

# Accelerated Genetic Algorithm Based on Search-space Decomposition for Change Detection in Remote Sensing Images

Caihong Mu, *Associate Member, IEEE*, Chengzhou Li, Yi Liu, Rong Qu, *Senior Member, IEEE*, and Licheng Jiao, *Fellow, IEEE*

**Abstract**—Detecting change areas among two or more remote sensing images is a key technique in remote sensing. It usually consists of generating and analyzing a difference image thus to produce a change map. Analyzing the difference image to obtain the change map is essentially a binary classification problem, and can be solved by optimization algorithms. This paper proposes an accelerated genetic algorithm based on search-space decomposition (SD-aGA) for change detection in remote sensing images. Firstly, the BM3D algorithm is used to preprocess the remote sensing image to enhance useful information and suppress noises. The difference image is then obtained using the logarithmic ratio method. Secondly, after saliency detection, fuzzy c-means algorithm is conducted on the salient region detected in the difference image to identify the changed, unchanged and undetermined pixels. Only those undetermined pixels are considered by the optimization algorithm, reducing the search space significantly. Inspired by the idea of the divide-and-conquer strategy, the difference image is also decomposed into sub-blocks with a method similar to down-sampling, where only those undetermined pixels are analyzed and optimized by SD-aGA in parallel. The category labels of the undetermined pixels in each sub-block are optimized according to an improved objective function with neighborhood information. Finally, the decision results of the category labels of all the pixels in the sub-blocks are remapped to their original positions in the difference image and then merged globally. Decision fusion is conducted on each pixel based on the decision results in the local neighborhood to produce the final change map. The proposed method is tested on five diverse remote sensing image benchmark datasets and compared against six state-of-the-art methods. Results demonstrate the excellent performance of the proposed SD-aGA on handling noises and detecting the changed areas accurately. In particular, compared with the traditional genetic algorithm, SD-aGA can

obtain a much higher detection accuracy with a much less computational time.

**Index Terms**—remote sensing image, change detection, evolutionary optimization, genetic algorithm, search space decomposition

## I. INTRODUCTION

CHANGE detection is one of the most important research topics in remote sensing image processing, and has been widely used in environmental monitoring [1], natural disaster damage assessment [2], forest resources monitoring [3], agricultural surveys [4] and so on. In change detection, two or more remote sensing images are acquired at different times in the same geographical area, to qualitatively or quantitatively analyze and determine surface changes. Remote sensing is a non-contact technique which acquires survey data of the target at distance. Images acquired by remote sensing can be categorized into microwave remote sensing images and optical remote sensing images. Synthetic Aperture Radar (SAR) acquires data actively with microwave, independent of atmospheric and sunlight conditions. However, speckle noises usually appear in SAR images, thus leads to difficulties to image interpretation. The resolution of optical remote sensing images is high, but the image quality can be affected by weather factors such as clouds, rain, fog, and haze. This paper focuses on SAR images and panchromatic spectral images.

Generally, change detection methods in remote sensing images can be classified into three types, namely supervised, semi-supervised and unsupervised methods. Both supervised and semi-supervised methods require class labels in the sample data to train effective classifiers. However, samples are often labeled manually and is in general costly. In contrast, unsupervised methods produce the final change map by just analyzing the difference image, thus are more widely used for change detection in remote sensing images. To analyze potential changes, a difference image is generated by comparing the coregistered multitemporal images pixel by pixel using differencing (subtraction operator) or rationing (ratio operator) techniques. We focus on unsupervised methods in this paper.

Unsupervised change detection in remote sensing images is consists of the following three main steps [5].

- 1) Preprocessing: two remote sensing images obtained at different time points in the same location are preprocessed by coregistration, radiation correction and denoising;
- 2) Difference image generation: the difference image of two remote sensing images is preliminarily generated for the

This work was supported by the National Natural Science Foundation of China (No. 61672405), Project supported the Foundation for Innovative Research Groups of the National Natural Science Foundation of China (61621005), the Fundamental Research Funds for the Central Universities (No. JB170204), the National Natural Science Foundation of China (U1701267, 61473215, 61876141, 61772399, 61773304, 61773300, and 61806156), the Fund for Foreign Scholars in University Research and Teaching Programs (the 111 Project) (No. B07048), the Major Research Plan of the National Natural Science Foundation of China (Nos. 91438201 and 91438103), and the Program for Cheung Kong Scholars and Innovative Research Team in University (No. IRT\_15R53).

Caihong Mu, Chengzhou Li, and Licheng Jiao are with Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, International Research Center for Intelligent Perception and Computation, Joint International Research Laboratory of Intelligent Perception and Computation, School of Artificial Intelligence, Xidian University, Xi'an, 710071, China (e-mail: mucaihongxd@foxmail.com; 1093533413@qq.com; lchjiao@mail.xidian.edu.cn).

Yi Liu is with School of Electronic Engineering, Xidian University, Xi'an, 710071, China (e-mail: yiliuxd@foxmail.com).

Rong Qu is with School of Computer Science, University of Nottingham, Nottingham, NG8 1BB, UK (e-mail: rong.qu@nottingham.ac.uk).

subsequent analysis;

3) Analyzing the difference image: the change map is produced according to the gray value and relevant neighborhood information of pixels in the difference image.

Among these, generating and analyzing difference image are the two key steps in change detection [6], [7]. In change detection in remote sensing images, analyzing the difference image to obtain the change map is essentially a binary classification problem. Threshold and clustering methods are among the most widely used unsupervised change detection methods in remote sensing images.

Threshold method builds a statistical model by analyzing the grey values of the difference image, and then uses the Kittler Illingworth (KI) or expectation maximization (EM) algorithm to find the appropriate threshold. Those pixels with gray values greater than the threshold are classified as the changed class; otherwise they are in the unchanged class. Bruzzone et al. [8] proposed an unsupervised change detection method by combining the EM and Markov random field (MRF) algorithms. Firstly, the EM algorithm was used to estimate the statistical distributions of changed and unchanged categories in the difference image. Then spatial texture information was analyzed by using MRF. Finally, Bayes rule is employed to produce the decision threshold that minimizes the probability of the overall error in the change detection process. Bazi et al. [9] produced the change map by a modified KI criterion based on the generalized Gaussian assumption to model the distributions of changed and unchanged classes. Moser et al. [10] developed an automatic thresholding technique in which the KI minimum-error threshold algorithm is generalized to take into account the non-Gaussian distribution of the amplitude values of SAR images. Sumaiya et al. [11] combined the Gabor filter and KI threshold (GF-KI) techniques to detect changes in the spatial domain of the SAR images. Generally, in threshold methods, the optimal threshold is determined mainly through the statistical model. It is simple in operation and low in computational complexity, however, its classification accuracy is not high [12].

K-means and fuzzy c-means are among the most widely used clustering analysis methods for clustering the difference image in change detection. K-means method generates two suitable cluster centers iteratively, where pixels are assigned to the changed or unchanged class based on a hard partition. Fuzzy c-means clustering (FCM) determines the cluster center based on a soft partition by generating a membership degree matrix using fuzzy sets. Both K-means and FCM only use the gray information of all the pixels in the image to obtain the clustering center. They do not consider the spatial information of pixels, thus are sensitive to noises in the image.

In the literature, neighborhood information of pixels in images, which can provide more useful spatial information for the central pixel, has been widely used in FCM algorithms to improve their anti-noise performance and reduce detection errors. Ahmed et al. [13] introduced spatial neighborhood information to the objective function, and proposed the improved FCM\_S algorithm, however, at an increased computational time. Szilagyí et al. [14] proposed an Enhanced

FCM which accelerated the image segmentation process. Krinidis et al. [15] proposed a fuzzy local information c-means (FLICM) method to balance noise suppression and detail preservation by adding fuzzy local spatial information and gray information to the objective function. Gong et al. [16] added MRF to improve the FCM approach, named as MRFFCM, by modifying the membership of each pixel according to the MRF-based spatial information built in the pivotal energy function. The energy function is altered by utilizing not only the memberships but also the number of neighborhood pixels in the same class. In summary, the accuracy of the FCM detection algorithms can be improved using image neighborhood information. However, since remote sensing images are severely affected by noises, using just improved FCM algorithms to analyze the difference map may still lead to high detection errors.

In recent years, various other techniques have also been applied to improve change detection in SAR images. Li et al. [17] exploited Gabor wavelet to yield the representation of difference image at multiple scales and orientations, and an advanced two-level cascade clustering to generate the final change map. In [18], Li et al. put forward a multiobjective fuzzy clustering method with two conflicting objectives to preserve details and remove noises, respectively. Membership values of pixels are updated to find the optimal trade-off between the two objectives. Gong et al. [19] used pre-classification to obtain samples from two original maps instead of using the difference image. A deep neural network was then established to determine change regions. This algorithm omits the step of generating a difference image, and thus the detection result is no longer dependent on the quality of the difference image. In [20], Gao et al. employed Gabor wavelets and FCM to select image patches with a center of high likeliness of being changed or unchanged. A PCA-Net classification model is then trained using these patches, results of which are combined with pre-classification results to form the final change map. Shang et al. [21] proposed a novel self-paced learning algorithm to detect changes in SAR images. FCM was first used to pre-classify the difference image and then self-paced learning is conducted on a high-quality training set with selected pixels to train a classifier.

