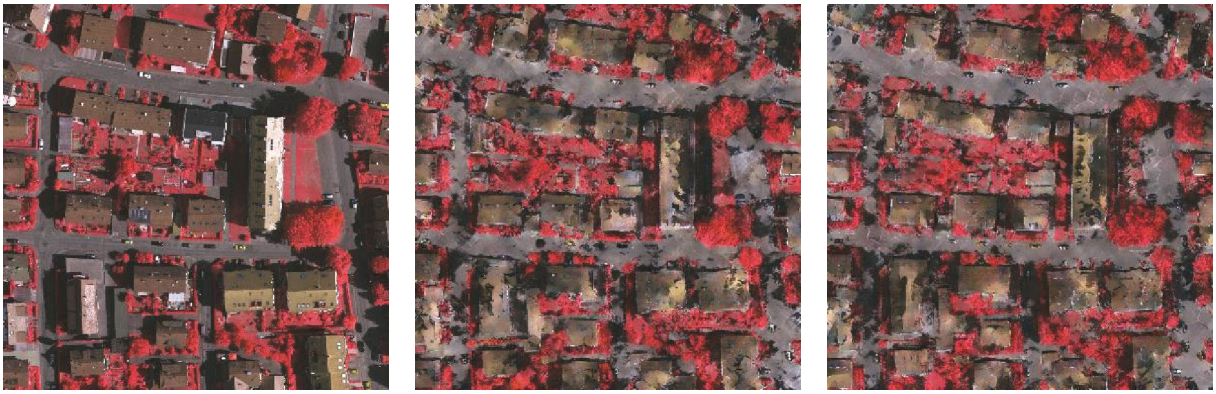
The result using DSM (see Figure 29(b)) was visually better than using other data (i.e., color label and B/W label). Also, both FID and SSIM indicate that using the DSM is better than using the other data. The smallest FID (i.e., 451.39) and the largest SSIM (i.e., 0.375) were obtained from DSM (see Experiments 1, 2, and 3 in Table 3). In addition, using the combined dataset of DSM and B/W label did not improve the results (see Figure 30(c) and Experiment 4 in Table 3).

As for the influence of the overlap, we compared the overlap rates between 50% and 60% with tile size of 256 × 256 (see Figures 31(b) and 31(c) and Experiments 1 and 5 in Table 3) for DSM. Even though it is not easy to tell the difference between the results according to the FID and SSIM, 60% overlap resulted in more or less clear edges of the buildings than 50% overlap by visual inspection. Similar results were obtained for the overlap rates between 50% and 70% with tile size of 512 × 512 (see Figures 32(b) and 32(c) and Experiments 6 and 7 in Table 2). However, for the case of overlap rates of 50% and 80% with tile size of 1024 × 1024, 80% overlap produced better results than 50% overlap (see Figures 33(b) and 33(c) and Experiments 8 and 9 in Table 2).



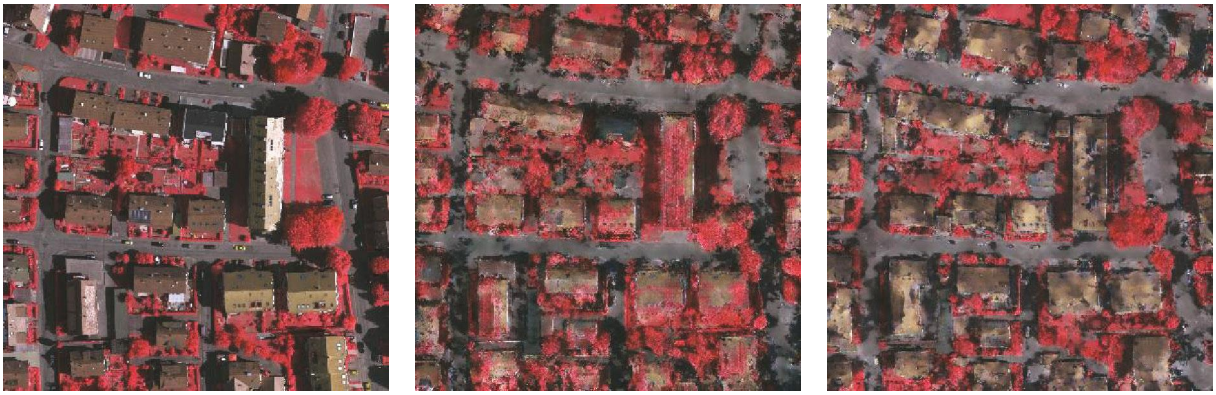
(a) (b) (c)

Figure 20: Generated true orthoimages with 256 × 256 tile size and 50% overlap: (a) real true orthoimage; (b) LiDAR intensity; (c) LiDAR DSM.



(a) (b) (c)

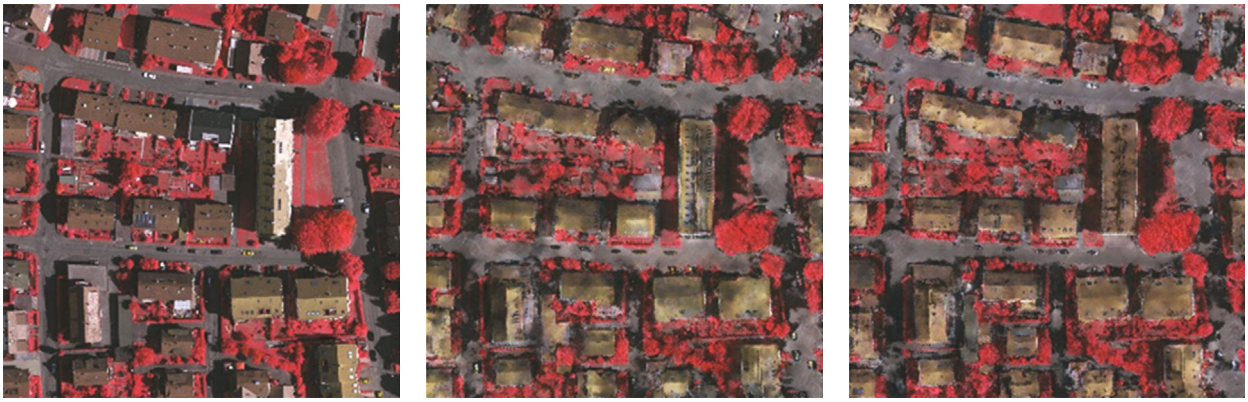
Figure 21: Generated true orthoimages with 256 × 256 tile size and 50% overlap: (a) real true orthoimage; (b) color label; (c) B/W label.



(a) (b) (c)

Figure 22: Generated true orthoimages with 256 × 256 tile size and 50% overlap: (a) real true orthoimage; (b) LiDAR intensity and LiDAR DSM; (c) LiDAR intensity and B/W label.

The best result was obtained from the model trained using all of the data (i.e., LiDAR intensity, LiDAR DSM, and B/W label) of the Vaihingen datasets with 256 × 256 and 80% overlap. Nevertheless, this trained model could not be applied to the unseen data because Potsdam datasets do not include LiDAR data. Therefore, the model trained using image matching DSM and B/W label, which is the second-best case of the experiments (see Experiment 14 in



(a) (b) (c)

Figure 23: Generated true orthoimages with 256 × 256 tile size and 50% overlap: (a) real true orthoimage; (b) LiDAR DSM and B/W label; (c) LiDAR intensity; LiDAR DSM and B/W label.

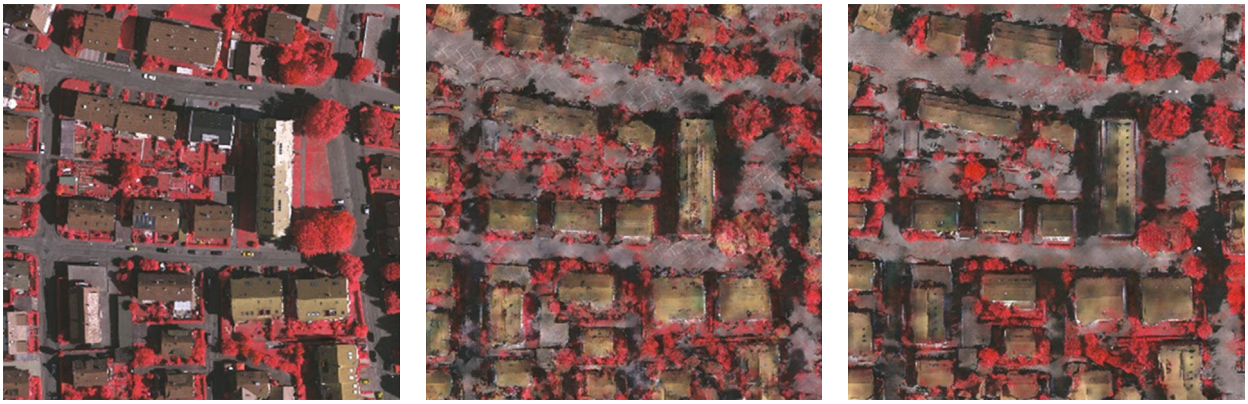
Table 2: Evaluation of training performance for Vaihingen test datasets.

Experiment	Training data type	Tile size (pixel)	Overlap (%)	FID	SSIM
1	Intensity	256 × 256	50	353.18	0.371
2	DSM	256 × 256	50	231.74	0.401
3	Color label	256 × 256	50	222.51	0.379
4	B/W label	256 × 256	50	194.76	0.383
5	Intensity+DSM	256 × 256	50	250.44	0.396
6	Intensity+B/W label	256 × 256	50	218.01	0.421
7	DSM+B/W label	256 × 256	50	214.77	0.431
8	Intensity+DSM+B/W label	256 × 256	50	131.81	0.434
9	DSM	256 × 256	60	197.63	0.427
10	DSM	512 × 512	50	258.59	0.395
11	DSM	512 × 512	70	174.79	0.406
12	DSM	1024 × 1024	50	369.21	0.368
13	DSM	1024 × 1024	80	148.41	0.377
14	DSM+B/W label	256 × 256	80	132.35	0.436
15	Intensity+DSM+B/W label	256 × 256	80	117.67	0.505

Table 2), was applied to the unseen data. As portrayed in Table 3, the smallest FID and SSIM values of 373.03 and 0.387, respectively, were produced in Experiment 10 which corresponds to the approach that produced the best outcome. The result is shown in Figure 34(b), and the difference between the real true orthoimage and generated true orthoimage is visualized (see Figure 34(c)). Figure 35 represents Table 3 with bar graphs.

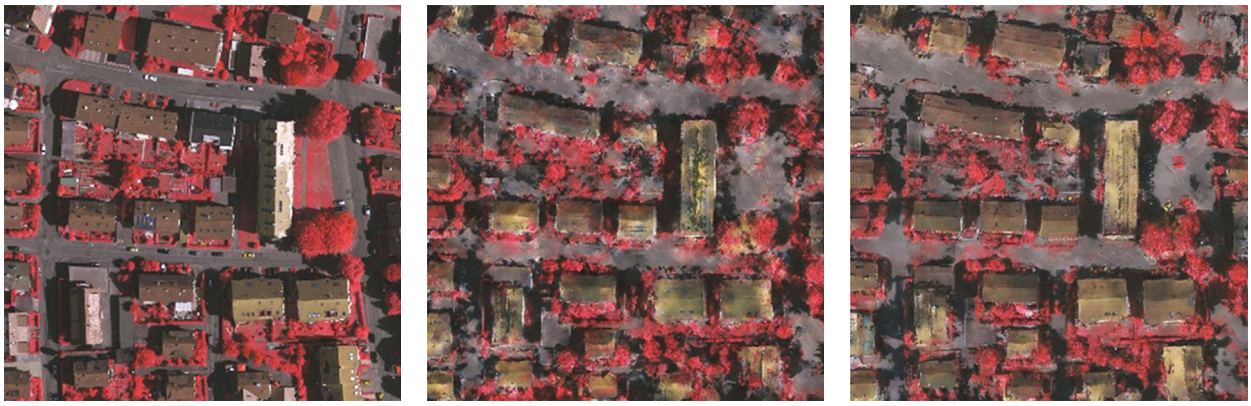
3.3. Results of Pseudocolor Image Generation. Conversion of the IR image to the pseudonatural image depends mostly on the target image used as ground truth such as time (or season), illumination condition, and camera characteristics. Since IR is beyond the visible spectrum, CIR images are displayed as the false color composite. The invisible NIR light of CIR can be visible by shifting the NIR light and the primary colors over. NIR wavelengths become visible as red while red wavelengths appear as green and green as blue. Blue wavelengths are shifted out of the visible portion of the spectrum, and so they appear as black. Most of the vegetation appears red while water generally appears black with artificial structures like buildings and roads showing as a light blue-green on the CIR images.

We carried out the color conversion using a neural network by training natural color images. The results of the test and unseen images are shown in Figures 36 and 37, respectively. The initial and final weights after training the neural networks are presented in Table 4. Regardless, the results of the pseudocolor images might not be the same as the natural color image of the primary visible red, green, and blue. In addition, various color schemes could be utilized by applying different weight constraints of the neural network model.



(a) (b) (c)

Figure 24: Generated true orthoimages with 256 × 256 tile size and 50% and 60% overlap: (a) real true orthoimage; (b) LiDAR DSM with 50% overlap; (c) LiDAR DSM with 60% overlap.

