Journal of Real-Time Image Processing (2021) 18:1577–1590
https://doi.org/10.1007/s11554-020-01062-y

SPECIAL ISSUE PAPER

Instant water body variation detection via analysis on remote sensing imagery

Yirui Wu[1] · Pengfei Han[1] · Zhan Zheng[1]

Received: 16 September 2020 / Accepted: 12 December 2020 / Published online: 16 January 2021

Abstract
Water resource is one of the most valuable natural resources for human beings, which requires to be monitored for careful protection. Inspired by a significant power of machine learning methods, researchers have successfully developed many applications to automatically perform identification on the water body via analyzing remote sensing images. Since a similar category of ground objects could show a large difference in spectral representation, researchers try to propose distinctive and effective features to offer redundant information for category classification. Moreover, large amount of high-resolution remote sensing images require analyzing algorithms to be parallel processed for instant feedback. Based on these requirements, we propose a novel water body variation detection via analysis on remote sensing images. Specifically, the proposed method firstly perform pixel-level classification to locate abnormal changes with thoughts of visual word patterns. Afterwards, the proposed method proposes block division method to construct parallel running version with Mapreduce structure. With high representational and parallel running abilities, the proposed method is capable to accurately detect variation areas on remote sensing images with instant feedback. Experiments on several self-collected datasets show the proposed method has achieved better efficiencies than comparative studies.

Keywords Real-time RS applications of detection and estimation theory · Remote sensing images · Water body variation detection method · Distributed processing

1 Introduction

Since there exists an increasing demand for water resource with fast economical development of human society, water ecosystem including rivers, lakes, and seas is of great significance to humans. However, unreasonable arrangements of water result in a huge waste of natural resource and might result in disasters, such as soil erosion and land desertification. To achieve better usage on limited water resource, how to effectively extract useful information from water body becomes a hot topic for researchers. In other words, users require highly distinguish models to automatically perform identification or diagnose on these valuable nature resource.

With a significant development of aerospace and imaging technology [1], both spatial and temporal resolution of remote sensing imagery have been continuously improved. On the basis of abundant information contained in high-resolution images, researchers and engineers have successfully built quantity of applications [2], such as land use analysis, geological map updating, ecological system monitoring and so on.

Based on their experience of utilizing remote sensing technology to solve monitoring problems, we pay special attention to perform water body variation detection via analysis on remote sensing imagery. Essentially, man-made buildings like reservoir, dams and so on lead to great variants in the water body and its surroundings, which could result in pollution or waste on a limited water resource. Waterbody monitoring application via remote sensing imagery thus offers an effective and instant way to accurately locate variations, due to features of rich appearance details and large visual fields.

Researchers have proposed lots of methods to detect water body variation based on information extracted from remote sensing imagery [3, 4]. However, their trials could result in low accuracy and efficiency, due to the following two main challenges:

✉ Yirui Wu
wuyirui@hhu.com

1 College of Computer and Information, Hohai University, Nanjing, China

– Similar category of ground objects could show a large difference in spectral representation, which leads to a great challenge for abnormal detection in surroundings of the water body. In other words, objects in the same category could own different spectral values, while objects in different categories might share similar spectral distributions.
– Large volume and complexity of remote sensing imagery leads to low efficiency of processing algorithm, which is against the initial purpose of instant water body monitoring. In other words, researchers are facing challenges brought by increasing volume and complexity of big remote sensing data.

To solve such problems, we propose a novel detection method based on bag of visual words to analyze variants in high-resolution remote sensing imagery, which helps locates abnormal changes around the lake and reservoir caused by human activities. To solve the first problem, we construct mixed feature space to offer abundant descriptive information on each pixel of input remote sensing images. Furthermore, the bag of visual words are adopted to further abstract neighboring feature information around each pixel, resulting in highly distinctive and effective visual word patterns for classifying ground objects.

Remote sensing applications often involve huge volume data obtained daily by numerous in-orbit satellites. This makes it a perfect area for data-driven applications. High-resolution images are widely used in the monitoring and analysis of a wide range of features due to their high resolution and multi-band image information [5]. The use of high-resolution images for variation detection can more effectively monitor changes in features in the surface area, which is of great help in research and applications. However, due to the imaging characteristics of remote sensing images, variation detection, along with other remote sensing image processing has been facing an inherent problem, namely, the large scale data processing volume. In contrast to the recent use of distributed computing to solve the processing problem of geometrically increasing network information, we consider the use of MapReduce for high-resolution image change detection, and construct a reasonable application scheme based on the characteristics of variation detection processing to improve execution efficiency.