Considering analyzing the difference image is essentially a binary classification problem, it can also be regarded as a combinatorial optimization problem and solved by evolutionary algorithms. That is, the pixels in the difference image are labeled as changed or unchanged class, denoted as 1 or 0, respectively. The change map is thus simply a matrix of binary numbers [22]. If the size of the remote sensing image is  $M \times N$ , the states of the change map will have  $2^{M \times N}$  combinations. Celik [23] proposed a simple GA to find the final change detection mask by evolving an initial random binary change detection mask. However, the algorithm only uses the gray values in the difference image. It is difficult to obtain good results in remote sensing images with strong noises. Moreover, when the size of remote sensing images is too large, the search space is too large for the simple GA to obtain good detection result. As one of the highly efficient global search algorithms,

evolutionary algorithms have been used successfully in image processing to optimize the parameters of image processing methods, including image segmentation [24], [25] and image classification [26]. However, directly using GA to optimize the binary matrix with the same size as that of the original remote sensing image is still a challenging attempt considering the very large search space.

In [27], Mu et al. proposed a method based on saliency map guidance and acceleration genetic algorithm (S-aGA) for change detection in SAR images to reduce the computational cost and mitigate noise sensitivity. Firstly, the saliency detection model based on a human visual mechanism was adopted to determine the salient region in the difference image and obtain the rough locations of the change areas. Then, pixels in the salient regions were classified into changed, unchanged, and undetermined classes using the FCM algorithm. GA was employed to optimize pixels in the undetermined class. At the same time, correlation of neighborhoods of pixels is utilized to provide heuristic information to adaptively generate mutation probability, thereby accelerating the GA. In addition, the neighborhood information of pixels is included in the objective function to reduce the negative impact caused by the noise points and thus improve noise immunity of the algorithm. Compared to simple GA, S-aGA significantly improved the time efficiency and quality of the resulting change map. However, the large search space of the algorithm still presents a challenging problem if there are too many undetermined pixels.

In this paper, we proposed an accelerated genetic algorithm based on search-space decomposition (SD-aGA) for change detection in remote sensing images. In SD-aGA, further improvements are made to enhance the anti-noise performance and efficiency by reducing the search space. The contributions of the proposed method are as follows:

(1) Block-matching and 3D filtering (BM3D) denoising is applied to the two remote sensing images in the pre-processing, thus to improve the quality of the difference image. Therefore, the process of obtaining the changed map from difference image analysis is less affected by the noise.

(2) The search space is further reduced based on sub-block decomposition. Using the divide-and-conquer strategy, the difference image is decomposed into multiple sub-blocks, thus the original optimization problem is divided into several smaller optimization problems on sub-blocks, reducing the search space significantly.

(3) Evolutionary optimization is independently applied in parallel to the optimization of sub-blocks, reducing the overall computational time of the algorithm.

(4) The results of sub-block optimization are merged globally to improve the noise immunity and detection accuracy of the change map.

Based on the experiments on five diverse real remote sensing images, it can be concluded that the anti-noise performance and efficiency of SD-aGA have been significantly improved compared with S-aGA in [27]. The improved performance is also demonstrated by comparisons against several state-of-the-art change detection methods in the existing literature.

The paper is organized as follows: Section 2 describes the proposed method in detail. The dataset and experimental setup are presented in Section 3. Section 4 provides parametric analysis, the experimental results and comparisons to the existing methods. Section 5 summarizes the paper and provides directions of future research.

## II. THE PROPOSED SD-AGA METHOD

Consider two remote sensing images  $I_1$  and  $I_2$ , acquired at two different time  $t_1$  and  $t_2$  in the same area, respectively, and coregistered for further analysis. The size of the images is  $M \times N$ .  $I_1(x, y)$  and  $I_2(x, y)$  denote the gray value of the pixel on the  $x$ th row and  $y$ th column in the two images, respectively,  $1 \leq x \leq M, 1 \leq y \leq N$ .

The purpose of change detection in remote-sensing images is to analyze the difference between images  $I_1$  and  $I_2$ , by labelling pixels at their corresponding positions as changed or unchanged, denoted as 1 or 0, respectively. A change map  $CM = \{CM(x, y), 1 \leq x \leq M, 1 \leq y \leq N\}$  is eventually obtained,  $CM(x, y) \in \{0, 1\}$ . The change map can then be used to visually display and evaluate the changes in the same region at time  $t_1$  and  $t_2$ .

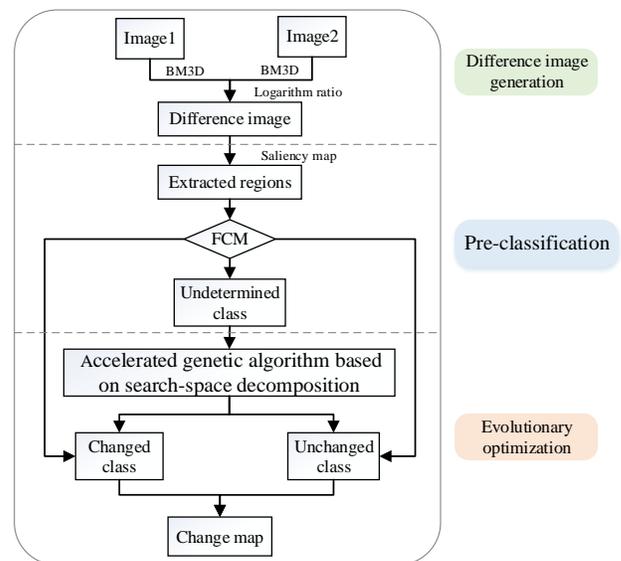


Fig. 1. Flow chart of the proposed SD-aGA algorithm

The flow chart of the proposed SD-aGA method is shown in Figure 1. It includes mainly three stages, namely the generation of difference image, pre-classification, and evolutionary optimization on the decomposed sub-blocks, as follows.

(1) Difference image generation. The BM3D algorithm is adopted to preprocess the two remote sensing images acquired at two different times in the same area to reduce the influence of noises and improve their quality. Logarithm ratio method is then employed to generate the difference image.

(2) Pre-classification. After applying the saliency map extraction and the FCM algorithm to pre-classify the pixels as changed, unchanged or undetermined class, the issue of analyzing the entire original difference image is thus converted into the classification of the undetermined pixels.

(3) Generation of the change map using SD-aGA on the decomposed search space. Firstly, the search space of the difference image is decomposed into sub-blocks. The class labels of undetermined pixels in each sub-block are then optimized by using the accelerated GA in parallel. In optimizing the sub-blocks, neighborhood information of pixels in the original difference image is considered. Finally, making use of local consistency of images, decision fusion is carried out based on the results of all the sub-blocks to generate the final change map using the majority voting strategy.

#### A. Difference Image Generation

To address the key issue of noise immunity when generating the difference image, denoising is required in the pre-processing of remote sensing images. Commonly used image denoising methods include average filtering, median filtering, and so on. Block-matching and 3D filtering (BM3D) [28] is applied here to denoise the two original images by filtering the images to find similar blocks. It performs well in suppressing noises and restoring details of the image, the resulting denoised image has a high peak signal-to-noise ratio.

On the two denoised images  $I_1'$  and  $I_2'$  generated using the BM3D algorithm, the logarithmic ratio operator as shown in (1) is used to obtain the difference image  $DI$ .

$$DI = \left| \log \frac{I_2' + 1}{I_1' + 1} \right|, \quad DI \in R^{M \times N} \quad (1)$$

#### B. Pre-classification

In SD-aGA, the automatic search space reduction method proposed in S-aGA [27] is applied as the pre-classification to reduce the search space. This includes two steps, salient regions extraction and FCM clustering.

Image saliency detection methods have been proposed based on machine vision [29]-[31]. Saliency detection analyzes the difference image to identify the salient area by identifying great changes in pixels, thus to extract regions of interest and narrow down the search space for evolutionary optimization.

In our proposed SD-aGA, the simple and effective context-ware saliency detection method developed in [31] is applied and the same parameter setting as that in [31] is adopted. The details can be referred to [31].

After obtaining the saliency map of the difference image, OTSU [32], a widely used image segmentation algorithm, is executed to obtain the binary saliency map  $S$ ,  $S \in R^{M \times N}$ , as defined in (2).

$$S(x,y) = \begin{cases} 1, & \text{saliency regions} \\ 0, & \text{non-saliency regions} \end{cases}, \quad 1 \leq x \leq M, 1 \leq y \leq N \quad (2)$$

The saliency based difference image  $DIS$  contains only the salient regions, and is obtained by using (3).

$$DIS(x,y) = \begin{cases} DI(x,y), & \text{if } S(x,y) = 1, 1 \leq x \leq M, 1 \leq y \leq N \\ 0, & \text{if } S(x,y) = 0 \end{cases} \quad (3)$$

$DIS$  is then analyzed by the FCM algorithm to further reduce the search space. In  $DIS$ , pixels can be classified by the FCM algorithm into three categories according to their gray values, namely changed class  $w_c$ , unchanged class  $w_u$ , and undetermined class  $w_{ud}$  as follows. Based on the three cluster

centers found by the FCM algorithm, all the pixels in  $DIS$  are divided into three classes, and the membership degrees  $U$  is obtained,  $U = [u_{x,y}^r]$ ,  $U \in R^{c \times MN}$ , where  $u_{x,y}^r$  denotes the membership of gray value of pixel  $DIS(x,y)$  in the  $r$ -th cluster, and  $v_r$  is the cluster center,  $r \in \{1,2,3\}$ . Then, setting  $num1 = \arg_r \{\max\{v_r\}\}$ ,  $num2 = \arg_r \{\min\{v_r\}\}$ , according to the membership matrix  $U$ , a pre-classified image  $C$  is obtained, each pixel of  $C$  is marked with a category label  $C_{x,y}$ ,  $C_{x,y} \in \{\omega_c, \omega_u, \omega_{ud}\}$  as in (4).

$$C_{x,y} = \begin{cases} \omega_c, & \text{if } num1 = \arg_r \{\max\{u_{x,y}^r\}\}, \max\{u_{x,y}^r\} > \tau \\ \omega_u, & \text{if } num2 = \arg_r \{\max\{u_{x,y}^r\}\}, \max\{u_{x,y}^r\} > \tau \\ \omega_{ud}, & \text{otherwise} \end{cases} \quad (4)$$

In (4),  $\tau$  is the threshold of the degree of membership.  $\tau = 0.90$  is set to ensure that the pixel marked as changed or unchanged class has a high degree of membership with the relevant cluster center.