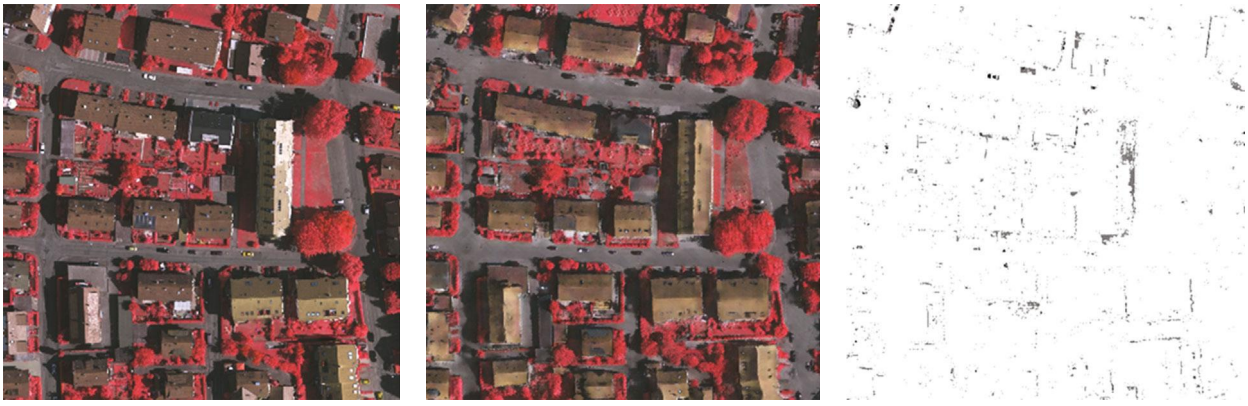


(a) (b) (c)

Figure 25: Generated true orthoimages with 512 × 512 tile size and 50% and 70% overlap: (a) real true orthoimage; (b) LiDAR DSM with 50% overlap; (c) LiDAR DSM with 70% overlap.

(a) (b) (c)

Figure 26: Generated true orthoimages with 1024 × 1024 tile size and 50% and 80% overlap: (a) real true orthoimage; (b) LiDAR DSM with 50% overlap; (c) LiDAR DSM with 80% overlap.



(a) (b) (c)

Figure 27: Generated true orthoimages with 256 × 256 tile size and 80% overlap: (a) real true orthoimage; (b) LiDAR intensity, LiDAR DSM, and B/W label; (c) difference image between (a) and (b).

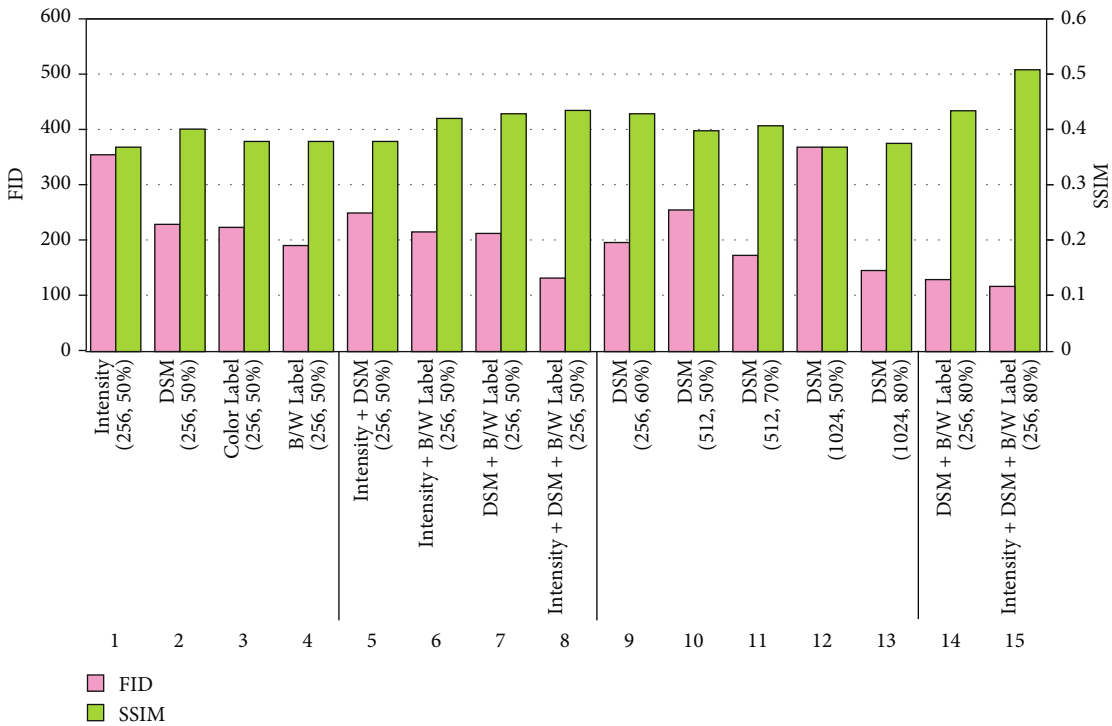
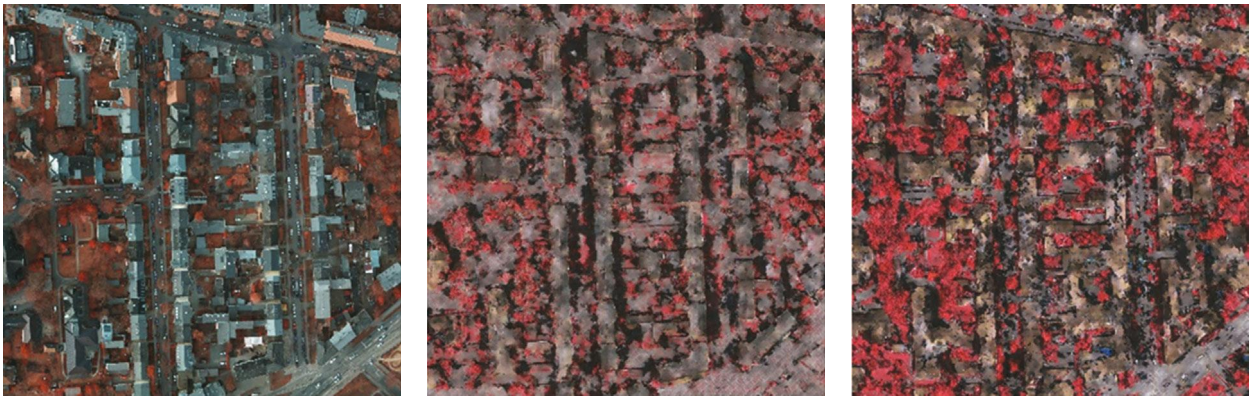


Figure 28: FID and SSIM of Vaihingen test datasets.

4. Conclusions and Discussions

Demand for true orthoimages is increasing since true orthoimages are important and useful products in the geospatial information system. However, generating true orthoimages is a challenging task using the conventional methods due to the complicated procedures involved in the process such as occlusion detection and recovery. This paper proposed a DL method for true orthoimage generation by utilizing the GAN-based Pix2Pix model. The crucial issues in DL are to improve the training efficiency and performance. The important factors in DL are characteristics, quantity, and quality of the training data. In this regard, the LiDAR data is one of the most essential data because LiDAR data provides precisely georeferenced 3D coordinates and intensity information based on the orthogonal projection. Therefore, both geometric and physical information could be utilized to improve the performance of the DL model training especially for generating true orthoimages. Intensive and systematic experiments were performed to find the optimal condition of the training datasets to generate true orthoimages including data partition size, overlap, and integration of multimodal data (i.e., LiDAR intensity, DSM, and label).

In order to train the CNN-based DL models such as GAN, the LiDAR data is to be resampled by interpolation. However, the boundaries of the building (i.e., break lines or

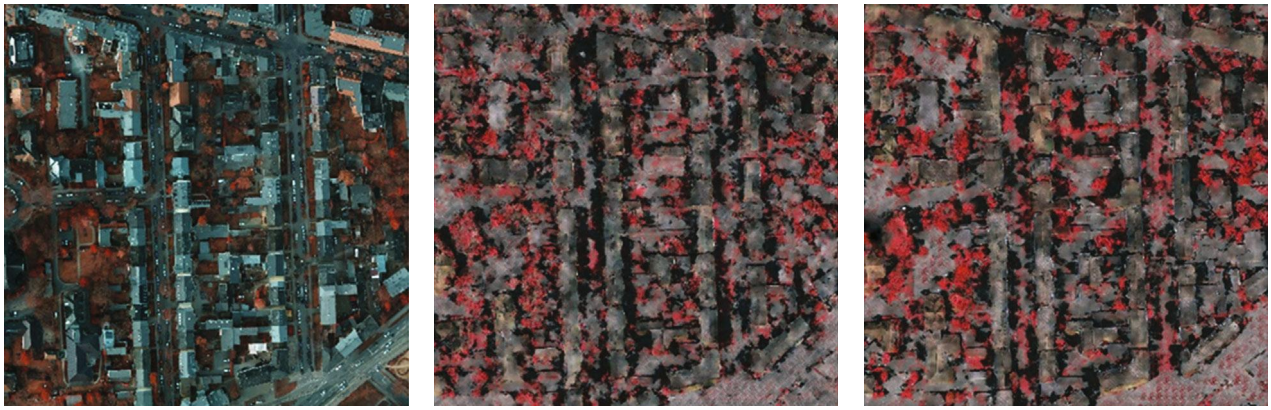


(a) (b) (c)

Figure 29: Generated true orthoimages with 256 × 256 tile size and 50% overlap: (a) real true orthoimage; (b) image matching DSM; (c) color label.

(a) (b) (c)

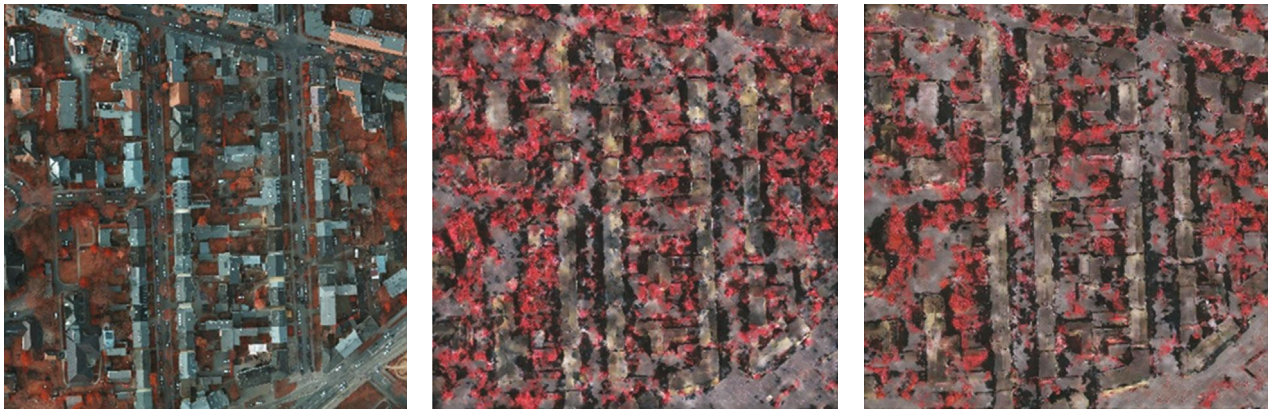
Figure 30: Generated true orthoimages with 256 × 256 tile size and 50% overlap: (a) real true orthoimage; (b) B/W label; (c) image matching DSM and B/W label.



(a) (b) (c)

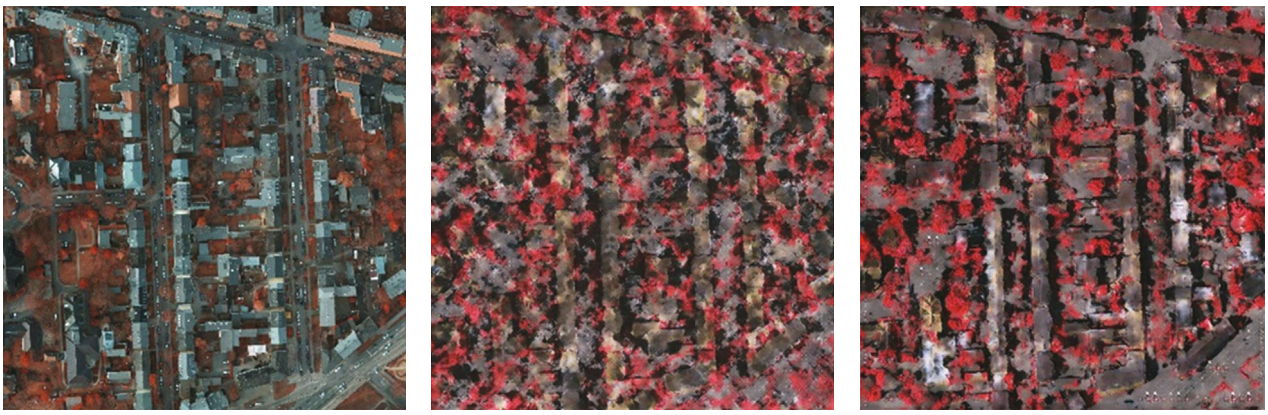
Figure 31: Generated true orthoimages with 256 × 256 tile size and 50% and 60% overlap: (a) real true orthoimage; (b) image matching DSM with 50% overlap; (c) image matching DSM with 60% overlap.

form lines) are not preserved due to the interpolation. In this regard, we carried out several experiments with different combinations of the training datasets to examine the contribution of each individual data. We concluded that the label contributes to preserving the boundaries of the objects. Therefore, combination of the multimodel datasets having



(a) (b) (c)

Figure 32: Generated true orthoimages with 512 × 512 tile size and 50% and 70% overlap: (a) real true orthoimage; (b) image matching DSM with 50% overlap; (c) image matching DSM with 70% overlap.



(a) (b) (c)

Figure 33: Generated true orthoimages with 1024 × 1024 tile size and 50% and 80% overlap: (a) real true orthoimage; (b) image matching DSM with 50% overlap; (c) image matching DSM with 80% overlap.

(a) (b) (c)

Figure 34: Generated true orthoimages with 256 × 256 tile size and 80% overlap: (a) real true orthoimage; (b) image matching DSM and B/W label; (c) difference image between (a) and (b).

different characteristics could result in mutual compensation of each individual data. In consequence, the synergistic effect during training DL models could be achieved.

Human beings have the ability to integrate various information to improve the learning process and the training performance on numerous applications. The same concept

Table 3: Evaluation of training performance for Potsdam unseen datasets.