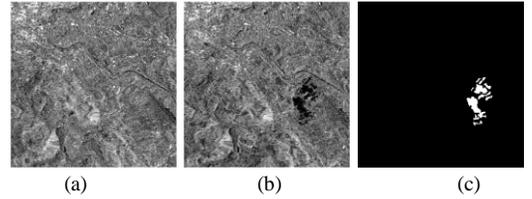


Fig. 2. The Bern dataset. (a) SAR image acquired in April 1999. (b) SAR image acquired in May 1999 after flood. (c) The reference image.

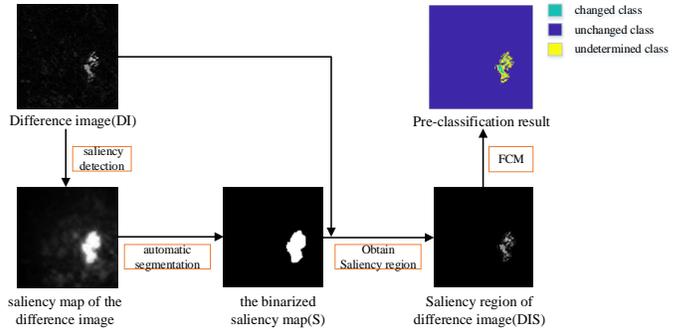


Fig. 3. Flowchart of pre-classification in the difference image

The SAR image dataset Bern as shown in Figure 2 is used to verify the effect of automatic search space reduction. The dataset contains two coregistered SAR images of size  $301 \times 301$  and also the available reference image. As shown in Figure 3, the search space is automatically reduced after saliency detection and FCM applied to  $DI$ . The pre-classification method can effectively reduce the search space thus reduce the computational cost for evolutionary optimization in the third stage.

#### C. Accelerated GA based on search-space decomposition (SD-aGA)

##### 1) The sub-block decomposition method

After the pre-classification, the search space is reduced greatly. However, when the size of  $DIS$  is large and the proportion of salient regions in  $DIS$  is high, the number of

pixels of the undetermined class may still be very large, presenting still a very large search space for the evolutionary optimization algorithm. In this subsection, sub-block decomposition is proposed to reduce the search space further.

Within the spatial neighborhood of an image, the gray values of the pixels generally describe the same object. That is, the pixel values within a region in the image have certain local consistency, thus one pixel can be used to represent its surrounding region.  $DIS$  is thus decomposed into several sub-blocks using a method similar to down-sampling as follows. First,  $DIS$  is edge-extended (i.e., the pixels on the edges are copied to extend  $DIS$  and ensure that  $DIS$  can be divided into an integer number of sampling windows), resulting into an extended image denoted as  $DIS'$ ; then  $DIS'$  is uniformly divided into non-overlapping sampling windows of size  $K \times K$ , and one pixel is extracted at each of these  $K \times K$  positions to represent each window. These representative pixels are then recombined respectively to form a set of  $K \times K$  sub-blocks, represented as set  $\{DIS_i\}$ ,  $i = 1, 2, \dots, K \times K$ . Figure 4 presents a schematic diagram of the sub-block decomposition method on  $DIS$  with sampling windows of  $2 \times 2$ . The downsized  $DIS_1$ ,  $DIS_2$ ,  $DIS_3$  and  $DIS_4$  are obtained by replacing each window in the original image with the representative pixel at position  $i = 1, 2, 3, 4$  in each window, respectively.

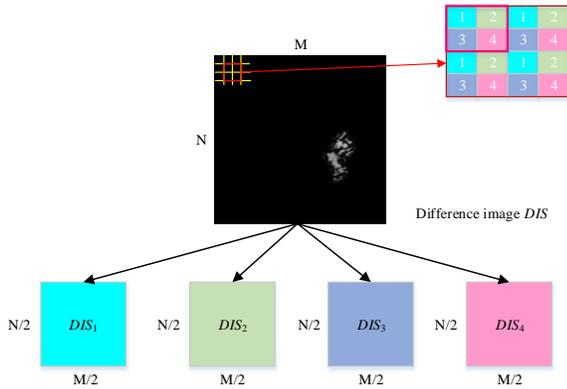


Fig. 4. When  $K = 2$ , sub-block decomposition diagram of  $DIS$

Table 1 shows the size of the search space for  $DIS$  and  $\{DIS_i\}$ . Assume that the image size of the difference image after the edge extension is  $M \times N$ , and all pixels in  $DIS$  are undetermined pixels for convenience. In  $\{DIS_i\}$ , the size of each sub-block is  $\frac{M}{K} \times \frac{N}{K}$ . The search space of the set  $\{DIS_i\}$  is reduced to  $K \times K \times 2^{\frac{M}{K} \times \frac{N}{K}}$ . The search space is thus reduced exponentially as the size of  $DIS$  increases. For a specific  $DIS$  with  $M = 100$ ,  $N = 100$ , and  $K = 4$ , the size of the original search space is  $2^{10000}$ , and the search space for the set of sub-blocks of difference image is reduced significantly to  $2^{629}$ .

Different from the decomposition method that directly divides  $DIS$  into several sub-blocks, each only containing local information of  $DIS$ , our decomposition method can obtain the sub-blocks that contain the global information of  $DIS$ , as each sub-block after the decomposition can be regarded as an

approximate representation of  $DIS$ . Thus, the undetermined pixels are nearly evenly distributed into all the sub-blocks rather than located in one or two sub-blocks. This is helpful for reducing the real search space of each sub-block. Furthermore, the obtained sub-blocks are extracted from different locations in the sampling window, respectively, and they are different from each other. When they are merged in the following stage of decision fusion, they will provide more useful information.

Table 1. The size of the search space for the algorithm before and after search-space decomposition

Image data	Size of search space
$DIS$	$2^{M \times N}$
sub-block $DIS$ sets $\{DIS_i\}$	$K \times K \times 2^{\frac{M}{K} \times \frac{N}{K}}$

Table 2. Statistics of pixel number in pre-classified image

		$DIS$	$DIS_1$	$DIS_2$	$DIS_3$	$DIS_4$
Salient regions	Changed class	351	90	92	83	86
	Unchanged class	2,040	520	509	505	506
	<b>Undetermined class</b>	<b>1,390</b>	<b>343</b>	<b>342</b>	<b>353</b>	<b>352</b>
Non-salient regions	Unchanged class	87,423	21,848	21,858	21,860	21,857
Total number of pixels		91,204	22,801	22,801	22,801	22,801

Using the Bern dataset shown in Figure 2 as an example,  $DIS$  is edge-extended firstly. Table 2 shows that the number of pixels in salient regions are 3,781, and the number of undetermined pixels found is 1,390. After using the sub-block decomposition method in  $DIS$ , the undetermined pixels of  $DIS_i$  is greatly reduced to below a quarter of the size of  $DIS$ .

## 2) Accelerated GA

In our previous work, S-aGA [27] used neighbor information as heuristic information to design the mutation operator in the accelerated GA. It showed to perform well on noise immunity and converge quickly. In S-aGA, three improvements have been made to determine the labels of the undetermined pixels in the difference image: (1) A two-dimensional encoding of individuals is used to effectively maintain the spatial structure of the individuals (one individual is actually one potential change map); (2) The neighborhood information is introduced to the objective function, to improve the anti-noise effect of the change detection method based on evolutionary optimization. (3) The neighborhood information of undetermined pixels in  $DIS$  is used as heuristic information to adjust the mutation probability of the undetermined pixels, thereby accelerating the convergence of GA. In this subsection, S-aGA will be improved and applied to the sub-blocks obtained by the proposed decomposition method.

**Binary two-dimensional encoding method.** The binary two-dimensional encoding method preserves the spatial structure of the individual. The size of the binary two-dimensional encoding matrix is the same as that of  $DIS_i$ , i.e.  $\frac{(M+m)}{K} \times \frac{(N+n)}{K}$ , where  $M \times N$  represents the size of  $DIS$ ,  $m$  and  $n$  represent two smallest positive integers of the rows and columns that make up  $(M+m)$  and  $(N+n)$  so that they are divisible by  $K$ , respectively. During the initialization of each

individual, the points at positions of all pixels in the non-saliency region detected by the saliency model and the unchanged / changed pixels determined using the FCM in the salient region are encoded as 0 / 1, respectively, and no evolution optimization operation is performed on them further. Those points at positions of the undetermined pixels pre-classified by the FCM in the salient region are randomly encoded as 0 or 1, and the position indices ( $Index_i$ ) of these undetermined pixels in the two-dimensional coding matrix are established. The evolutionary optimization will be implemented at only these positions of undetermined pixels to search for the optimal solution.