Experiment	Training data type	Tile size (pixel)	Overlap (%)	FID	SSIM
1	DSM	256 × 256	50	451.39	0.375
2	Color label	256 × 256	50	510.94	0.352
3	B/W label	256 × 256	50	515.57	0.363
4	DSM+B/W label	256 × 256	50	519.37	0.384
5	DSM	256 × 256	60	515.69	0.377
6	DSM	512 × 512	50	498.72	0.368
7	DSM	512 × 512	70	442.75	0.377
8	DSM	1024 × 1024	50	526.68	0.346
9	DSM	1024 × 1024	80	492.61	0.370
10	DSM+B/W label	256 × 256	80	373.03	0.387

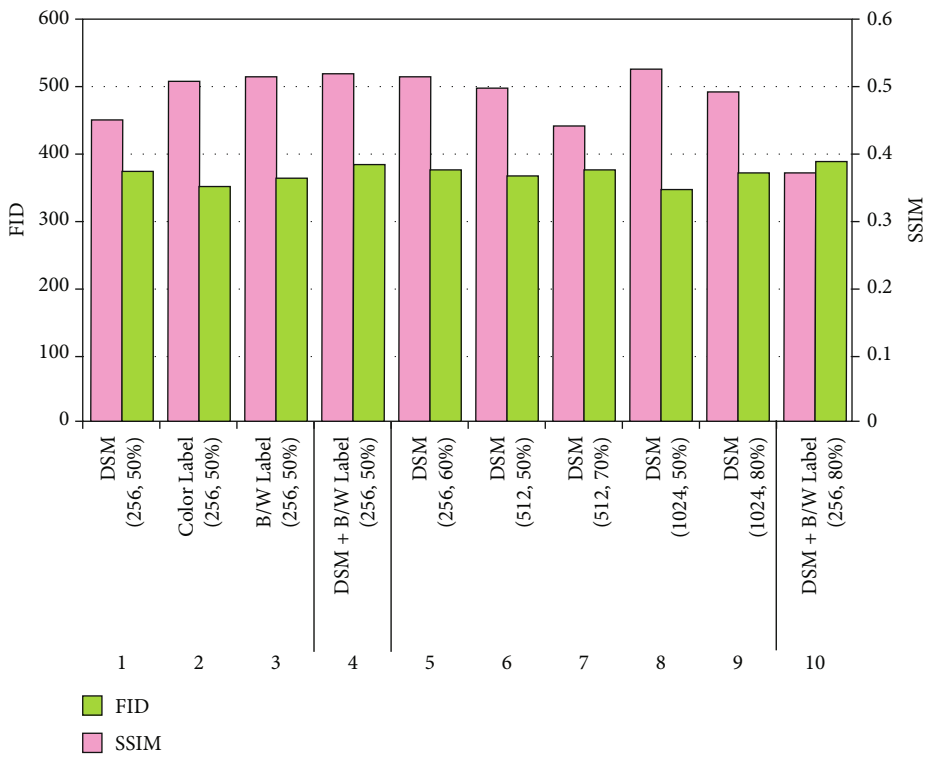
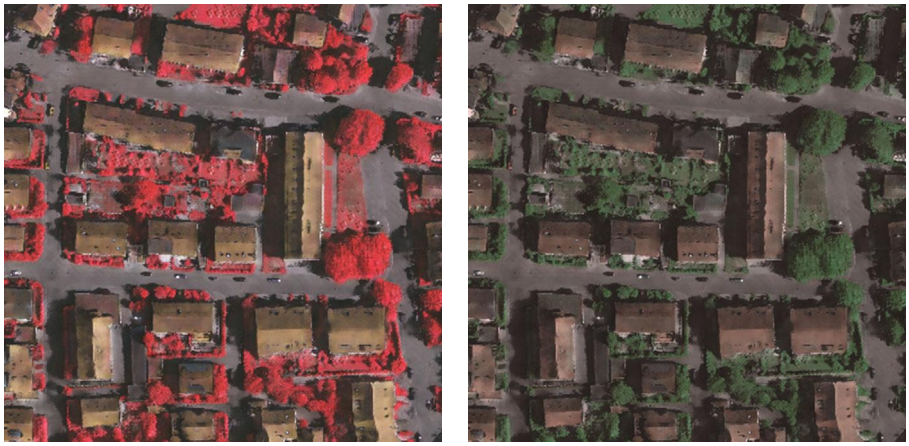


Figure 35: FID and SSIM of Potsdam unseen datasets.

could be applicable to DL. The major problem of the DL is data dependency. In other words, the results generated from the same DL model would be different even if the same training data is used. In this matter, various experiments were performed to provide trained models with different settings of the training datasets which could be provided including different tile sizes with overlap rates and different combinations of the datasets. It is obvious that the partitioning with overlap of the training datasets is beneficial because increasing the number of the datasets and diversity is desirable for training DL models. Also, training with multimodal datasets could provide improved results due to synergistic effect of each data. However, determining the optimal condition is not a trivial task. Based on the experiments with the specific dataset used in this paper, the case of using smaller tile size with higher overlap rate and combining all data results in the best training performance.

In most cases, labels are essential in DL. Sometimes, the labels are not available and intensive work such as manual screen digitizing of the orthoimages is required to obtain the label information. However, the labels could be created efficiently using the digital vector maps as shown in Figure 38. Since the digital maps are composed of classified layers, various labels could be created by selecting appropriate layers for specific purposes. Different colors of the features can be assigned to the label.



(a) (b)

Figure 36: IR to natural color conversion of generated true orthoimage of test data: (a) IR true orthoimage; (b) pseudonatural color.

(a) (b)

Figure 37: IR to natural color conversion of generated true orthoimage of unseen data: (a) IR true orthoimage; (b) pseudonatural color.

Table 4: Initial and final weights of spectral conversion matrix.

Spectral weights	Initial weights	Final weights after training	Learning rate	Epoch
$\begin{bmatrix} w_{11} & w_{21} & w_{31} \\ w_{12} & w_{22} & w_{32} \\ w_{13} & w_{23} & w_{33} \end{bmatrix}$	$\begin{bmatrix} 0.1 & 0.9 & 0.0 \\ 0.3 & 0.0 & 0.7 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}$	$\begin{bmatrix} 0.078 & 0.887 & -0.013 \\ 0.239 & -0.041 & 0.659 \\ 0.040 & 0.059 & 0.759 \end{bmatrix}$	0.1	100

(a) (b)

Figure 38: Example of label created from digital vector map: (a) building, road, and vegetation layer of digital map; (b) label.

The ultimate goal of DL is that the trained models are to be applicable to other datasets (i.e., unseen or new datasets) that are not involved in the model training. Unfortunately, use of the trained or pretrained models might be limited to the specific datasets that have similar properties to the training datasets. Therefore, various trained models to generate true orthoimages with multimodal datasets that take into account the regional and topographic properties could be more useful. In this matter, this paper might contribute to some strategies to improve the performance of the DL by providing an appropriate form of the training datasets.

Data Availability

The datasets generated during the current study are available from the corresponding author on reasonable request. The Vaihingen and Potsdam datasets are provided by the ISPRS Test Project on Urban Classification, 3D Building Reconstruction, and Semantic Labeling, https://www2.isprs.org/commissions/comm2/wg4/benchmark.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2018R1D1A1B07048732). The Vaihingen dataset was provided by the German Society for Photogrammetry, Remote Sensing and Geoinformation (DGPF) [23]: http://www.ifp.uni-stuttgart.de/dgpf/DKEPAllg.html.

References

[1] E. Mikhail, J. Bethel, and J. McGlone, *Introduction to Modern Photogrammetry*, John Wiley & Sons, 2001.

[2] K. Novak, "Rectification of digital imagery," *Photogrammetric Engineering and Remote Sensing*, vol. 58, no. 3, pp. 339–344, 1992.

[3] G. Zhou, W. Schickler, A. Thorpe, P. Song, W. Chen, and C. Song, "True orthoimage generation in urban areas with very tall buildings," *International Journal of Remote Sensing*, vol. 25, no. 22, pp. 5163–5180, 2004.

[4] Y. Liu, X. Zheng, G. Ai, Y. Zhang, and Y. Zue, "Generating a high-precision true digital orthophoto map based on UAV images," *ISPRS International Journal of Geo-Information*, vol. 7, no. 9, pp. 1–15, 2018.

[5] M. Haggag, M. Zahran, and M. Salah, "Towards automated generation of true orthoimages for urban areas," *American Journal of Geographic Information System*, vol. 7, no. 2, pp. 67–74, 2018.

[6] F. Amhar, J. Jansa, and C. Ries, "The generation of true orthophotos using a 3D building model in conjunction with a conventional DTM," *International Archives of Photogrammetry and Remote Sensing*, vol. 32, Part 4, pp. 16–22, 1988.

[7] J. Rau, N. Chen, and L. Chen, "True orthophoto generation of built-up areas using multi-view images," *Photogrammetric Engineering and Remote Sensing*, vol. 68, no. 6, pp. 581–588, 2002.

[8] K. Bang and C. Kim, "A new true ortho-photo generation algorithm for high resolution satellite imagery," *Korean Journal of Remote Sensing*, vol. 26, no. 3, pp. 347–359, 2010.

[9] A. Habib, E. Kim, and C. Kim, "New methodologies for true orthophoto generation," *Photogrammetric Engineering and Remote Sensing*, vol. 73, no. 1, pp. 25–36, 2007.

[10] P. Kuzmin, A. Korytnik, and O. Long, "Polygon-based true orthophoto generation," in *International Society for Photogrammetry and Remote Sensing (ISPRS) Congress*, pp. 529–531, Istanbul, Turkey, 2004.

[11] H. Gharibi and A. Habib, "True orthophoto generation from aerial frame images and LiDAR data: an update," *Remote Sensing*, vol. 10, no. 4, p. 581, 2018.

[12] E. Yoo and D. C. Lee, "True orthoimage generation by mutual recovery of occlusion areas," *GIScience & Remote Sensing*, vol. 53, no. 2, pp. 227–246, 2016.

[13] G. Zhou, W. Chen, J. Kelmelis, and D. Zhang, "A comprehensive study on urban true orthorectification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 9, pp. 2138–2147, 2005.

[14] A. Gunay, H. Arefi, and M. Hahn, "Semi-automatic true orthophoto production by using LIDAR data," in *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 2873–2876, Barcelona, Spain, 2007.

[15] Y. Sheng, "Minimising algorithm-induced artefacts in true ortho-image generation: a direct method implemented in the vector domain," *The Photogrammetric Record*, vol. 22, no. 118, pp. 151–163, 2007.

[16] L. Barazzetti, M. Brovelli, and M. Scaioni, "Generation of true-orthophotos with LiDAR high resolution digital surface models," *The Photogrammetric Journal of Finland*, vol. 21, no. 1, pp. 26–36, 2008.

[17] H. Oliveira, M. Galo, and A. Dal Poz, "Height-gradient-based method for occlusion detection in true orthophoto generation," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 11, pp. 2222–2226, 2015.

[18] C. Zhong, H. Li, and X. Huang, "A fast and effective approach to generate true orthophoto in built-up area," *Sensor Review*, vol. 31, no. 4, pp. 341–348, 2011.

[19] H. Oliveira, A. Poz, M. Galo, and A. F. Habib, "Surface gradient approach for occlusion detection based on triangulated irregular network for true orthophoto generation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 2, pp. 443–457, 2018.

[20] H. Bagheri and S. Sadeghian, "Ortho image and DTM generation with intelligent methods," in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XL-1/W3, pp. 475–480, Tehran, Iran, 2013.

[21] Y. Shin, S. Hyung, and D. C. Lee, "True orthoimage generation from LiDAR intensity using deep learning," *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, vol. 38, no. 4, pp. 363–373, 2020.

[22] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza et al., "Generative adversarial nets," in *Proceedings of the International Conference on Advances in Neural Information Processing Systems (NIPS)*, pp. 2672–2680, Montreal, Canada, 2014.

[23] M. Cramer, "The DGPF-test on digital airborne camera evaluation overview and test design," *Photogrammetrie - Fernerkundung - Geoinformation*, vol. 2010, no. 2, pp. 73–82, 2010.

[24] F. Rottensteiner, G. Sohn, J. Jung et al., "The ISPRS benchmark on urban object classification and 3D building reconstruction," *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. I-3, pp. 293–298, 2012.

[25] D. Lee, Y. Shin, and D. C. Lee, "Land cover classification using SegNet with slope, aspect, and multidirectional shaded relief images derived from digital surface model," *Journal of Sensors*, vol. 2020, Article ID 8825509, 21 pages, 2020.

[26] P. Isola, J. Zhu, T. Zhou, and A. Efros, "Image-to-image translation with conditional adversarial networks," in *2017 IEEE conference on computer vision and pattern recognition (CVPR)*, pp. 5967–5976, Honolulu, USA, 2017.

[27] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative adversarial networks: an overview," *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 53–65, 2018.

[28] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015), Lecture Notes in Computer Science*, vol. 9351, Springer, 2015.

[29] P. Nagorny, T. Lacombe, H. Favrelière et al., "Generative adversarial networks for geometric surfaces prediction in injection molding: performance analysis with discrete modal decomposition," in *2018 IEEE International Conference on Industrial Technology (ICIT)*, pp. 1514–1519, Lyon, France, 2018.

[30] V. Nair and G. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proceedings of the 27th International Conference on International Conference on Machine Learning (ICML)*, pp. 807–814, Haifa, Israel, 2010.

[31] A. Maas, A. Hannun, and A. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proceedings of the 30th International Conference on Machine Learning (ICML)*, Atlanta, USA, 2013.

[32] C. Aguilera, X. Soria, A. D. Sappa, and R. Toledo, "RGBN multispectral images: a novel color restoration approach," in *International Conference on Practical Applications of Agents and Multi-Agent Systems*, pp. 1–9, Porto, Portugal, 2017.

[33] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[34] D. Brunet, E. Vrscay, and Z. Wang, "On the mathematical properties of the structural similarity index," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1488–1499, 2012.

[35] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs trained by a two time-scale update rule converge to a local nash equilibrium," in *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS)*, pp. 6629–6640, Long Beach, CA, USA, 2017.