#### The objective function with neighborhood information.

The selection of individuals in evolutionary optimization is based on the fitness values estimated by the objective function, which then determines the probabilities, based on which individuals are selected to execute genetic operators. In  $DIS$ , pixel values of the non-saliency region have been set to 0, so the objective function only evaluates pixels of the salient region. Suppose  $S_i$  is the binary saliency map of sub-block  $DIS_i$ , the number of individuals in the population is  $H$ , and  $CM_h$  denotes an individual in SD-aGA for sub-block  $DIS_i$ ,  $h \in \{1, 2, \dots, H\}$ . The objective function of  $CM_h$  is  $OF$ . The individual fitness value  $Fitness$  is calculated by formula (5).

$$Fitness = \frac{1}{OF} \quad (5)$$

$$OF = \frac{1}{M_1 \times M_2} \sum_{r=0}^1 M_r \sum_{\forall (x,y) \in R_r} [(DIS_i(x,y) - v_r)^2 + G_r(x,y)] \quad (6)$$

$$G_r(x,y) = \frac{1}{D} \sum_{\substack{j \in L_{(x,y)} \\ j \neq center(L_{(x,y)})}} \frac{1}{d_j + 1} \times (1 - u_{jr})^2 \times (L_j - v_r)^2 \quad (7)$$

where  $r \in \{0,1\}$ ,  $R_0$  denotes all unchanged pixels in the salient region, and  $R_0 = \{(x,y) | CM_h(x,y) = 0, S_i(x,y) = 1\}$ ,  $R_1$  represents all changed pixels in the salient region,  $R_1 = \{(x,y) | CM_h(x,y) = 1, S_i(x,y) = 1\}$ .  $L_{(x,y)}$  denotes the  $3 \times 3$  neighborhood of position  $(x,y)$  in  $DIS$  mapped by using the position in  $DIS_i$ .  $j \neq center(L_{(x,y)})$  means that  $j$  is not at the center position of  $L_{(x,y)}$ .  $d_j$  represents the spatial distance between the  $j$ -th pixel in the neighborhood and the center point  $(x,y)$ . The statistical variable  $M_r$  for the number of changed and unchanged pixels, the  $r$ -th cluster center  $v_r$  and the spatial distance weight  $D$  are defined, respectively, in (8)-(10) as follows.

$$M_r = \sum_{\forall (x,y) \in R_r} 1, r \in \{0,1\} \quad (8)$$

$$v_r = \frac{1}{M_r} \sum_{\forall (x,y) \in R_r} DIS_i(x,y), r \in \{0,1\} \quad (9)$$

$$D = \sum_{\substack{j \in L_{(x,y)} \\ j \neq center(L_{(x,y)})}} \frac{1}{d_j + 1} \quad (10)$$

Inspired by the FLICM method [15], the neighborhood information of pixels is considered in the objective function. The first term in the objective function (6) is the intraclass variance of pixels in the salient region of  $DIS_i$ , and the second term is the neighborhood information of position  $(x,y)$  in  $DIS_i$ .  $d_j$  in (7) is the spatial distance between neighborhood pixel and the center pixel. The spatial distance between the pixels in the four-neighborhood and the central pixel is 1, and  $\sqrt{2}$  between

the pixel in the diagonal position and the central pixel.  $1/(d_j + 1)$  denotes the influence weight determined by the spatial distance between the neighborhood pixel and the central pixel. The bigger the distance is, the smaller the influence weight is, and  $1/D$  is the normalization of the weight. The degree of membership of the  $j$ -th pixel of neighborhood  $L_{(x,y)}$  belonging to the  $r$ -th cluster is denoted as  $u_{jr}$  in (7), and calculated using formula (11).

$$u_{jr} = 1 - \frac{((DIS_{i,j}(x,y) - v_r)^2)}{((DIS_{i,j}(x,y) - v_1)^2 + ((DIS_{i,j}(x,y) - v_2)^2)} \quad (11)$$

#### Calculation of mutation probabilities based on neighborhood information.

In the literature, most of genetic algorithms use the global mutation probability which operates on randomly selected gene with the same probability. Some methods have been developed to automatically change the mutation probability. The mutation probability may be adaptive, and be decreased along the evolution [33], using simulated annealing [34]. To obtain the change map from the difference image, the universal global mutation probability will slow down the algorithm convergence rate when the number of undetermined pixels increases, thus is not applicable to the problem concerned.

Considering the local consistency between the neighborhood and the central pixel in the difference image, a novel mutation operator is proposed to use the neighborhood information in  $DIS$  as heuristic information and automatically calculate a suitable mutation probability for each undetermined pixel. Denote the label of each undetermined pixel  $(x,y)$  in individual  $CM_h$  as  $CM_h(x,y)$ , the mutation probability  $P_m(x,y)$  to  $CM_h(x,y)$  is calculated according to neighborhood information of  $(x,y)$  in  $DIS$  with the following steps.

Step 1: Map the pixel  $(x,y)$  of  $DIS_i$  to the corresponding position in  $DIS$  and consider its  $3 \times 3$  neighborhood  $L_{(x,y)}$  in  $DIS$ . For  $j \in L_{(x,y)}$ ,  $j = 1, 2, 3, \dots, 9$ , the membership degree  $u_{j0}$  and  $u_{j1}$  of the neighboring pixels belonging to the individual cluster centers  $v_0$  and  $v_1$  (obtained according to (9)) are calculated using (12), respectively.

$$u_{jr} = \frac{1}{\sum_{k=0}^1 \frac{(L_j - v_k)^2}{(L_j - v_k)^2}}, r \in \{0,1\} \quad (12)$$

Step 2: Using the degree  $u_{jr}$  of the membership of neighborhood  $L_{(x,y)}$ , the category  $C_j$  of each neighborhood pixel is determined by (13).

$$C_j = \begin{cases} 1 & u_{j1} > u_{j0} \\ 0 & u_{j0} > u_{j1} \end{cases} \quad (13)$$

Step 3: Calculate the mutation probability  $P_m(x,y)$  by using (14).

$$P_m(x,y) = p \times Z^\lambda \quad (14)$$

$$Z = CM_h(x,y) \times \sum_{\substack{j \in L_{(x,y)} \\ j \neq center(L_{(x,y)})}} \frac{1}{d_j} - \sum_{\substack{j \in L_{(x,y)} \\ j \neq center(L_{(x,y)})}} \frac{1}{d_j} C_j \quad (15)$$

where  $p = 0.00001 \sim 0.001$  is the basic mutation probability (i.e. global mutation probability in most existing algorithms). The parameter  $\lambda$  is the exponent of  $Z$ , indicating the degree of influence from the neighborhood information to the mutation

probability of the central pixel. If  $\lambda = 0$ ,  $P_m(x, y)$  will become a global mutation probability.  $d_j$  is the spatial distance between the neighborhood pixel and the center point. The spatial distance between the pixel in the four neighborhoods and the center point is 1, and the space distance between the diagonal pixel and the center point is  $\sqrt{2}$ . Using the neighborhood information of the difference map  $DIS$ ,  $Z$  can be determined as  $|Z| \in [0, 4 + 2\sqrt{2}]$ .

By using the above mechanism, the mutation probabilities of undetermined pixels change adaptively according to the category of neighboring pixels  $L_{(x,y)}$ . When the category  $CM_h(x, y)$  is consistent with most categories of neighborhood  $L_{(x,y)}$ , the probability of mutation is small thus to maintain the original regional consistency. Otherwise, the mutation probability increases, so  $CM_h(x, y)$  tends to become consistent with most of the categories of neighborhood  $L_{(x,y)}$ . This can significantly speed up the convergence of the algorithm.

For each sub-block  $DIS_i$ , the steps of accelerated GA optimization are presented in Algorithm 1.

---

#### Algorithm 1: Procedure of the accelerated GA

---

**input:**  $DIS_i$ , location index of undetermined pixel  $Index_x$ , basic mutation probability  $p$ , crossover probability  $P_c$ , number of individuals  $H$ , number of maximum generation  $G$

**output:** sub-block of change map  $CM_i$

---

#### Begin

1. Generation of the initial population with  $H$  individuals, each encoded as a two-dimension binary matrix. The size of the individual is  $\frac{(M+m) \times (N+n)}{K}$  (consistent with the size of sub-block  $DIS_i$ ). Set the counter of generation  $g = 1$ .
2. The fitness values of the individuals are calculated using (5) and (6).
3. The parent individuals are selected by tournament selection, with the elite retention strategy.
4. The uniform crossover is applied to the selected individuals.
5. The mutation probability  $P_m(x, y)$  of the undetermined pixels in the individual is computed using (14) and (15), and then the mutation operation is performed.
6. Set  $g = g + 1$ . If  $g > G$ , then the algorithm is terminated, and the best individual of the current generation is output as the final change map. Otherwise, go to Step 2.

#### End

---

### 3) Decision fusion of sub-block results based on neighborhood information

For each sub-block  $DIS_i$  of the difference image, the accelerated GA is executed in parallel on the undetermined pixels to obtain the sub-block change map set  $\{CM_i\}$ . The category labels of the pixels in each sub-block, i.e. decision results of the pixels in each sub-block, are then remapped to the corresponding pixel position in  $DIS$  to obtain  $CM$ . The size of  $CM$  is the same as  $DIS$ , i.e.  $M \times N$ .

In the remote sensing image, the pixel values in the neighborhood of a region have certain local consistency, based on which the decision results of all the sub-blocks can be fused as follows.

After  $CM$  is synthesized by  $\{CM_i\}$ , the local neighborhood information of each pixel is used to determine the final category label of the pixel. The pixel  $CM^*(x, y)$  of the final change map

$CM^*$  at position  $(x, y)$  can be obtained by using formula (16):

$$CM^*(x, y) = \begin{cases} 1, & \text{if } n_1 > n_0 \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

where  $n_1$  and  $n_0$  represent the total number of pixels labeled as 1 and 0 in the local neighborhood centered at  $CM(x, y)$ , respectively. If  $n_1 > n_0$ , it means that most of the pixels in the local neighborhood have a category label of 1, so  $CM^*(x, y) = 1$ ; otherwise  $CM^*(x, y) = 0$ . Figure 5 shows an instance of decision fusion based on neighborhood information. There are six pixels labelled 1 around the central pixels, but the label of the central pixel is 0, meaning the central pixel may be classified wrongly or is just a noisy pixel. According to the principle of local consistency, the final category label of the central pixel is changed to 1.

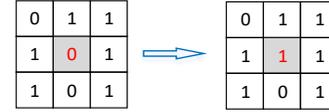


Fig. 5. An instance of decision fusion based on neighborhood information

The procedure described above is called decision fusion of sub-block results based on neighborhood information. The final decision result of the central pixel is determined together based on all the decision results of pixels around it. This procedure can correct the potential wrongly classified result or remove the potential noise in the change map  $CM$ , leading to the final change map  $CM^*$  with a higher detection accuracy.

During the decision fusion of sub-block results, the size of the local neighborhood  $L_{(x,y)}$  is an important parameter that affects the detection accuracy of the final change map. If the size of  $L_{(x,y)}$  is set too large, the referenced local neighborhood information is more abundant, and the noise suppression effect is stronger, however not effective on keeping the detail of the image. If the size of  $L_{(x,y)}$  is too small, decision results from multiple sub-block cannot be fully utilized, and the noise suppression is not sufficient. The size  $3 \times 3$  of local neighborhood is used in this paper based on preliminary experiments.

## III. DATASETS AND EXPERIMENTAL SETTING

### A. The Datasets

Table 3. Remote Sensing Image Datasets

Dataset	Data type	Size	Location	Cause
Bern	SAR image	301×301	Bern, Switzerland	Flood
Ottawa	SAR image	350×290	Ottawa	Flood
Farmland 1	SAR image	291×306	Yellow River estuary	Farming
Farmland 2	SAR image	289×257	Yellow River estuary	Farming
Shimen Reservoir	Panchromatic image	252×349	Taiwan, China	Typhoon

Five datasets of remote sensing images shown in Table 3 are used to verify the accuracy and effectiveness of the proposed SD-aGA algorithm. These include Bern (shown in Figure 2), Ottawa, farmland 1, farmland 2, and Shimen reservoir. Figures

6 - 9 present the remote sensing images at the first time point, the second time point, and the corresponding reference image, respectively.

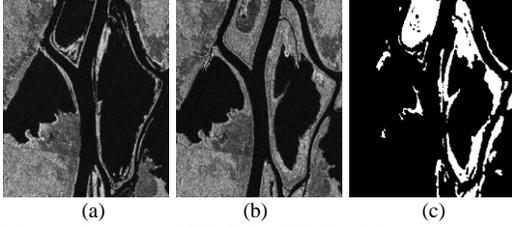


Fig. 6. Ottawa dataset, (a) 1997.5, (b) 1997.8, (c) the reference image

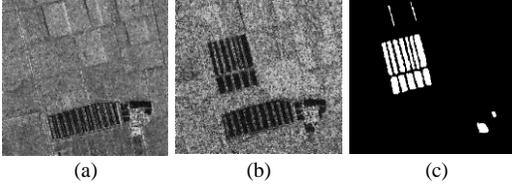


Fig. 7. Farmland 1 dataset, (a) 2008.6, (b) 2009.6, (c) the reference image

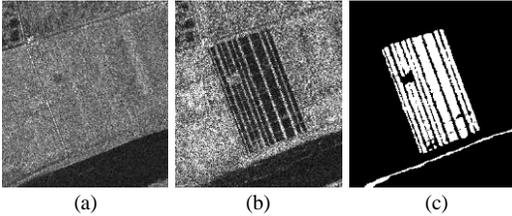


Fig. 8. Farmland 2 dataset, (a) 2008.6, (b) 2009.6, (c) the reference image

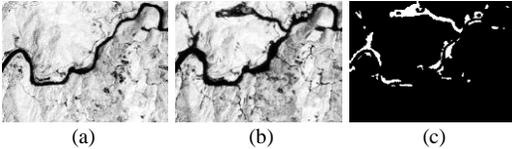


Fig. 9. Shimen Reservoir dataset, (a) 2004.8, (b) 2004.9, (c) the reference image

Table 4. Parameter settings for BM3D Denoising

Dataset	Parameter $\sigma$ of BM3D	
	Remote sensing image 1	Remote sensing image 2
Bern	25	25
Ottawa	25	25
Farmland 1	20	50
Farmland 2	20	50
Shimen Reservoir	30	30

The noise intensities of each of these five datasets are different. Table 4 lists the parameter settings of the BM3D used to denoise in pre-processing these five datasets.  $\sigma$  is the standard deviation of the noise, which is a parameter used in BM3D [28], and the other parameter settings of the BM3D are the same as those in [28].

Results are evaluated using both qualitative and quantitative analysis between the change map obtained by the change detection algorithm and the reference image. Qualitative assessment is subjectively evaluated by visual observation manually. Quantitative analysis adopts some criteria used in the literature [35], and the commonly used five evaluation measures include False Positive ( $FP$ , the number of pixels that

are mistakenly labeled as changed pixels in the detected change map), False Negative ( $FN$ , the number of pixels that are falsely classified as unchanged pixels in the detected change map), Total Error ( $OE$ , the sum of  $FP$  and  $FN$ ), Percentage Correct Classification ( $PCC$ , the ratio between the number of correctly classified pixels and the total number of pixels) and  $Kappa$  coefficient.  $Kappa$  coefficient shows the degree of similarity between the change map obtained by the change detection method and the reference image. The closer to 1 the  $Kappa$  is, the more similar the obtained change map is to the reference image.

### B. Existing methods and parameter settings

To verify the effectiveness of the proposed method, SD-aGA is compared with six existing methods including FCM, FLICM [15], MRFFCM [16], S-FCM (cluster salient regions using FCM after saliency detection), simple GA [23], and S-aGA [27]. FCM, FLICM, MRFFCM and S-FCM are all fuzzy clustering algorithms. FLICM, MRFFCM are the variants of FCM and they consider the neighborhood information during clustering. S-aGA and SD-aGA are the variants of simple GA. The parameter settings of FLICM, MRFFCM, simple GA and S-aGA are the same as those reported in the original literature of their best performance. The parameter settings of simple GA, S-aGA and SD-aGA are provided in Table 5.  $H$  is the population size,  $P_c$  is the crossover probability,  $P_m$  is the mutation probability,  $p$  is the basic mutation probability,  $\lambda$  is the exponent of  $Z$ , and  $K \times K$  is the number of decomposed sub-blocks.

Table 5. The parameter settings of simple GA, S-aGA and SD-aGA

Method	$H$	$P_c$	$P_m$	$p$	$\lambda$	$K \times K$
Simple GA	20	0.8	0.01	-	-	-
S-aGA	20	0.8	-	0.0001	2	4
SD-aGA	20	0.8	-	0.0001	2	4

### C. Experimental environment

All experiments were run on a personal computer with Intel(R) Core(TM) i7-4720HQ CPU, and 8GB RAM. Matlab is used to implement functions and evaluations needed in SD-aGA.

The proposed SD-aGA runs independently on each sub-block obtained by the decomposition of the difference image with no information exchange during the optimization. Serial programming is used to simulate the parallel execution of SD-aGA, and the converted time of parallel execution is used as the computational time for the proposed algorithm.

## IV. PARAMETRIC ANALYSIS AND EXPERIMENTAL RESULTS

### A. Parametric Analysis

#### 1) Number of Decomposed Sub-blocks

This subsection analyzes the effect of the number of decomposed sub-blocks on algorithm performance.

The number of sub-blocks is denoted as  $K \times K$ . When  $K$  is 1, 2, 3, 4, and 5, the number of sub-blocks is 1, 4, 9, 16 and 25, respectively. To analyze the effect of  $K$ , the crossover probability  $P_c$  is set to 0.8, the basic mutation probability  $p =$

0.0001, the exponent  $\lambda = 2$ , and the number of individuals in the population  $H = 20$  in SD-aGA. In SD-aGA, optimization of the accelerated GA stops if the objective function value of the best individual on each generation does not improve for 50 consecutive generations. Figure 10 shows the evaluation of  $Kappa$  on the five datasets when the number of sub-blocks increases. Figure 11 shows the number of generations for each sub-block for the five datasets on sub-block decomposition size  $K \times K$ .

It can be seen from Figure 10, that along with an increasing number of the sub-blocks, the  $Kappa$  values of the detection results slightly increase in the Bern, Ottawa, farmland 1, and farmland 2 datasets, and then tend to be stable at larger values of  $K$ . In the Shimen Reservoir dataset, the  $Kappa$  coefficient remains stable. Figure 11 shows, by using decomposition, as the number of the sub-block increases, the number of generations required for convergence decreases rapidly and then flattens.