[36] A. Borji, "Pros and cons of GAN evaluation measures," *Computer Vision and Image Understanding*, vol. 179, pp. 41–65, 2019.

Hindawi
Journal of Sensors
Volume 2021, Article ID 5524057, 14 pages
https://doi.org/10.1155/2021/5524057

Research Article

Analysis of Surface Deformation in East Dongting Lake Based on 2016-2019 Sentinel-1A Dataset

Heng Zhang,[1,2] **Qian Sun**,[1,2] **and Jun Hu**[3]

[1]*College of Geographic Science, Hunan Normal University, Changsha 410081, China*
[2]*Key Laboratory of Geospatial Big Data Mining and Application, Hunan Normal University, Changsha 410081, China*
[3]*School of Geosciences and Info-Physics, Central South University, Changsha 410083, China*

Correspondence should be addressed to Qian Sun; sandra@hunnu.edu.cn

Received 8 February 2021; Revised 29 March 2021; Accepted 19 April 2021; Published 3 May 2021

Academic Editor: Hyung-Sup Jung

In this study, 84 scenes Sentinel-1A satellite datasets from October 2016 to September 2019 were used to analyze surface deformation in East Dongting Lake, China, by employing the small baseline subset interferometric synthetic aperture radar (SBAS-InSAR) method. The data are divided into two seasons, i.e., the flood and dry seasons. It was suggested that the surface deformation is related to the distribution of the river network and water flow activities. During the dry season, the water flow is active along the internal river, scouring the surrounding soil. During the flood season, the water flow basically occurs around the external drainage network. By qualitatively comparing surface deformation and precipitation changes as well as changes in soil erosion, it is found that the deformation was highly related to soil erosion and seasonal precipitation. The precipitation in the flood period is heavy than that in the dry season. Therefore, the runoff with amount silt will scour the soil in the passing area, resulting obvious surface deformation. During the dry period, surface deformation is smaller due to the less precipitation.

1. Introduction

Surface deformation in lake region associated with mud and sand is a common problem in the world nowadays, seriously threatening agriculture irrigation, natural resources protection, and ecosystem environments regulation [1, 2]. Lake region's surface deformation mainly includes two processes. One is surface settlement, which is caused by score and erosion of flowing water to soft soil in the uppermost layer of the lake region. The surface soil flows to other areas with flowing water, thus, causing surface settlement in the passing area. Another deformation is surface sedimentation, which is due to the increase of sediment content brought by the flowing water. Due to gravity and sedimentation, silt deposits on the surface and raises the lake bed. Lake area will decrease due to sedimentation. Continuous silt uplift of the river bed also brings huge challenges to flood control. Over the past years, numerous mud blockage river events had been reported in many lakes around the world [3]. In China, more than 5 billion tons of mud and sand are lost to drainage network each year, which causes the silting of rivers, raising of river beds, silting of reservoirs, and reduction of the reservoir capacity [4]. Therefore, surface deformation in lake region seriously endangers water conservancy construction and the flood control capacity [5, 6].

Traditional ground leveling [7] and global positioning system (GPS) measurements cannot provide sufficient samples required by surface deformation monitoring. Interferometric synthetic aperture radar (InSAR) technology is a modern space geodetic survey technology that has developed rapidly in recent decades [8]. InSAR can not only obtain a wide range of high-precision digital elevation models (DEM) [9] but also monitor small topographic changes that range from a millimeter to meter through Differential InSAR (D-InSAR) [10]. InSAR has greatly improved its ability to monitor tiny ground deformations, especially with the development of the multitemporal (MT-InSAR) algorithm [11, 12], which overcomes temporal decorrelation and atmospheric disturbance with D-InSAR [13]. The MT-InSAR algorithm analyzes a time series of targets with high coherence points in the interferograms and then obtains the deformation velocity and cumulative deformation of the earth

surface [14]. InSAR has been successfully applied in earthquakes, landslides, and other fields [15–17].

Dongting Lake, which is the second largest freshwater lake in China, is an important regulatory and storage lake in the Yangtze River basin. Dongting Lake is a lake group composed of multiple lakes, including East Dongting Lake, South Dongting Lake, West Dongting Lake, and DaTong Lake. Dongting Lake is a wetland with a dense river network and complex biodiversity, known as "the kidney of the Yangtze River" [18]. Recently, due to the drainage silt into the East Dongting Lake area in a large amount of sediment, the river channel is continuously silted up, and the riverbed is raised [19]. The lake area continues to shrink, and the flood control efficiency is gradually weakened, which has become a threat to the ecological, economic, and social benefits of Hunan province and even the whole country [20]. Therefore, it is necessary to continuously monitor the surface deformation of the lake area to understand the process of erosion and sedimentation changes in Dongting Lake, which can qualitatively and quantitatively reflect the area, range, distribution, and dynamic changes of sediment erosion and sedimentation. This can effectively improve the scientific and effectiveness of disaster prevention flooding of Dongting Lake. However, continuous monitoring of surface deformation covering the areas of East Dongting Lake is lacking.

In this study, 84 Sentinel-1A satellite images were exploited to monitor the surface deformation of East Dongting Lake with a classical MT-InSAR algorithm, i.e., small baseline subset InSAR [12] (SBAS-InSAR). The use of SBAS-InSAR can reduce the phenomenon of temporal and spatial decoherence. However, for a study area with large surface changes such as the Dongting Lake, if the long-term deformation is directly monitored by using SBAS-InSAR, the lake areas would be basically incoherent due to the alternating process of dry and flood periods. If it is not separated for data processing, lake areas will have no results. Therefore, in this article, considering that the lake water storages in East Dongting Lake are quite different during the seasons of high and low water lines, the deformation with respect to dry and flood seasons during 2016-2019 were estimated, respectively. In addition, by processing three flood periods and three dry periods, respectively, we can compare the changes of the flood or dry seasons among different years, which were used to reflect on the intensity of surface deformation and reveal the spatial-temporal variations in the surface deformation of East Dongting Lake. Erosion and sedimentation changes in East Dongting Lake, precipitation, and human activities were considered as influencing factors, and their effects on surface deformation were studied.

2. Study Area and Datasets

2.1. Site Description. East Dongting Lake, which is located in the northeast of Hunan province, south of the Jingjiang River section of the middle reaches of the Yangtze River. The area of East Dongting Lake is 1328 km^2. The East Dongting Lake nature reserve is one of the 21 wetlands nature reserves of international importance designated by the Chinese government [21]. The specific location of East Dongting Lake is shown in Figure 1. The climate of this area is a continental subtropical monsoon humid climate. The average annual temperature is 17°C, the annual precipitation is about 1200-1300 mm, and the annual frost-free period is 285 days [22]. Dongting Lake, as one of only two river lakes in the middle and lower reaches of the Yangtze River, plays an important role in regulating the flood runoff of the Yangtze River and protecting the biodiversity of species. Previous research has indicated that the annual mean sediment is 1.7×10^9 t, of which 80% comes from the channels, 18% from the tributaries, and 2% from the local area, whereas 26% of the total sediments are transported into the Yangtze River and deposited into the lake [23]. However, due to the soil erosion in the Dongting Lake area, a large amount of sediment is discharged, the river channel is continuously silted up, and the riverbed is raised [20], which will cause surface deformation of East Dongting Lake. At the same time, there are many polders in the East Dongting Lake Plain, which is one of the state commercial grain bases. Human activities will accelerate surface fragmentation, increasing river sediment [24].

2.2. Data Used. In this study, 84 Sentinel-1A images acquired from 07 October 2016 to 10 September 2019 were collected for the study area. Sentinel-1A is a medium-high resolution C-band satellite launched by the European Space Agency (ESA) in 2014. Table 1 shows the specific parameters of Sentinel-1A used in this study. Except for resolution, another advantage is that the Sentinel-1A satellite's return cycle is only 12 days. With the launch of Sentinel-1B in 2016, a temporal resolution of 6 days could be approached. However, the Sentinel-1B data are unavailable for this study, so this study only employed Sentinel-1A data. The Sentinel-1 data are available free of charge to users around the world. In this study, considering the variation of lake water storage in East Dongting Lake during the different seasons, the time series images were divided into two groups, i.e., dry and flood seasons. It was assumed that from March to September of each year is the flood season, and that from October to February of the next year is the dry season. 12 images were selected for each one of the three dry seasons during 2017 and 2019, and 16, 18, and 14 images were selected for years 2017, 2018, and 2019, respectively. Figure 1 shows the location and geomorphology of the study area and the Sentinel-1A data coverage of part of the study area.

3. Methodology

3.1. SBAS-InSAR Processing. SBAS-InSAR inherited the advantages of the conventional D-InSAR, which generates an appropriate combination of differential interferograms produced by SAR data pairs based on the spatiotemporal baseline threshold. The interferogram satisfying the threshold condition is taken as the initial quantity, and the least square (LS) method or singular value decomposition (SVD) method is then used to obtain the average deformation rates and accumulative variables of the observation periods [24, 25].

Although SBAS-InSAR can suppress the spatiotemporal decoherence by setting the spatiotemporal baseline

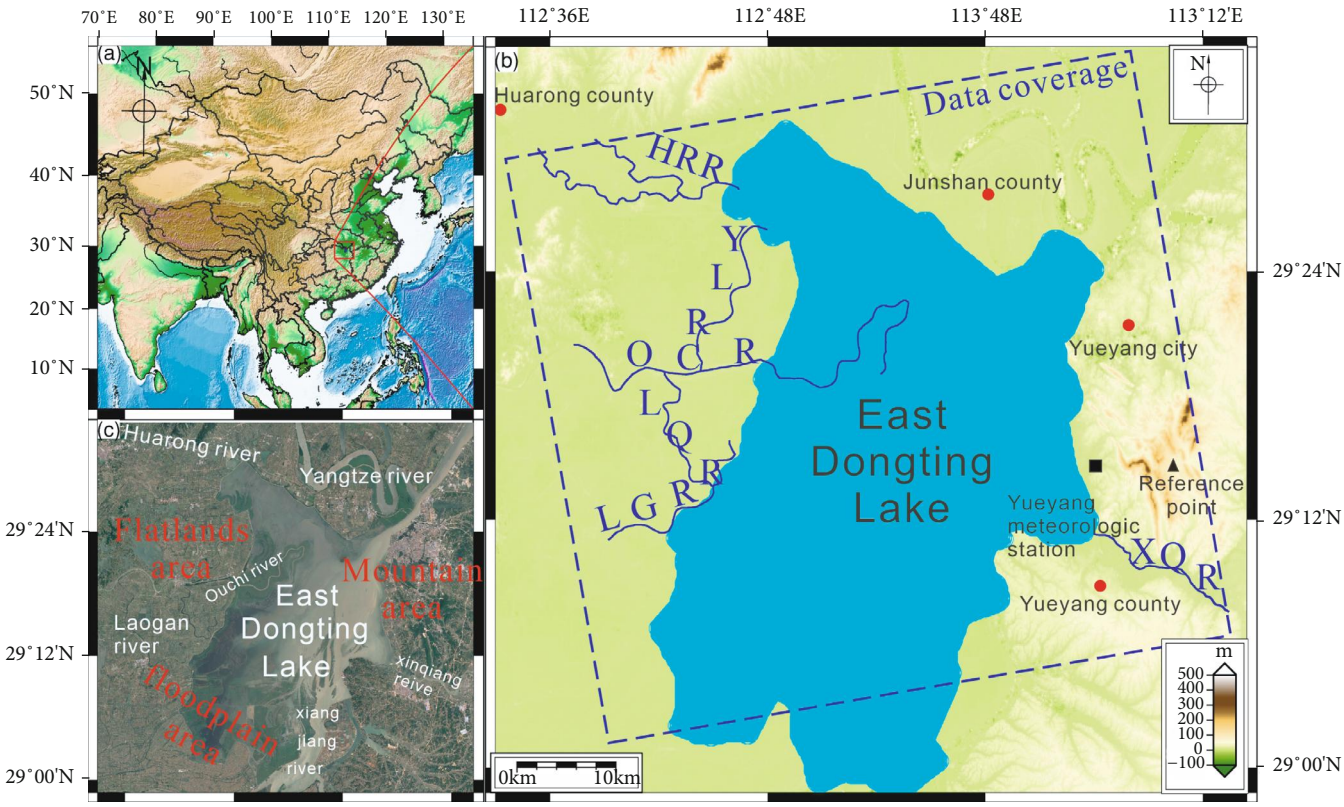


Figure 1: (a) Location of East Dongting Lake in China. (b) Terrain of the study area. Due to the large changes in the area of East Dongting Lake during the flood and dry periods, in this study, the embankment of the East Dongting Lake is considered as the lake boundaries, shown in light blue. Blue lines are the rivers around East Dongting Lakes. HRR: Huarong River; LQR: Longqing River; LGR: Laogan River; XQR: Xinqiang River; YLR: Yuelai River. Dark blue dashed line is the research area after subsetting of the satellite images in this paper, and red dots show the main towns. Black rectangle represents the location of the meteorological station. Black triangle represents the location of unwrapping reference point. (c) Google Earth image of East Dongting Lake acquired in flood erosion. The east is the mountainous area, the west is the flatlands, the middle south is the lake floodplain terrace, and the north is the Yangtze River floodplain.

Table 1: Specific parameters of Sentinel-1A data.

Parameters	Description
Product mode	Sentinel-1A IW mode
Track number	11
Azimuth angle (degree)	90
Incidence angle (degree)	43.94
Orbit direction	Ascending
Polarization	VV

thresholds, seasonal coherence decorrelation is quite serious in some specific research areas (e.g., seasonal lake areas and reservoir), where it is impossible to obtain effective interferometric pairs in the study area. Some scholars had proposed that the Intermittent Small Baseline Subset (ISBAS) method can be used to select appropriate interferometric pairs [26]. However, in this study, there are obvious dry and flood periods in East Dongting Lake. During the alternating process of dry and flood periods, the lake areas are severely incoherent. While in the flood period, this situation is even more serious. Therefore, time sequence analysis cannot be performed by combining all data.

Therefore, in this article, the SBAS-InSAR method is applied to the dataset in single season rather than the whole investigated period. At first, the data were divided into two seasons (i.e., example and flood dry periods), and then we obtained the surface deformation of each flood or dry period. There are two main points to be noted in this paper. On the one hand, flood and dry periods were classified based on the obtained images. On the other hand, the interferometric pair generated according to the spatiotemporal threshold also needs to be further selected according to the quality of the interferogram, especially in the flood season, because the temporal decorrelation is obvious.

3.2. Data Processing. All Sentinel-1A data were registered to the same common master image of each period. The acquisition data of the super master image for the dry and flood seasons selected in this study were December 30, 2016; June 4, 2017; December 13, 2017; May 6, 2018; December 20, 2018; and June 18, 2019, respectively. In order to avoid the phase jump phenomenon caused by different bursts with Terrain Observation by Progressive scans SAR (TOPSAR) modes of Sentinel-1A data, Shuttle Radar Topography Mission (SRTM) DEM data provided by National Aeronautics and Space Administration (NASA) were used for assistant

registration in this study. As mentioned above, the ground coverage corresponding to the Sentinel-1A image is 250 km × 250 km. In order to reduce the unnecessary data and calculation, the data located outside the research area were removed in this experiment.

In order to reduce the phase noise, 10 × 2 (range × azimuth) multilook rate was adopted in data processing. In the differential process, 30 m SRTM data were used as external DEM data to remove the topographic phase. In order to avoid the generation of small isolated baselines and maintain the requirement of coherence, the threshold of the perpendicular baseline was 80 m. In addition, decorrelation will also occur when the time interval between two images is too long, which will affect the final accuracy. Therefore, the threshold of the temporal baseline was set as 60 days for both the dry and flood period. The temporal and perpendicular baselines of the used interferograms are shown in Figure 2.

The interferograms were generated according to the small baseline set image pairs selected [26]. In order to improve the quality of the interferogram, the spatial adaptive filtering method was used to eliminate random noise in the interferogram [27]. Then, phase unwrapping of the interferogram was carried out with the minimum cost flow method. In this method, a stable point in the research period is required as the start point of unwrapping. Phase unwrapping was conducted by choosing a pixel located in Yueyang city as a reference point. The location of the reference point is shown as a black triangle in Figure 1(b).

According to the coherence value, the interferograms with a relatively high coherence were selected as the unwrapping mask of the flood and dry season data, and the average coherence threshold in dry and flood seasons was 0.5 and 0.3, respectively. According to the characteristics of the lake water line in the flood and dry periods, different unwrapping masks should be determined in advance, according to the coherence of differential interferograms before phase unwrapping. Meanwhile, interferometric pairs with low coherence and poor unwrapping were removed. Finally, we obtained 41, 32, and 32 differential interferograms in the dry season of 2017, 2018, and 2019, respectively, and 47, 64, and 40 differential interferograms in the flood season of 2017, 2018, and 2019, respectively. Figure 3 shows the representative differential interferograms for each season. Moreover, the differential phase includes the elevation residual phase, atmospheric delay phase, and noise phase. Therefore, it is necessary to remove other phases to obtain an accurate deformation [12]. In order to reduce the error caused by low coherence in the differential interferograms, the SBAS-InSAR algorithm only calculates the high coherence points. In this study, the coherence coefficient average threshold method was used to select high coherence points [28, 29].

4. Results

4.1. Mean Rate in the Dry Season. Figure 4 shows the monthly average surface deformation rate the of the study area, during the dry season. The map is superimposed on the Google Earth image of East Dongting Lake acquired in each dry season, and the white line is the embankment of East Dongting Lake. Positive values represent that the surface is getting closer to the radar along the line of sight (LOS) direction, and negative value surface is further away from the radar in the LOS direction. The water line of East Dongting Lake started to rise at the beginning of the annual flood season and then reached the peak at the end of July and early August. After that time, Chenglingji (one of the East Dongting Lake outlets) began to open floodgates to reduce the volume of East Dongting Lake water. Subsequently, in the dry season, the lake area began to decrease, and the running water carried away the lake silt in the area, causing surface settlement in the lake floodplain.

Region A is in the floodplain area at the south-central side of East Dongting Lake, and the main settlement is distributed along the inner channel with an average deformation velocity of about -3 mm/month in 2017. In the 2018 dry season, the intensity of deformation was even greater around the region, and the average rate of the settlement reached -6 mm/month. The settlement in this area mainly occurs on the southern floodplain terraces and on both banks of the rivers inside the lake. The main cause is the river water entering the lake during the dry season and the internal water flow of the lake scouring the surrounding soil. In the 2019 dry season, the average rate decreased to about -2 mm/month. There is another channel along the embankment in the southwest. During the dry season, there is also flowing water to flow through the surrounding silt, which could cause the settlement center moved southwest in 2019.

Region B is located at the entrance of Xinqiang River in the east of East Dongting Lake, and the river twists and turns in the region. The average rates of settlement in 2017 and 2018 were about -5 and -6 mm/month, respectively. There are two reasons for this settlement. First, the river has twists and turns, making it easier for sediment to accumulate in the water during the flood period. When the sediment is not completely integrated into the lake surface during the dry period, it is more likely to be eroded. Second, the flowing water outside Xinqiang River will directly scour this area. However, surface deformation was hardly observed in this region in the 2019 dry season. This can be ascribed to the gentle water flow of this tributary and Xinqiang River caused by the relatively less precipitation in the previous year (i.e., 2018). In addition to the deformation at the entrance of the Xinqiang River, the deformation is found in the entrance of the Ouchi River.

Region C (see Figure 4(b)) is a small area of surface settlement in the northwest corner of East Dongting Lake, where there is a small reservoir. If the water line of East Dongting Lake is too high to threaten other areas, this reservoir can regulate and store East Dongting Lake for water storage and begin to drain outwards in the dry period. In the reservoir, an obvious settlement velocity of about -5 mm/month was only detected in 2018. It is reported that there was a serious flooding disaster in East Dongting Lake in the summer of 2017, which caused the increase of silt in the entire lake area [30]. This small reservoir also contained a large amount of mud and soil. Therefore, when the water level drops during the dry season of the next year (i.e., 2018), the drainage

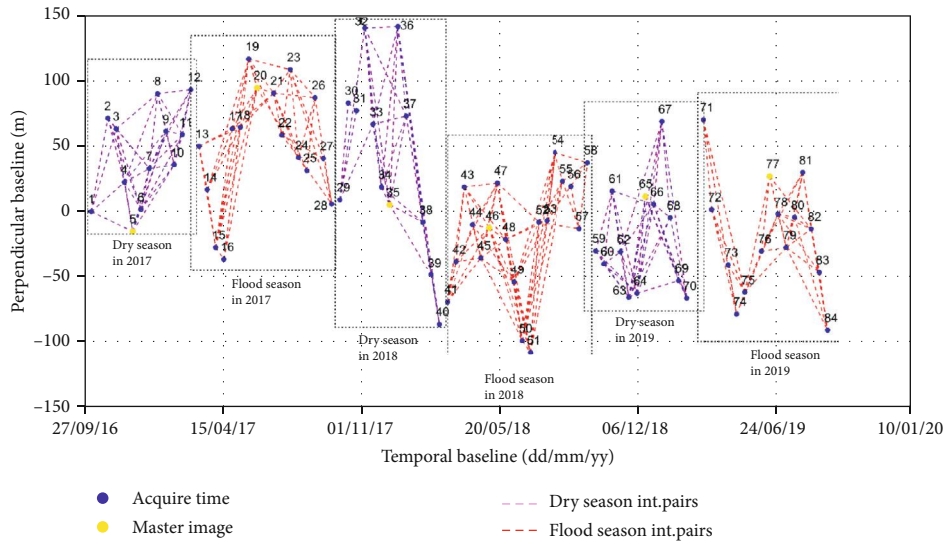


Figure 2: Temporal and perpendicular baselines of study periods. Blue dots represent the image acquisitions, and yellow dots represent the super master images. Purple and red dashed lines represent interferometric pairs with the dry and flood seasons, respectively.

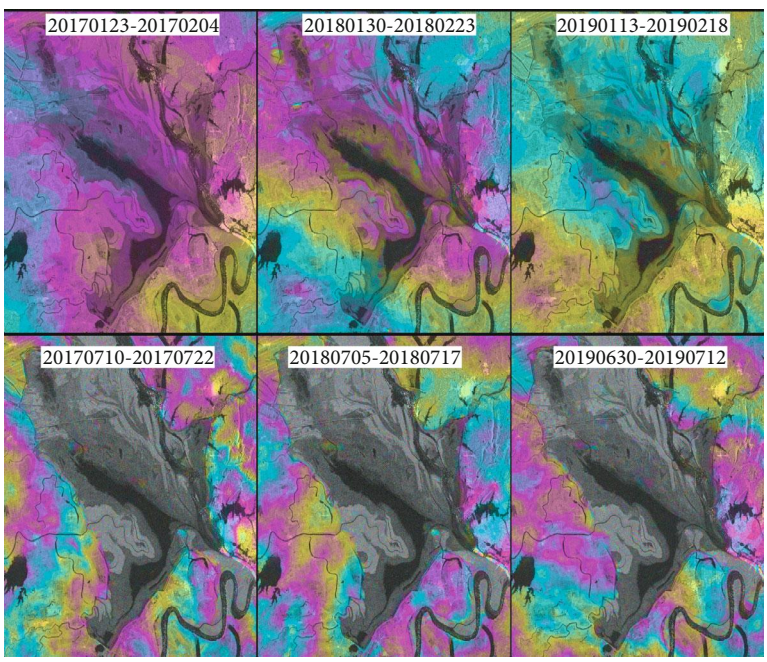
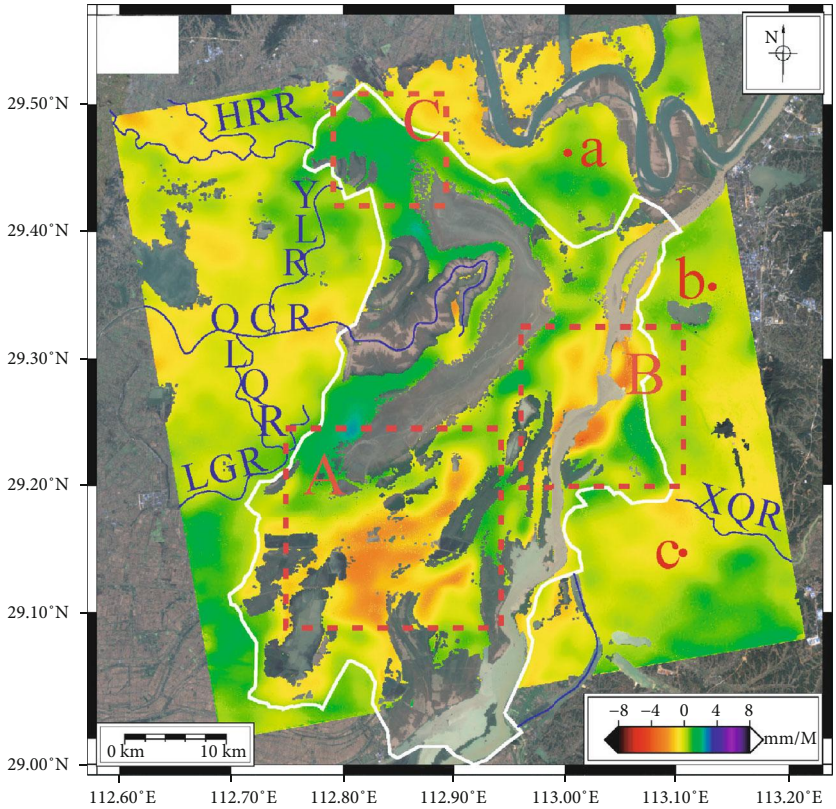


Figure 3: Representative differential interferograms for each season, the first row is the differential interferogram in the dry period, and the second row is the differential interferogram in the flood period. Acquired time of differential interferogram is presented at the top of each subplot.

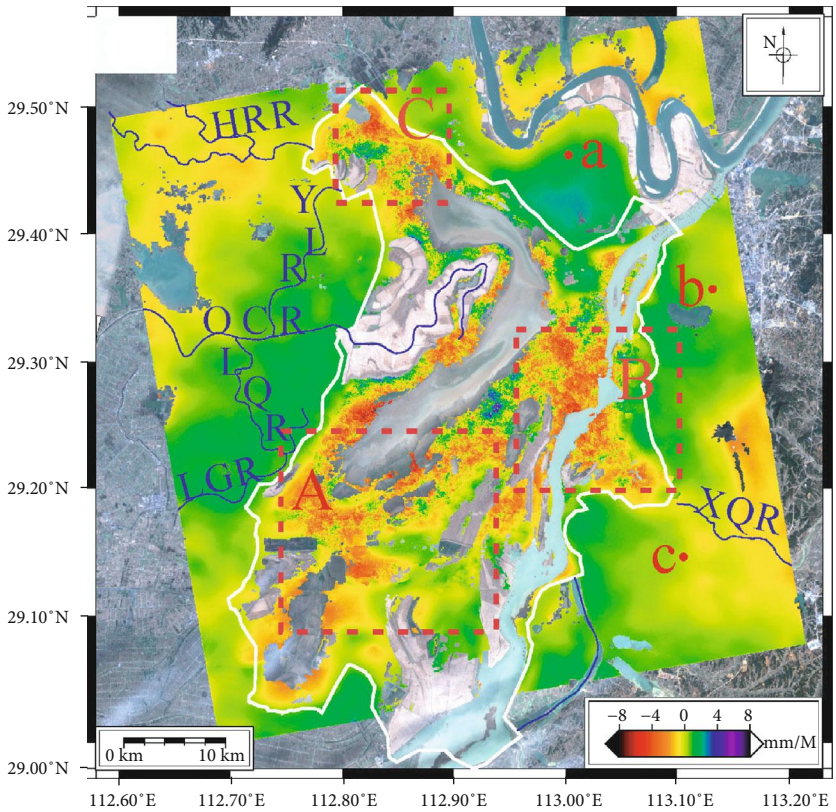
process will wash out the coastal soft soil, leading to coastal erosion and collapse.