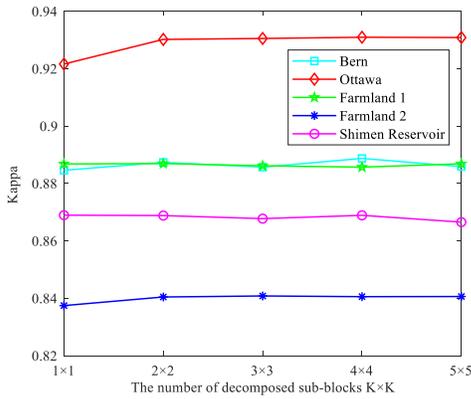


Fig. 10. With the increasing number of decomposed sub-blocks, evaluation of  $Kappa$  on the five datasets

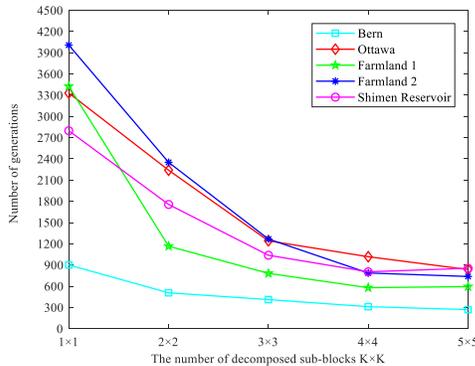
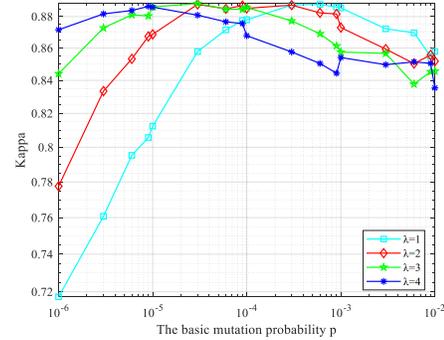


Fig. 11. With the increasing number of sub-blocks  $K \times K$ , the number of generations for each sub-block

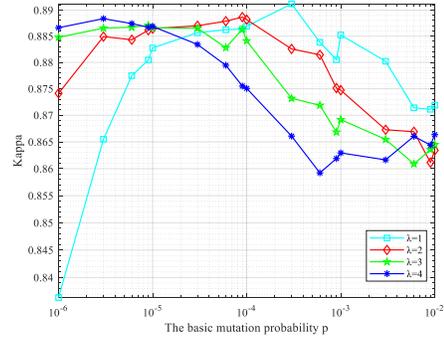
This shows that the decomposition method reduces the search space of the difference image, thereby reducing the difficulty of solving the original problem (according to Figure 11), and increasing the detection accuracy slightly (according to Figure 10). However, it does not mean that the more sub-blocks, the better the detection effect will be. GA is a stochastic global optimization method. The ability of search

from a single generation of the population is limited. The evolutionary efficiency of GA is not proportional to the simplification of the problem. Therefore, we set the parameter of sub-block decomposition as  $K = 4$ , which means that  $DIS$  is globally decomposed into 16 sub-blocks.

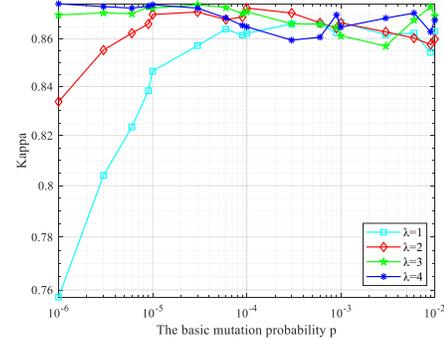
## 2) Adaptive mutation probability operator



(a) Bern



(b) Farmland 1



(c) Shimen Reservoir

Fig. 12. Analysis of parameter  $p$  and  $\lambda$  in adaptive mutation probability operator

In SD-aGA, the neighborhood information of undetermined pixels in  $DIS$  is employed as heuristic information to guide the automatic setting of the mutation probability using formula (14), thereby accelerating the convergence of GA. In formula (14), the basic mutation probability  $p$  and the exponent  $\lambda$  are used to control the range of the mutation probability. When the values of  $p$  and  $\lambda$  are set larger, the maximum value of  $P_m(x, y)$  is higher and thus encourages exploration in GA, and vice versa.

To analyze appropriate values for  $p$  and  $\lambda$  and effect of adaptive mutation operator, we fix the number of sub-blocks  $K \times K = 1 \times 1$ , the crossover probability  $P_c = 0.8$ , and the number of individual in the population  $H = 20$ . In the adaptive mutation

probability operator, let  $p \in [1 \times 10^{-6}, 1 \times 10^{-2}]$ ,  $\lambda \in \{1, 2, 3, 4\}$ . For three datasets (Bern, Farmland1 and Shimen Reservoir), a fixed number of evolutions was set to analyze the detection effect of SD-aGA under different mutation probability parameter settings. The number of generations for the three datasets is set as follows: 1000 for Bern, 4,000 for farmland 1, and 3,000 for Shimen Reservoir.

Figure 12 shows the detection results of  $Kappa$  on the five datasets from SD-aGA with different values for  $p$  and  $\lambda$ . From the experimental results, it can be seen that as  $\lambda$  increases, the effect of the neighborhood information is greater, and the best basic probability value gradually changes from high to low. For a fixed  $\lambda$ ,  $Kappa$  increases first and then decreases when the basic mutation probability  $p$  increases, due to the fact that if  $p$  is too small, evolutionary optimization will be easily stuck into a local optimum; while if  $p$  is too large, the good patterns may be broken and the optimization may miss the global optimal solution. From the results, a reasonable basic mutation probability range is  $p \in [1 \times 10^{-5}, 1 \times 10^{-3}]$ , and the exponent is  $\lambda \in \{2, 3\}$ . Considering the degree of influence from the neighborhood information to the mutation probability, we set  $p = 0.0001$ ,  $\lambda = 2$ .

### B. Experimental Results

The proposed SD-aGA algorithm are compared against six existing methods on the five real remote sensing image datasets. The accuracy of the detection and improvement of speed of SD-aGA is evaluated. In the experiments, the parameter of SD-aGA is set as  $K = 4$ ,  $P_c = 0.8$ ,  $p = 0.0001$ ,  $\lambda = 2$ ,  $H = 20$ , and the optimization of three GA based algorithms (i.e., Simple GA, S-aGA and SD-aGA) stops if the objective function value of the optimal individual on each generation does not improve for 50 consecutive generations.

#### 1) The Bern dataset

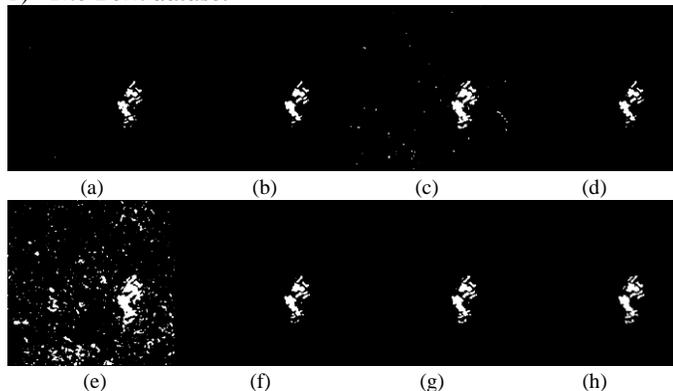


Fig. 13. Change detection results on the Bern dataset. (a) FCM, (b) FLICM, (c) MRFFCM, (d) S-FCM, (e) simple GA, (f) S-aGA, (g) SD-aGA, (h) the reference image

Table. 6. Evaluation results of seven different change detection methods on the Bern dataset

Criteria Method	OE	FN	FP	PCC	Kappa	Time(s)
FCM	357	315	42	0.9961	0.8228	0.62
FLICM	329	293	36	0.9964	0.8379	4.22
MRFFCM	344	63	281	0.9962	0.8337	1.98
S-FCM	347	306	41	0.9962	0.8284	32.29
Simple GA	4430	20	4410	0.9511	0.3246	26515.43
S-aGA	279	189	90	0.9969	0.8723	356.42

SD-aGA	251	138	113	0.9972	0.8888	158.27
--------	-----	-----	-----	--------	--------	--------

Figure 13 (a)-(g) and Table 6 present the change detection map and evaluation result of the seven methods on the Bern dataset, and Figure 13(h) is the reference image. Compared with S-aGA and simple GA, SD-aGA shows a significant improvement in the efficiency of detection, i.e.  $OE = 251$ , and the  $Kappa$  coefficient is 0.8888. The computational time is also reduced to 158.27 seconds. FCM and S-FCM use only the information of gray value, thus the robustness of the algorithms is poor, and  $FN$  is high. Compared with FCM, S-FCM uses saliency detection model to analyze only those salient areas to avoid the influence from those pixels in non-salient areas, and thus the accuracy of detection is improved. The FLICM and MRFFCM methods consider the local neighborhood, which is also beneficial to improve the accuracy of detection. On the whole, SD-aGA achieves the best result in terms of  $OE$ ,  $PCC$  and  $Kappa$ .

#### 2) The Ottawa dataset

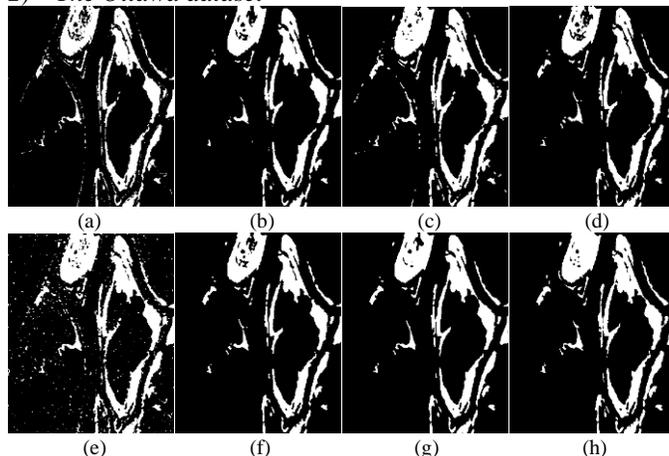


Fig. 14. Change detection results on the Ottawa dataset. (a) FCM, (b) FLICM, (c) MRFFCM, (d) S-FCM, (e) simple GA, (f) S-aGA, (g) SD-aGA, (h) the reference image

Table. 7. Evaluation results of seven different change detection methods on the Ottawa dataset

Criteria Method	OE	FN	FP	PCC	Kappa	Time(s)
FCM	2882	2019	863	0.9716	0.8901	0.62
FLICM	2239	2075	164	0.9779	0.9129	4.10
MRFFCM	2288	770	1518	0.9775	0.9169	1.74
S-FCM	2579	2359	220	0.9746	0.8991	34.60
Simple GA	4008	757	3251	0.9605	0.8605	38764.97
S-aGA	2096	1720	376	0.9793	0.9197	2914.99
SD-aGA	1826	1333	493	0.9820	0.9310	388.26

Figure 14 (a)-(g) and Table 7 present the change maps and evaluation results for the Ottawa dataset using the seven change detection methods, respectively. The Ottawa dataset contains more change areas as shown in Figure 14 (h). Therefore, the running time of the detection method based on evolutionary optimization has increased. SD-aGA requires 388.26 seconds, which is much shorter compared with the simple GA and S-aGA. SD-aGA uses BM3D denoising to improve the quality of the difference image generated, and the fusion of the optimization results from multiple sub-blocks also improves the noise immunity of the detection. The detection result,

$Kappa = 0.9310$ , is the best compared with the other six detection methods. SD-aGA also obtains the best  $OE$  and  $PCC$  among the seven algorithms.