Trends in the East Dongting Lake surface deformation intensity grade dynamic changes in time and space could be indicated by analyzing the settlement and sedimentation grade in East Dongting Lake with dry season. We use a method similar to the equidistant division to divide the deformation rate. The purpose of this division is to assess the degree of surface deformation of the East Dongting Lake in comparison with the water regime. The deformation intensity are divided into six classes, including a severe settlement area (<-4 mm/month), moderate settlement area (-4 to -2 mm/month), slight settlement area (-2 to -0.5 mm/month), stable area (-0.5 to 0.5 mm/month), slight sedimentation area (0.5 to 2 mm/month), and severe sedimentation area (>2 mm/month). The areas of each surface deformation intensity grade of dry seasons are compared in Table 2.

In the dry season, the severe settlement area, moderate settlement area, slight sedimentation area, and severe sedimentation area all experienced the change of increasing from 2017 to 2018 and then decreasing from 2018 to 2019. On the

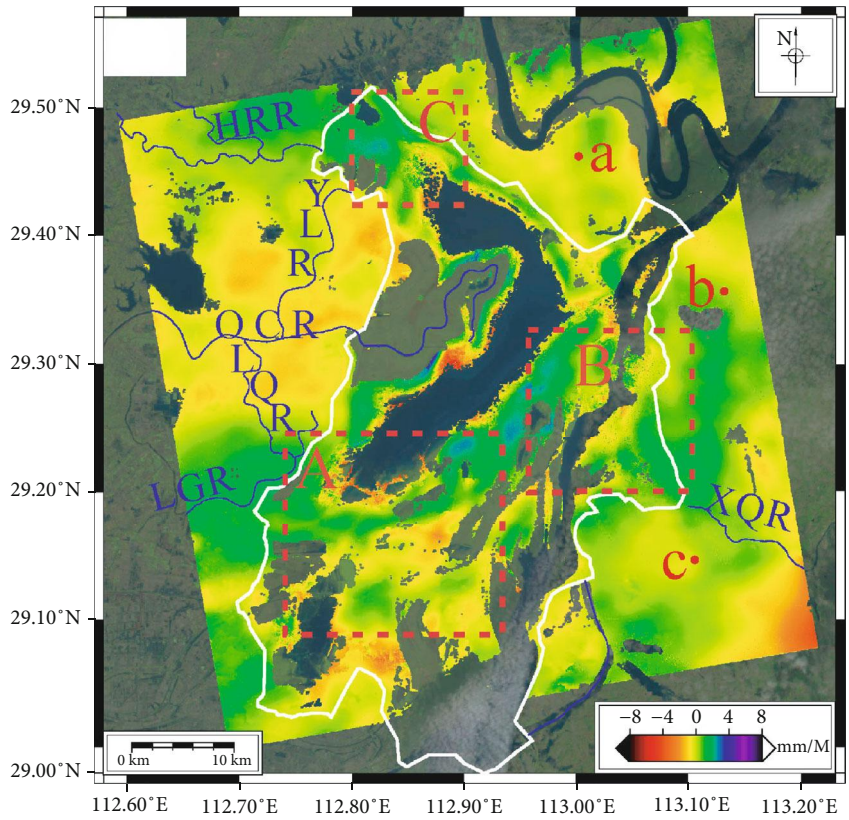


(a)



(b)

Figure 4: Continued.



(c)

Figure 4: The monthly average deformation rate along LOS direction across East Dongting Lake during the dry season of 2017 (a), 2018 (b), and 2019 (c). The base image is an optical image of each dry season. White line represents the embankment of East Dongting Lake, and the blue lines are the rivers around East Dongting Lakes. HRR: Huarong River; LQR: Longqing River; LGR: Laogan River; XQR: Xinqiang River; YLR: Yuelai River. Red dots represent the main towns. a: Junshan district, b: Yueyang city, and c: Yueyang county. Region A: southern floodplain area, Region B: entrance of the Xinqiang River, and Region C: northwest catchment area.

Table 2: Comparison of the intensity grades of surface deformation in the dry seasons.

Intensity grade of surface deformation	Surface deformation area in the dry season					
	2017		2018		2019	
	Area (km^2)	%	Area (km^2)	%	Area (km^2)	%
Severe settlement	0.04	0.002	6.62	0.370	0.88	0.050
Moderate settlement	19.99	1.121	67.68	3.781	15.82	0.895
Slight settlement	356.10	19.968	349.78	19.542	335.50	18.985
Stable area	959.70	53.814	717.26	40.072	843.28	47.720
Slight sedimentation	440.93	24.725	619.87	34.631	549.04	31.069
Severe sedimentation	6.59	0.370	28.70	1.604	22.64	1.281

contrary, the areas of slight settlement and stable area experienced a decrease followed by an increase during the three years' dry seasons. According to the deformation areas presented in Table 2, it can be found that the deformation intensity was 2018 > 2019 > 2017 for the dry season. Comparing the changes in the area of sedimentation and settlement, we can find that the sedimentation and settlement in East Dongting Lake are synergistic and simultaneous. Meanwhile, the stable areas still account for at least 40% of areas for the dry season. These stable areas are mainly located in the outer areas of East Dongting Lake.

4.2. Mean Rate in the Flood Season. The water level of East Dongting Lake rises during the flood season. Due to the specular reflection on the water surface, the echo signal is weak, and even local areas have no coherence. Therefore, there are no coherence points inside East Dongting Lake. As a result, surface deformation in the flood season can only be obtained by analyzing the deformation around East Dongting Lake basin. Flood runoff will scour away loose surface soil along slopes and river valleys and cause river silt. Moreover, in the flood season, the water flow is large and the flow velocity is fast, exhibiting great transportation and scour ability.

This means that the flood not only carries the sediment carried by the storm runoff but also scours the riverbed and riverbank, bringing the original sediment from the river valley into the turbulent flow and East Dongting Lake.

During the flood season, the tributary water system of East Dongting Lake plays an important role in the process of surface deformation. The first is Huarong River basin, which is on the northwest of East Dongting Lake; the second is the Ouchi River basin, which include three tributaries, i.e., Laogan River, Longqing River, and Yuelai River; the third is the basin of Xinqiang River on the eastern of East Dongting Lake; and the last is the sandbank in the southeast of East Dongting Lake. Figure 4 shows the surface deformation of East Dongting Lake in each flood season.

Region D shows the surface deformation of the Huarong River Basin, which is located in the northwest of East Dongting Lake, and is one of the "four mouths" of the Yangtze River flowing into East Dongting Lake, which has played an important role in regulating the flood season of the Yangtze River. There was obvious settlement area mainly concentrated in the northern part of Huarong River during the flood season in 2017 and 2019. The settlement rate in 2017 was about -3 mm/month. In 2019, settling velocity phenomenon was more serious, with the value exceeding -4 mm/month. There was no obvious settlement around the Huarong River in 2018, which can be ascribed to a relatively smaller Yangtze River flood run-off in 2018, as indicated by the relatively less precipitation in 2018. Region E is the surface deformation area around Ouchi River and its three tributaries. Ouchi River is located in the west of East Dongting Lake. It is also one of the "four mouths" of the Yangtze River flowing to East Dongting Lake. Compared with the Huarong River, the Ouchi River is longer and has a wider basin and more tributaries. During the flood season in 2017, there was obvious surface settlement around the Laogan River and Longqing River, with the settlement rate being about -3 mm/month. In 2018, the areas with settlement center were mainly concentrated in the west of the Yuelai River, and the rate was also nearly -3 mm/month. In the 2019 flood season, there was a relatively serious settlement area in south of Laogan River and Longqing River, with a velocity of about -6 mm/month. This is mainly an agricultural planting area, where irrigation water will accelerate the flow of water and have a stronger scouring effect on the surrounding area. In addition, human activities in agricultural areas may also loosen the surface soil and be more easily washed by running water.

The Xinqiang River located in Region F is an important river for the East Dongting Lake Basin. During the flood season, floods in the eastern mountainous areas flow through the Xinqiang River into the East Dongting Lake. In Figure 5, we found that the settlement phenomenon became more and more serious in the Xinqiang River Basin from 2017 to 2019. In 2019, the rate of settlement reaches up to -4 mm/month. In 2017 and 2018, the maximum deformation settlement rates were close to -2 mm/month. However, the surface settlement area was larger in 2017 than that in 2018. We can find that there are more serious erosion gullies in the floodplain of Xinqiang River watershed.

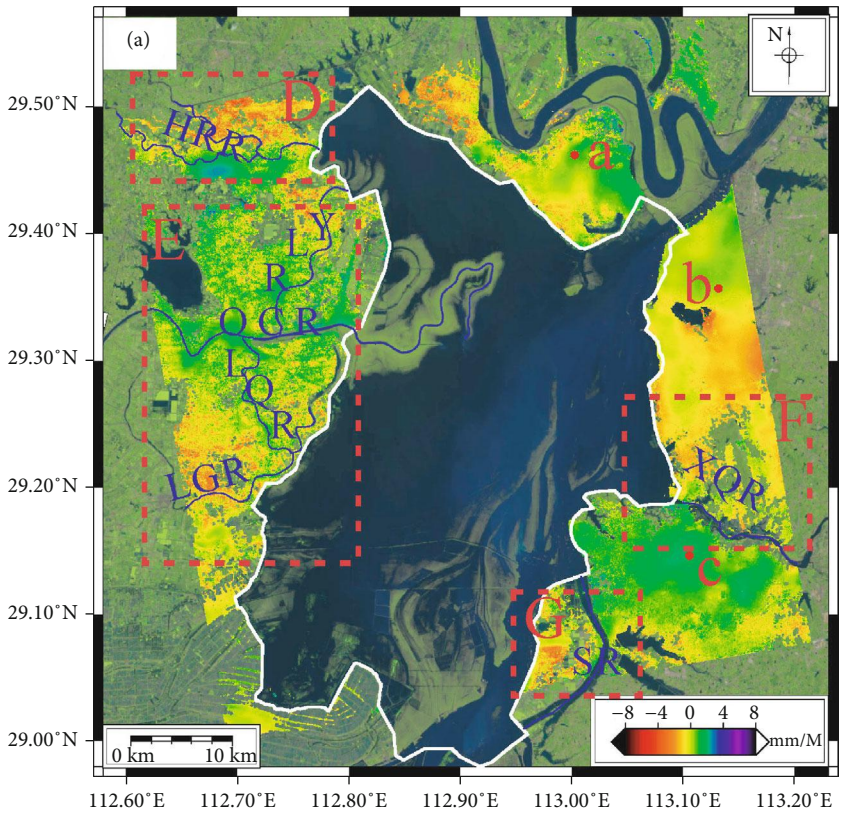
In addition to the three main drainage networks mentioned above, there was also a perennial settlement area in the sandbank at the southeast of East Dongting Lake (see Region G in Figure 5). The surface deformation rate of this area is up to about -4 mm/month. Besides, this region is basically a rice planting area with agricultural planting and human activity; therefore, the surface deformation can be related to the Laogan River area as well as irrigation and agricultural activities.

As shown in Table 3, the intensity grades of surface deformation are also classified during the flood season. The severe settlement area, moderate settlement area, slight settlement area, slight sedimentation area, and severe sedimentation area all experienced the change of first decreasing and then increasing from 2017 to 2019, and the stable areas experienced a decrease followed by an increase during flood seasons. By comparing the area of surface deformation, it could be found that the deformation intensity was 2017 > 2019 > 2018 for the flood season. We also noticed that, before the flood season of 2018, East Dongting Lake experienced the most serious subsidence phenomenon in the 2018 dry season. This indicates that after having experienced strong settlement and sedimentation, the earth surface in East Dongting Lake maintains a period of relative stability. Therefore, we found that, during the 2018 flood season, the stable area accounted for 56.994%, and slight settlement and slight sedimentation accounted for 21.282% and 21.061%, respectively. Those three deformation intensity grade areas dominated more than 99% of the observed area.

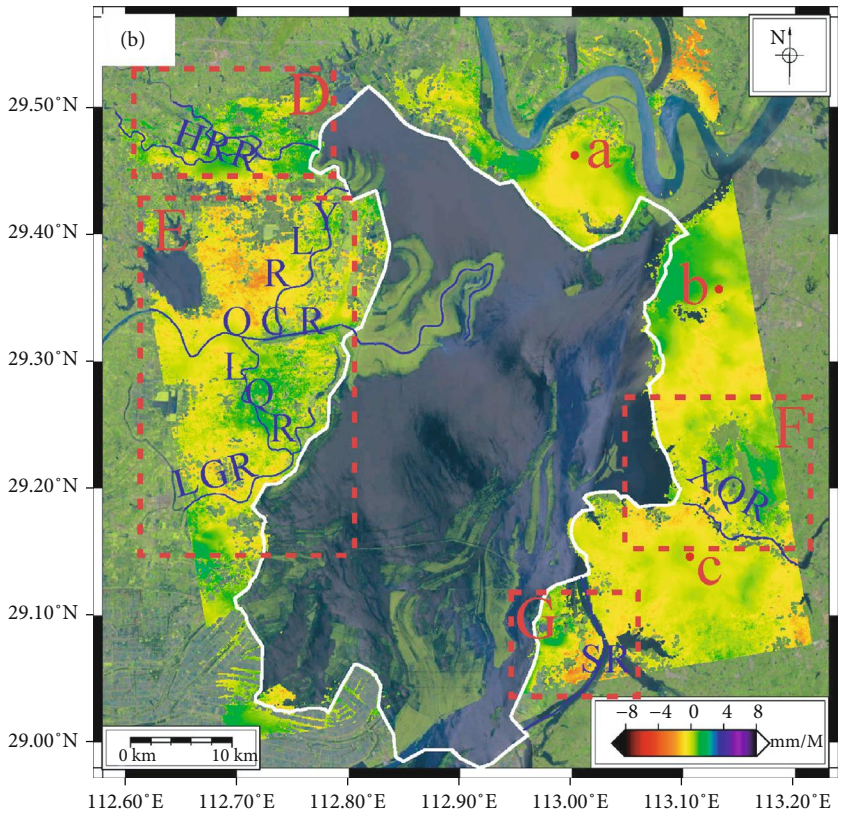
4.3. Time Series Deformation of East Dongting Lake. Although the method proposed in this study can obtain the results of surface deformation in different periods, the time series deformation of the study period cannot be proved. In this section, we use the conventional SBAS-InSAR method to combine all images to obtain the time series of the surface deformation.