### 3) The Farmland 1 dataset

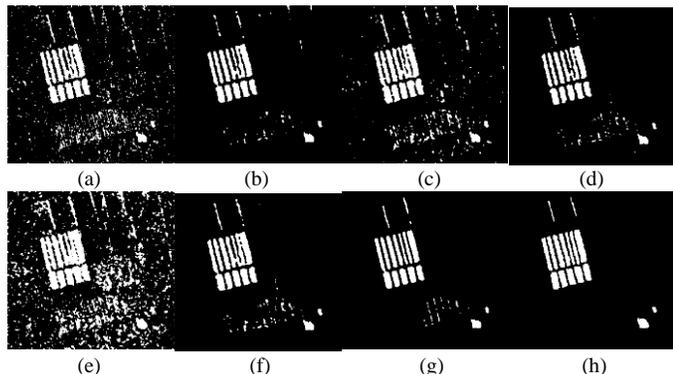


Fig. 15. Change detection results on the Farmland 1 dataset, (a) FCM, (b) FLICM, (c) MRFFCM, (d) S-FCM, (e) simple GA, (f) S-aGA, (g) SD-aGA, (h) the reference image

From the results of each detection method in Figure 15 (a)-(g) and the evaluation results in Table 8, it can be seen that the background of the farmland 1 dataset can be mistakenly detected as changed category easily. The FCM, FLICM, MRFFCM, and simple GA methods all directly performed on the entire difference image. The background noise of the difference image interferes with the detection, leading to a higher false detection rate in the detection results from these methods. S-FCM uses saliency detection models to obtain salient regions. Non-salient background regions are not considered during the analysis, thus effectively reducing false detection rates. The neighborhood information is employed in FLICM to achieve a certain degree of anti-noise effect, and there is only a small amount of punctate erroneous points in the background of the change map. The S-aGA method uses GA to optimize the salient region, considering the neighborhood information in the optimization objective function. Detection accuracy is thus better than S-FCM and FLICM. SD-aGA uses the BM3D algorithm to enhance the useful information in the difference image and suppress part of the background noise. The strategy of decomposition and fusion reduces the computational cost. It also achieves the highest accuracy of detection in strong noise remote sensing images.

Table 8. Evaluation results of the seven different change detection methods on the Farmland 1 dataset

Criteria Method	$OE$	$FN$	$FP$	$PCC$	$Kappa$	$Time(s)$
FCM	5002	462	4540	0.9438	0.6367	1.02
FLICM	1276	777	499	0.9857	0.8717	4.51
MRFFCM	2926	442	2484	0.9671	0.7558	2.15
S-FCM	1354	862	492	0.9848	0.8627	74.44
Simple GA	12229	140	12089	0.8627	0.4087	27341.77
S-aGA	1302	523	779	0.9854	0.8750	1498.02
SD-aGA	1122	772	350	0.9874	0.8857	171.61

### 4) The Farmland 2 dataset

From Figure 16 (a)-(g) and Table 9, it can be seen that the proposed SD-aGA algorithm again achieves the highest accuracy in the detection. The background noise is effectively suppressed, with a better retention on details. Compared with

other algorithms, the detection accuracy of SD-aGA is significantly improved. Overall, the backgrounds in the results of S-FCM, S-aGA and SD-aGA are cleaner and more consistent than those of FCM, FLICM, MRFFCM and simple GA. This is due to that S-FCM, S-aGA and SD-aGA use the saliency detection model to extract the saliency map, thus effectively prevent the influence of background noise on the analysis of the difference image. As for the three GA based algorithms (simple GA, S-aGA and SD-aGA), SD-aGA can obtain the best detection accuracy with the least time.

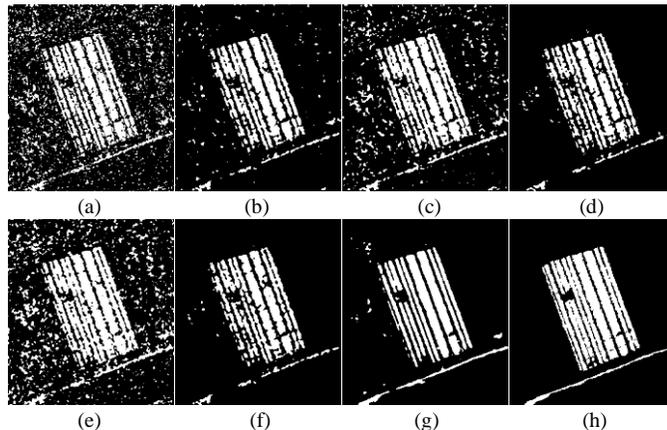


Fig. 16. Change detection results on the Farmland 2 dataset, (a) FCM, (b) FLICM, (c) MRFFCM, (d) S-FCM, (e) simple GA, (f) S-aGA, (g) SD-aGA, (h) the reference image

Table 9. Evaluation results of seven different change detection methods on the Farmland 2 dataset

Criteria Method	$OE$	$FN$	$FP$	$PCC$	$Kappa$	$Time(s)$
FCM	10280	2483	7797	0.8616	0.5952	0.45
FLICM	4460	3105	1355	0.9400	0.7865	3.38
MRFFCM	6860	2298	4562	0.9076	0.7075	1.29
S-FCM	4236	3618	618	0.9430	0.7892	53.60
Simple GA	9902	1672	8230	0.8667	0.6219	22161.74
S-aGA	4085	3577	508	0.9450	0.7962	1830.94
SD-aGA	3324	2567	757	0.9552	0.8406	212.79

### 5) The Shimen Reservoir dataset

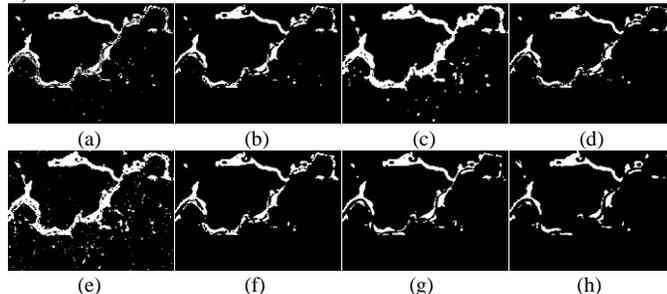


Fig. 17. Change detection results on the Shimen Reservoir dataset, (a) FCM, (b) FLICM, (c) MRFFCM, (d) S-FCM, (e) simple GA, (f) S-aGA, (g) SD-aGA, (h) the reference image

From Figure 17 (a)-(g) and Table 10, it can be observed that S-FCM, S-aGA and SD-aGA obtain results with cleaner backgrounds and FLICM also obtains good result. Overall, the result of the SD-aGA algorithm on the Shimen Reservoir dataset is the most similar to the reference image in Figure 17 (h), and the  $Kappa$  coefficient reaches 0.8689, which is much

larger than the others. SD-aGA also obtains the best *OE* and *PCC* comparing with other algorithms.

Table. 10. Evaluation results of the seven different change detection methods on the Shimen Reservoir dataset

Criteria Method	<i>OE</i>	<i>FN</i>	<i>FP</i>	<i>PCC</i>	<i>Kappa</i>	<i>Time(s)</i>
FCM	2293	389	1904	0.9739	0.7812	0.37
FLICM	1810	524	1286	0.9794	0.8155	4.15
MRFFCM	5351	32	5319	0.9392	0.6133	1.43
S-FCM	1793	801	992	0.9796	0.8073	34.83
Simple GA	5687	51	5636	0.9353	0.5968	29034.26
S-aGA	1969	365	1604	0.9776	0.8076	2207.00
SD-aGA	1289	251	1038	0.9853	0.8689	240.91

Considering the computational time of different algorithms shown in Tables 6 - 10, it can be found that the FCM based algorithms (FCM, FLICM, MRFFCM and S-FCM) run much faster than the GA based algorithms (simple GA, S-aGA and SD-aGA); however, their detection results are generally inferior to those obtained by S-aGA and SD-aGA. Among the three GA based algorithms, SD-aGA obtains the best results with the least time. Generally, the running time of SD-aGA is about one tenth of that of S-aGA and about one hundredth of that of Simple GA. The experimental results demonstrate the effectiveness and efficiency of SD-aGA.

## V. CONCLUSIONS

This paper proposes an accelerated genetic algorithm based on search-space decomposition (SD-aGA) for change detection in remote sensing images. Based on an existing accelerated genetic algorithm with salient map guidance (S-aGA), further improvements have been proposed to promote the noise immunity and speed of the change detection in remote sensing images. The BM3D algorithm has been introduced to enhance useful information to a certain extent and suppress the noise, so that the quality of the generated difference image will be higher. The salient regions in the difference image is decomposed into sub-blocks. The accelerated GA then optimizes the undetermined pixels in each sub-block in parallel. The change map of the sub-block is remapped to the positions in the original difference image, and results of sub-blocks are fused in the local neighborhood by using the majority vote method to acquire the final change map.

Five widely used real remote sensing images datasets are used in experiments to verify and demonstrate the efficiency and effectiveness of the proposed SD-aGA. Compared to the six existing change detection methods, the proposed SD-aGA algorithm fully demonstrates its superior anti-noise performance and accuracy of detection. The search space of SD-aGA can be effectively reduced, thus significantly reduces the computational time.

Based on all the experiments carried out, some further conclusions can be drawn as follows.

(1) The decomposition method can reduce the search space of the difference image, thereby reduce the difficulty of solving the original problem, and thus increase the detection accuracy slightly.

(2) If all the sub-blocks are processed in parallel, the overall

run time decreases as the number of sub-blocks increases.

(3) As the number of sub-blocks increases, the running time of each block decreases first and then flattens, meaning it is not necessary to set a very large number of blocks.

(4) Parameter  $p$  and  $\lambda$  control the maximum value of  $P_m(x, y)$  together, and  $\lambda$  reflects the degree of influence of the neighborhood information. As  $\lambda \in \{2, 3\}$ , the algorithm can obtain good detection results within a wide range of  $p$ , indicating the proposed SD-aGA is not sensitive to parameter settings.

(5) Compared with Simple GA and S-aGA, SD-aGA can achieve a higher detection accuracy with less time, indicating that SD-aGA is more executable and can be applied in the area of change detection in remote sensing images.

In the future, the improved SD-aGA can be further investigated and analyzed with adaptive parameters such as the number of sub-blocks,  $p$  and  $\lambda$ .

## ACKNOWLEDGMENT

The authors would like to thank the School of Computer Science, University of Nottingham for providing the essential research facilities to this joint research.

## REFERENCES

- [1] R. S. Lunetta, C. D. Elvidge, "Remote sensing change detection: environmental monitoring methods and applications. Ann Arbor Mi Ann Arbor, 1999.
- [2] J.-F. Zhang, L.-L. Xie, X.-X. Tao, "Change detection of earthquake-damaged buildings on remote sensing image and its application in seismic disaster assessment", Proc. IGARSS, 2013, vol. 4, pp. 2436-2438.
- [3] R. Salvador, X. Pons, "On the reliability of landsat TM for estimating forest variables by regression techniques: A methodological analysis", IEEE Trans. Geosci. Remote Sensing, vol. 36, no. 6, pp. 1888-1897, Nov. 1998.
- [4] H. Yang, J. Shi, Z. Li, H. Guo, "Temporal and spatial soil moisture change pattern detection in an agricultural area using-temporal RADARSAT ScanSAR data", Int. J. Remote Sens., vol. 27, no. 19, pp. 4199-4212, Oct. 2006.
- [5] L. Bruzzone, D. F. Prieto, "An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images", IEEE Trans. Image Process., vol. 11, no. 4, pp. 452-466, Mar. 2002.
- [6] Y. Bazi, F. Melgani, H. Al-Sharari, "Unsupervised change detection in multispectral remotely sensed imagery with level set methods", IEEE Trans. Geosci. Remote Sens., vol. 48, no. 8, pp. 3178-3187, Aug. 2010.
- [7] M. Gong, Z. Zhou, J. Ma, "Change detection in synthetic aperture radar images based on image fusion and fuzzy clustering", IEEE Trans. Image Process., vol. 21, no. 4, pp. 2141-2151, Apr. 2012.
- [8] L. Bruzzone, D. F. Prieto, "Automatic analysis of the difference image for unsupervised change detection", IEEE Trans. Geosci. Remote Sens., vol. 38, no. 3, pp. 1171-1182, May 2000.
- [9] Y. Bazi, L. Bruzzone, F. Melgani, "An unsupervised approach based on the generalized Gaussian model to automatic change detection in multitemporal SAR images", IEEE Trans. Geosci. Remote Sens., vol. 43, no. 4, pp. 874-887, Apr. 2005.
- [10] G. Moser, S. B. Serpico, "Generalized minimum-error thresholding for unsupervised change detection from SAR amplitude imagery", IEEE Trans. Geosci. Remote Sens., vol. 44, no. 10, pp. 2972-2982, Sep. 2006.

- [11] M.N. Sumaiya, R. Shantha Selva Kumar, "Gabor filter based change detection in SAR images by KI thresholding", *Optik*, no. 130, pp. 114-122, 2017.
- [12] G. Moser, S. B. Serpico, "Unsupervised Change Detection From Multichannel SAR Data by Markovian Data Fusion", *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no.7, pp. 2114-2128, 2009.
- [13] M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag, T. Moriarty, "A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data", *IEEE Trans. Med. Imaging*, vol. 21, no. 3, pp. 193-199, 2002.
- [14] L. Szilagyi, Z. Benyo, S. Szilagyii, H. Adam, "MR brain image segmentation using an enhanced fuzzy C-means algorithm", in *Proc. 25th Annu. Int. Conf. IEEE EMBS*, 2003, pp. 17-21.
- [15] S. Krinidis, V. Chatzis, "A robust fuzzy local information C-means clustering algorithm", *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1328-1337, May 2010.
- [16] M. Gong, L. Su, M. Jia, W. Chen, "Fuzzy clustering with a modified MRF energy function for change detection in synthetic aperture radar images", *IEEE Trans. Fuzzy Syst.*, vol. 22, no. 1, pp. 98-109, Feb. 2014.
- [17] H.-C. Li, T. Celik, N. Longbotham, W. J. Emery, "Gabor feature based unsupervised change detection of multitemporal SAR images based on two-level clustering", *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 12, pp. 2458-2462, Dec. 2015.
- [18] H. Li, M. Gong, Q. Wang, J. Liu, L. Su, "A multiobjective fuzzy clustering method for change detection in SAR images", *Appl. Soft Comput.*, vol. 46, pp. 767-777, Nov. 2015.
- [19] M. Gong, J. Zhao, J. Liu, Q. Miao, L. Jiao, "Change Detection in Synthetic Aperture Radar Images Based on Deep Neural Networks", *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 1, pp. 125-138, Jun. 2015.
- [20] F. Gao, J. Dong, B. Li, Q. Xu, "Automatic change detection in synthetic aperture radar images based on PCANet", *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1792-1796, Oct. 2016.
- [21] R. Shang, Y. Yuan, L. Jiao, Y. Meng, A. M. Ghalamzan, "A Self-paced Learning Algorithm for Change Detection in Synthetic Aperture Radar Images", *Signal Process.*, vol. 142, pp. 375-387, Jan. 2017.
- [22] W. Ren, J. Song, S. Tian, W. Wu, "Survey on unsupervised change detection techniques in SAR images1", in *Signal and Information Processing (ChinaSIP) 2014 IEEE China Summit & International Conference on*, 2014, pp. 143-147.
- [23] T. Celik, "Change detection in satellite images using genetic algorithm", *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 2, pp. 386-390, Apr. 2010.
- [24] B. Bhanu, S. Lee, J. Ming, "Adaptive image segmentation using a genetic algorithm", *IEEE Trans. Syst. Man Cybern.*, vol. 25, no. 12, pp. 1543-1567, Dec. 1995.
- [25] A. Sheta, M. S. Braik, S. Aljahdali, "Genetic Algorithms: A tool for image segmentation", in *IEEE International Conference on Multimedia Computing and Systems*, 2012, pp. 84-90.
- [26] S. Cho, Z. Chi, "Genetic evolution processing of classification", *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 2, pp. 216-231, 2005.
- [27] C. Mu, C. Li, Y. Liu, M. Sun, L. Jiao, R. Qu, "Change detection in SAR images based on the salient map guidance and an accelerated genetic algorithm", in *IEEE Cong. on Evolut. Comput.*, Donostia, Spain, 2017, pp. 1150-1157.
- [28] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, "Image denoising by sparse 3d transform-domain collaborative filtering", *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080-2095, Aug. 2007.
- [29] N. Bruce, J. Tsotsos, "Saliency Based on Information Maximization", in *International conference on Neural Information Processing Systems*, 2006, vol. 18, pp. 155-162.
- [30] R. Achanta, S. Hemami, F. Estrada, S. Susstrunk, "Frequency-Tuned Salient Region Detection", in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2009, pp. 1597-1604.
- [31] S. Goferman, L. Zelnik-Manor, A. Tal, "Context-aware saliency detection", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 10, pp. 1915-1926, Oct. 2012.
- [32] N. Otsu, "A threshold selection method from gray-level histograms", *IEEE Trans. Syst. Man Cybern.*, vol. 9, no. 1, pp. 62-66, Jan. 1979.
- [33] M. Liu, Z. Sun, J. Yan, J. Kang, "An adaptive annealing genetic algorithm for the job-shop planning and scheduling problem", *Expert Syst. Appl.*, vol. 38, no. 8, pp. 9248-9255, 2011.
- [34] D. Sirag, P. Weisser, "Toward a unified thermodynamic genetic operator", in *International Conference on Genetic Algorithms on Genetic Algorithms and Their Application*, L. Erlbaum Associates Inc., 1987, pp. 116-122.
- [35] P. Rosin, E. Ioannidis, "Evaluation of global image thresholding for change detection", *Pattern Recognit. Lett.*, vol. 24, no. 14, pp. 2345-2356, Oct. 2003.