Figure 6 shows the average deformation rate during the whole investigated period and the time series deformation of the feature points.

As shown in Figure 6(a), the deformation area obtained from the whole investigated period is similar to that obtained from the flood season. It can be found that the deformation area is located at the external river network of East Dongting Lake (e.g., Huarong River, Ouchi River, Xinqiang River, and Langan River). The average rate of deformation during the whole investigated period is smaller than those during the dry or flood seasons. This can be ascribed that the sedimentation and the settlement effects in the dry and flood seasons offset each other in the long-term deformation results. In order to show more details of the detected surface deformation, the time series deformations over the three typical feature points located at the external river network are shown in Figure 6(b). It was found that the selected three feature points basically showed a linear settlement during the dry period. However, the time series deformation shows obvious fluctuations during the flood period, which is highly correlated to the precipitation increase at the same time. This also demonstrates that the precipitation in the flood season

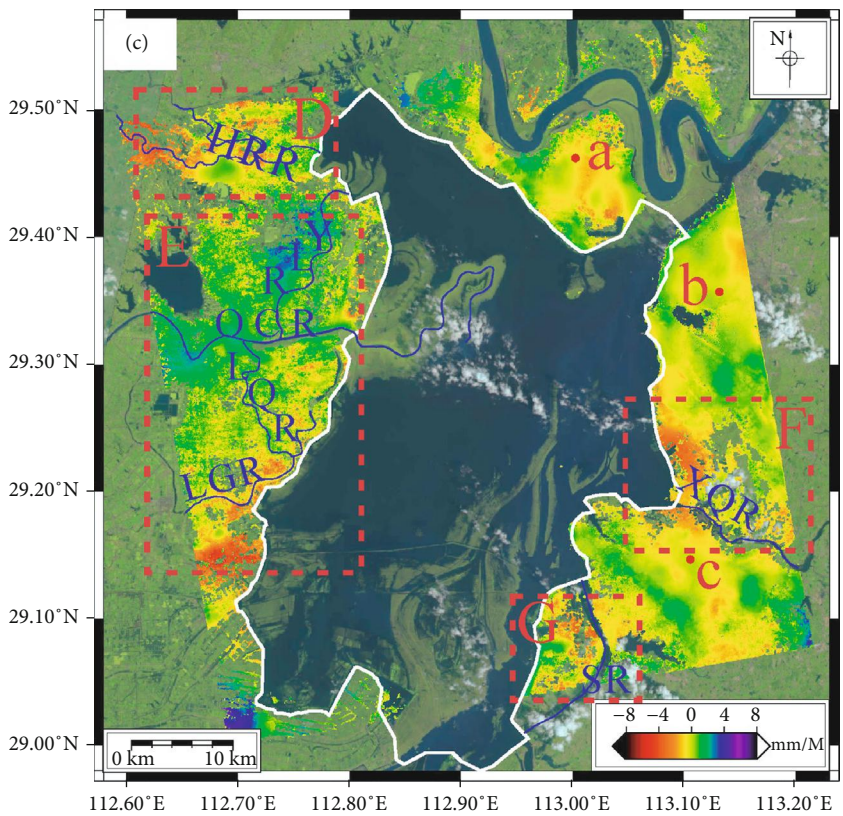


(a)



(b)

Figure 5: Continued.



(c)

Figure 5: Surface deformation rate caused by soil erosion in East Dongting Lake during the flood season of 2017 (a), 2018 (b), and 2019 (c). The base image is an optical image of each flood season. The white line represents the embankment of East Dongting Lake, and the blue lines are the rivers around East Dongting Lake. HRR: Huarong River; LQR: Longqing River; LGR: Laogan River; XQR: Xinqiang River; SR: sandbank area; YLR: Yuelai River. Region D: Huarong River Basin, Region E: Ouchi River and its tributaries, Region F: Xinqiang River Basin, and Region G: sandbank area. Red dots represent the main towns. a: Junshan district, b: Yueyang city, and c: Yueyang county.

Table 3: Comparison of the intensity grades of surface deformation in the flood seasons.

Intensity grade of surface deformation	Surface deformation area in the flood season					
	2017		2018		2019	
	Area (km^2)	%	Area (km^2)	%	Area (km^2)	%
Severe settlement	0.16	0.019	0	0	3.11	0.352
Moderate settlement	20.75	2.493	4.85	0.550	32.80	3.717
Slight settlement	211.37	25.396	187.88	21.282	200.53	22.727
Stable area	309.07	37.135	503.14	56.994	348.10	39.451
Slight sedimentation	277.50	31.805	185.93	21.061	257.34	29.165
Severe sedimentation	13.44	1.802	1.00	0.113	40.49	4.588

dominates the surface deformation in the East Dongting Lake area.

5. Discussion

5.1. Influence of Precipitation on Surface Deformation. Figure 7 shows the daily precipitation and deformation intensity grade percent during the study period. In Figure 7, it is observed that the number of days with precipitation in the flood season is more than that in the dry season. We discovered that, in the flood season, the surface settlement is positively correlated with precipitation events, indicating that deformation may be caused by flood events after heavy precipitation. Since the precipitation cannot infiltrate in time, a large amount of sediment is mixed with running water and promoted the erosion of surrounding river embankments. While in the dry season, the deformation is negatively correlated with precipitation events. Previous studies have pointed out that a small amount of precipitation will increase the soil viscosity, thereby playing a role in inhibiting soil erosion [31]. In the dry season, with less precipitation and weak intensity, after the tiny sediment particles are wetted by

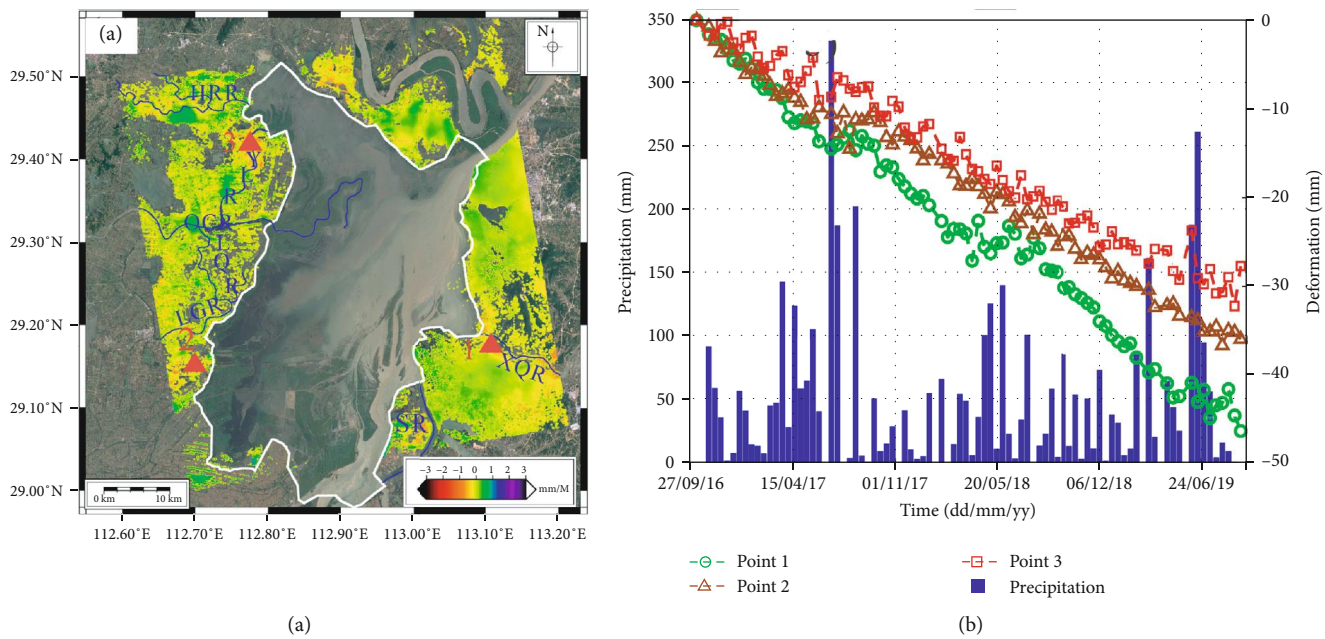


Figure 6: (a) Deformation rate obtained from the whole investigated period. White line represents the embankment of East Dongting Lake, and the blue lines are the rivers around East Dongting Lakes. HRR: Huarong River; LQR: Longqing River; LGR: Laogan River; XQR: Xinqiang River; YLR: Yuelai River; SR: sandbank area. Red triangles represent the feature points. (b) Time series deformations of the points of interest and the precipitation during the whole investigated period.

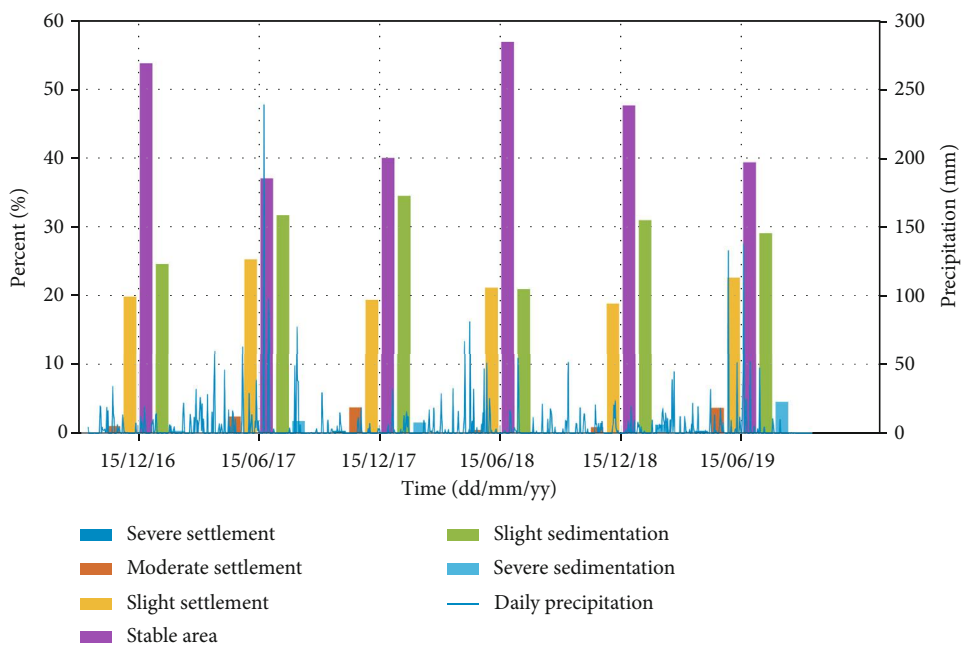


Figure 7: Daily precipitation and surface deformation intensity during the study period.

precipitation, the upper soil can more easily adhere to the surface with particles, and the cohesion is stronger.

Generally, the area is more stable in the dry season than the flood season, as the results in 2017 and 2019. However, there is an abnormal phenomenon in 2018, and the percent of the stable area is larger in the dry season than in the flood season. This is somewhat expected since the precipitation days and the maximum precipitation during the flood season in 2018 were both smaller than in 2017 and 2019. The precipitation and the number of precipitations during the dry season in 2018 were also smaller than in 2017 and 2019. The less precipitation in the dry season, the more obvious the deformation phenomenon [32]. Therefore, the surface deformation during the dry season in 2018 is the most obvious. Another reason is that during the flood season in 2017, there was a catastrophic flood disaster event in the Dongting Lake

area. This disaster event showed that a lot of silt and sediment remained in East Dongting Lake [30]. Therefore, during the dry season, the settlement of the soil occurred on the surface.

5.2. Causes of Surface Deformation in East Dongting Lake

5.2.1. Soil Erosion. As a geomorphic process persistently occurring over the earth's surface, soil erosion is one of the most serious environmental problems in the world nowadays, seriously threatening agriculture, soil nutrients, and surface integrity [33]. The wetland of Dongting Lake Plain is a key prevention zone for soil erosion in Hunan Province. Dongting Lake undertakes the sediment of major rivers in Hunan Province, and the soil erosion modulus of the basin is about 3970 t/km^2 [34]. The main type of soil erosion is water erosion, including surface erosion and gully erosion. There are also a small amount of soil erosions such as collapse and landslides in local areas. The flood season is the key sedimentation period of East Dongting Lake, and the deposited sediment exceeds the annual average sediment sedimentation. However, in the dry season, the erosion effect is greater than the sedimentation effect, and 10% of the sedimentation amount in the flood season are washed out of East Dongting Lake [35]. In addition, the sedimentation during the flood season indicates that soil erosion and transportation in the eastern mountainous area and western plain area are mainly related to storm runoff and flood runoff. The surface deformation monitored in this study shows in good agreement with the soil erosion feature. Therefore, we can speculate that soil erosion and siltation are the main reason for the surface deformation of East Dongting Lake.

The lake water contains a lot of sediment during the flood season, and in the dry season, the water line of the lake then drops, so the remaining sediment on the floodplain will flow into the catchment area with the flowing water and may erode the flowing area. Especially at the junction of the floodplain and the catchment area, the erosion effect becomes more obvious under the dual action of external flowing water and catchment water circulation. Settlement in East Dongting Lake area is mainly centralized in the south and west of the catchment, which indicates that the erosion in those areas is very significant. These two catchments are the main source of water flow in the dry season. Another settlement region is in the middle of the eastern part of East Dongting Lake. When the flowing water passes through the bend in the dry season, the pressure of the flowing water increases, the river water exerts a greater erosion force on the river bank, and the erosion effect is more obvious.

During the flood season, all the surface settlement regions are distributed in East Dongting Lake of the main external drainage. There are two main sources of river flooding. The first is the flooding of Yangtze River, which mainly occurs in the Ouchi River and its tributaries and Huarong River in the west of East Dongting Lake. There are also floods caused by the eruption of mountain torrents during the flood season in the eastern mountainous area, which mainly flow through the Xinqiang River area. Among them, only Xinqiang River and Laogan River keep a state of continuous erosion, and the erosion velocity is gradually increasing, indicating that surface settlement in the basin is getting more and more serious. In addition, according to the Yangtze River Sediment Loss Bulletin, the sediment flow in East Dongting Lake in 2018 was less than that in other years [36], which agreed with the results of this study.

The soil erosion of East Dongting Lake has the following characteristics based on the above analysis. First, the surface soil has been eroded and sedimentation both in the flood and dry periods. Second, in the dry season, erosion and sedimentation mainly occur in the internal river channels and catchment of the lake, and in the flood season, erosion mainly occurs in the area where the river flows. Third, flood disasters in flood season will have an impact on erosion and sedimentation in the later dry season.

5.2.2. Human Activities. In addition to soil erosion, human activities could be another reason for surface deformation. It had been found that human activities would affect the hydrological conditions of lakes [37]. Moreover, agricultural reclamation will cause soil destruction in the reclamation area, which is more susceptible to erosion, and increases in sediments in the river [38]. There are many polders in the East Dongting Lake Plain, which is one of the state commercial grain bases. Human activities and irrigation water extraction in the process of agricultural planting can also cause surface deformation in some areas. As expected, this part of the deformation is mainly concentrated in the agricultural planting stage, which corresponds to the flood period in this study. In addition, the study area also includes some urban areas. The rapid urbanization process, infrastructure construction, groundwater extraction, or other urban development processes may cause regional surface deformation. As shown in Figures 4 and 5, it is found that there are some small deformation areas in the urban area.

6. Conclusion

East Dongting Lake area is selected as the area of deformation monitoring with InSAR. It is found that the unique lake morphology of Dongting Lake, i.e., "river facies in dry period, lake facies in flood period," has a greater impact on the coherence. If the SBAS-InSAR method is used to the multi-years of dataset, serious errors will be caused in the subsequent deformation result estimation. Therefore, the Sentinel-1A data acquired on the East Dongting Lake are divided into different groups according to dry and flood seasons. Surface deformations are estimated for the flood and dry seasons, respectively.

The results show that during the dry season of Dongting Lake, the surface deformation was the most serious in 2018. Almost the entire lake area had settlement and sedimentation. The flood disasters in 2017 and 2019 were relatively slight, so the surface deformation areas are mainly concentrated around the already catchment areas along the rivers in the lake. During the flood season, we found that the most severe surface deformation occurred in 2017, indicating that the flood disaster not only had a serious impact on the

surface deformation inside the lake but also had an impact on the drainage network outside the East Dongting Lake.

Through the analysis of the deformation feature and precipitation, it is believed that the surface deformation detected by the SBAS-InSAR method is basically caused by lake mud erosion and sedimentation. Therefore, InSAR could be a potential method for the monitoring of soil erosion. Compared with the current methods [39, 40], SBAS-InSAR can provide quantitative results of soil erosion with high spatial resolution and avoid the time-consuming and laborious problem of manual verification in the general model.

Data Availability

Copernicus Sentinel data 2016-2019 processed by the European Space Agency (ESA) were retrieved from the ASF Distributed Active Archive Center (DAAC) (https://search.asf.alaska.edu/#/). The precipitation data are provided by the China Meteorological Data Service Center (http://data.cma.cn).

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

Copernicus Sentinel data 2016-2019 processed by the European Space Agency (ESA) were retrieved from the ASF Distributed Active Archive Center (DAAC) (https://vertex.daac.asf.alaska.edu/#). The precipitation data are provided by the China Meteorological Data Service Center (http://data.cma.cn). This study was supported by the National Natural Science Foundation of China (Nos. 41704001 and 42030112), the Hunan Natural Science Foundation (Nos. 2020JJ2043 and 2018JJ3347), the Special funds for the construction of Hunan innovative province (No. 2019GK5006), and the Project of Innovation-driven Plan of Central South University (No. 2019CX007).

References

[1] M. R. Rahman, Z. H. Shi, and C. Chongfa, "Soil erosion hazard evaluation—an integrated use of remote sensing, GIS and statistical approaches with biophysical parameters towards management strategies," *Ecological Modelling*, vol. 220, no. 13-14, pp. 1724–1734, 2009.

[2] Y. Gao, Z. Liang, B. Wang, Y. Wu, and P. Wu, "Wetland change detection using cross-fused-based and normalized difference index analysis on multitemporal Landsat 8 OLI," *Journal of Sensors*, vol. 2018, Article ID 8130470, 8 pages, 2018.

[3] W. H. McAnally, C. Friedrichs, D. Hamilton et al., "Management of fluid mud in estuaries, bays, and lakes. I: present state of understanding on character and behavior," *Journal of Hydraulic Engineering ASCE*, vol. 133, no. 1, pp. 9–22, 2007.

[4] L. G. Zheng, G. J. Liu, Y. Kang, and R. K. Yang, "Some potential hazardous trace elements contamination and their ecological risk in sediments of western Chaohu Lake, China," *Environmental Monitoring and Assessment*, vol. 166, no. 1-4, pp. 379–386, 2010.

[5] D. Fan, C. Li, D. Wang, P. Wang, A. W. Archer, and S. F. Greb, "Morphology and sedimentation on open-coast intertidal flats of the Changjiang Delta, China," *Journal of Coastal Research*, vol. 43, pp. 23–35, 2004.

[6] F. Xu, J. Tao, Z. Zhou, G. Coco, and C. Zhang, "Mechanisms underlying the regional morphological differences between the northern and southern radial sand ridges along the Jiangsu Coast, China," *Marine Geology*, vol. 371, pp. 1–17, 2016.

[7] G. Desir, F. Gutiérrez, J. Merino et al., "Rapid subsidence in damaging sinkholes: measurement by high-precision leveling and the role of salt dissolution," *Geomorphology*, vol. 303, pp. 393–409, 2018.

[8] K. Yang, L. Yan, G. Huang, C. Chen, and Z. Wu, "Monitoring building deformation with InSAR: experiments and validation," *Sensors*, vol. 16, no. 12, p. 2182, 2016.

[9] H. A. Zebker and R. M. Goldstein, "Topographic mapping from interferometric synthetic aperture radar observations," *Journal of Geophysical Research*, vol. 91, no. B5, pp. 4993–4999, 1986.

[10] A. K. Gabriel, R. M. Goldstein, and H. A. Zebker, "Mapping small elevation changes over large areas: differential radar interferometry," *Journal of Geophysical Research*, vol. 94, no. B7, pp. 9183–9191, 1989.

[11] A. Ferretti, C. Prati, and F. Rocca, "Permanent scatterers in SAR interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 1, pp. 8–20, 2001.

[12] P. Berardino, G. Fornaro, R. Lanari, and E. Sansosti, "A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 11, pp. 2375–2383, 2002.

[13] J. Dong, L. Zhang, M. S. Liao, and J. Y. Gong, "Improved correction of seasonal tropospheric delay in InSAR observations for landslide deformation monitoring," *Remote Sensing of Environment*, vol. 233, p. 111370, 2019.

[14] C. H. Lu, C. F. Ni, C. P. Chang, J. Y. Yen, and R. Y. Chuang, "Coherence difference analysis of sentinel-1 SAR interferogram to identify earthquake-induced disasters in urban areas," *Remote Sensing*, vol. 10, no. 8, p. 1318, 2018.

[15] X. Yu, J. Hu, and Q. Sun, "Estimating actual 2D ground deformations induced by underground activities with cross-heading InSAR measurements," *Journal of Sensors*, vol. 2017, Article ID 3170506, 12 pages, 2017.

[16] B. Hu, B. Yang, X. Zhang, X. Chen, and Y. Wu, "Time-series displacement of land subsidence in Fuzhou downtown, monitored by SBAS-InSAR technique," *Journal of Sensors*, vol. 2019, Article ID 3162652, 12 pages, 2019.

[17] M. Tzouvaras, C. Danezis, and D. G. Hadjimitsis, "Small scale landslide detection using sentinel-1 interferometric SAR coherence," *Remote Sensing*, vol. 12, no. 10, p. 1560, 2020.

[18] Y. Tang, Y. H. Xie, F. Li, and X. S. Chen, "Spatial distribution of emergent herbaceous wetlands in the East Dongting Lake during the last twenty years based landsat data," *Resources and Environment in the Yangtze Basin*, vol. 22, pp. 1484–1492, 2013.

[19] B. F. Wu, J. L. Huang, and L. B. Shen, "Analysis and evaluation of flood-control functions of wetland taking Dongting Lake as an example," *Geographical Research*, vol. 19, pp. 189–193, 2000.

[20] J. Li, H. Yin, J. Chang, C. Lu, and H. Zhou, "Sedimentation effects of the Dongting Lake area," *Journal of Geographical Sciences*, vol. 19, no. 3, pp. 287–298, 2009.

[21] Y. Yuan, G. Zeng, J. Liang et al., "Effects of landscape structure, habitat and human disturbance on birds: a case study in East Dongting Lake wetland," *Ecological Engineering*, vol. 67, pp. 67–75, 2014.

[22] C. Hu, Z. M. Deng, Y. H. Xie, X. S. Chen, and F. Li, "The risk assessment of sediment heavy metal pollution in the East Dongting Lake Wetland," *Journal of Chemistry*, vol. 2015, Article ID 835487, 8 pages, 2015.

[23] G. Jun-Feng, Z. Chen, J. Jia-Hu, and H. Qun, "Analysis of deposition and erosion of Dongting Lake by GIS," *Journal of Geographical Sciences*, vol. 11, no. 4, pp. 402–410, 2001.

[24] A. Hooper, "A multi-temporal InSAR method incorporating both persistent scatterer and small baseline approaches," *Geophysical Research Letters*, vol. 30, article L16302, 2008.

[25] F. Casu, M. Manzo, and R. Lanari, "A quantitative assessment of the SBAS algorithm performance for surface deformation retrieval from DInSAR data," *Remote Sensing of Environment*, vol. 102, no. 3-4, pp. 195–210, 2006.

[26] A. Sowter, L. Bateson, P. Strange, K. Ambrose, and M. F. Syafiudin, "DInSAR estimation of land motion using intermittent coherence with application to the South Derbyshire and Leicestershire coalfields," *Remote Sensing Letters*, vol. 4, no. 10, pp. 979–987, 2013.

[27] R. M. Goldstein and C. L. Werner, "Radar interferogram filtering for geophysical applications," *Geophysical Research Letters*, vol. 25, no. 21, pp. 4035–4038, 1998.

[28] O. Mora, J. J. Mallorqui, and A. Broquetas, "Linear and nonlinear terrain deformation maps from a reduced set of interferometric sar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, pp. 2243–2253, 2003.

[29] R. F. Hanssen, *Radar Interferometry: Data Interpretation and Error Analysis (Vol. 2)*, Springer Science & Business Media, 2001.

[30] J. Li, W. Zheng, and X. Xu, "Information extraction and analysis of food disaster in Dongting Lake area base on sentinel-1," *Anhui Agricultural Science Bulletin*, vol. 26, pp. 151–153, 2020.

[31] H. El Kateb, H. Zhang, P. Zhang, and R. Mosandl, "Soil erosion and surface runoff on different vegetation covers and slope gradients: a field experiment in Southern Shaanxi Province, China," *Catena*, vol. 105, pp. 1–10, 2013.

[32] J. H. Xiao, S. X. Zhang, Y. Zhang, M. K. Li, and M. Hong, "Research on the response of land water reserve to rainfall in Yunnan Province," in *Proceedings of the 2019 China Earth Sciences Joint Academic Conference*, vol. 12, pp. 4–6, Beijing, China, 2019.

[33] L. Wang, J. Huang, Y. Du, Y. Hu, and P. Han, "Dynamic assessment of soil erosion risk using Landsat TM and HJ satellite data in Danjiangkou Reservoir Area, China," *Remote Sensing*, vol. 5, no. 8, pp. 3826–3848, 2013.

[34] S. Zhang and C. Zhu, "Soil loss and its effect on flooding catastrophe in Yangtze Drainage Basin," *Journal of Soil and Water Conservation*, vol. 15, pp. 9–13, 2001.

[35] W. F. Zhou, "Soil erosion and sediment deposition in Dongting Lake," *Water and Soil Conservation in China*, vol. 6, pp. 22–25+65, 1992.

[36] "Changjian water resources commission of the ministry of water resources," in *Annual Yangtze River Sediment Bulletin*, OLE, 2020, http://www.cjw.gov.cn/zwzc/bmgb/2020gb/49234.html.2020-07-31.

[37] Y. Zhou, J. Li, Y. Zhang, X. Zhang, and X. Li, "Enhanced lakebed sediment erosion in Dongting Lake induced by the operation of the Three Gorges Reservoir," *Journal of Geographical Sciences*, vol. 25, no. 8, pp. 917–929, 2015.

[38] Y. Du, H. P. Xue, S. J. Wu, F. Ling, F. Xiao, and X. H. Wei, "Lake area changes in the middle Yangtze region of China over the 20th century," *Journal of Environmental Management*, vol. 92, no. 4, pp. 1248–1255, 2011.

[39] B. Wang, F. Zheng, and Y. Guan, "Improved USLE-K factor prediction: a case study on water erosion areas in China," *International Soil and Water Conservation Research*, vol. 4, no. 3, pp. 168–176, 2016.

[40] C. G. Karydas, T. Sekuloska, and G. N. Silleos, "Quantification and site-specification of the support practice factor when mapping soil erosion risk associated with olive plantations in the Mediterranean island of Crete," *Environmental Monitoring and Assessment*, vol. 149, no. 1-4, pp. 19–28, 2009.