Although many researchers have demonstrated and implemented distributed processing of remote sensing images using platforms such as Hadoop, most of these efforts have focused on the processing of massive images rather than on the processing of a single large image [6], where the processing of each image is independent of each other and cannot maintain inter-connectivity between images at different processing nodes, and cannot satisfy the requirements of a single large image. Some data requirements for distributed processing of large-format images. In addition, distributed processing for remote sensing image variation detection is even less known. Therefore, we propose to perform water body variation detection with distributed processing for real-time performance, where we successfully transform the original version of variation detection algorithm to Map and Reduce tasks following the guidance of MapReduce structure. Such parallel running version of the algorithm greatly improve efficiency to complete variation detection task, especially when processing with a large amount of high-resolution remote sensing images.

The contributions of this paper are twofold:

– We propose a novel water body variation detection method based on input remote sensing images, which utilizes multiple features and bag of visual words to first classify pixel-level labels for ground objects and then locate abnormal changes inside images. Experiments have proved the effectiveness of the proposed method to complete abnormal detection task with instant and accurate responses.
– We transform the proposed method to a parallel running version, which introduces the idea of distributed processing with the proposed block division method under restrictions of MapReduce structure. We believe such block-based idea for MapReduce transformation can be applied in most of the application scenarios to deal with remote sensing images.

The rest of the paper is organized as follows. Section 2 gives an overview of the related work on relative aspects. In Sect. 3, details of the proposed water body variation detection is discussed. Section 4 explains how to transform the original algorithm into a parallel running version under MapReduce structure. Section 5 shows our experimental results with several comparative methods, and finally Sect. 6 concludes the paper.

2 Related work

This section will introduce the relevant research that inspired us to design the proposed method, including Introduction to Processing of remote sensing imagery and bag of visual words model.

2.1 Introduction to processing of remote sensing imagery

In early remote sensing processing applications, classification methods generally use low-level visual features, like shape, texture, color and so on. For example, Luo et al. [7] combine six different low-level visual features to form a multi-feature representation for accurate classification on remote sensing imageries. However, classifiers built on these low-level visual features are not powerful enough to deal with complex remote sensing imagery. Researchers thus propose to apply distinguish visual features for classification, such as scale-invariant feature transform (SIFT), local binary pattern (LBP) and other feature representations. For example, Xu et al. [8] use the modified normalized difference water index (MNDWI) to extract water surface information for further study.

Meanwhile, the quantity of classical machine learning methods are adopted to deal with the complexity of remote sensing imagery. For example, He et al. [9] use support vector machine (SVM) as a classifier to segment semantical parts in remote sensing imagery. Comparing with classifiers built on maximum likelihood estimation (MLE) and neural network, they argue that SVM is more efficient in such task, due to its ability in mathematic modeling on high-dimensional feature space. Later on, Gokhan et al. [10] focus on methods to solve land cover problem, where they employ a number of manifold learning algorithms to perform feature extraction and conduct extensive comparative experiments on a dataset with 200-band hyper-spectral images. They finally get a conclusion that SVM is the best classifier for dealing with remote sensing imagery.

Recently, deep learning methods have emerged and been widely applied [11], due to its significant power to extract distinctive and informative features from images. In addition, deep learning methods are capable of abstract high-level semantic information from input images, which avoids to extract semantic meanings by designing artificial features. Based on these advantages, many researchers are devoted to design appropriate neural networks for analysis on remote sensing imagery.

Early, Mnih et al. [12] use patch-based CNN classifiers to perform segmentation experiments on multi-spectral images to locate targets like road and house. Later, Souleyman et al. [13] firstly use VGGNet to extract deep features from remote sensing imagery, and then use discriminant correlation analysis (DCA) to transform characteristics of different feature layers, where they construct a novel classifier by fusing the transformed features for accurate recognition results. Afterwards, Sakrapee et al. [14] train classifiers by fusing manual and CNN extracted features. Afterwards, they utilized conditional random field (CRF) [15] to optimize graph to perform classification tasks, where they achieve remarkable results on the ISPRS 2D semantic annotation dataset. Following ideas of involving strength from both graph-based learning and deep learning, Zhao et al. [16] use combination of CNN and CRF to perform over-segmentation at first for remote sensing imagery and then merge parts with the guidance of semantic meanings.

Recently, Zhao et al. [17] utilize their proposed PSPNet to extract global context information through information aggregation from different regions, which successfully achieve high accuracy on benchmark datasets such as PASCAL VOC 2012 and Cityscapes. For the goal of fusing multi-layer features, Sun et al. [18] introduce maximum fusion strategy for Fully-Connected Network, which effectively fuses semantic information extracted from deep layers and detailed information from shallow layers. To apply context information for higher accuracy, Wang et al. [19] propose an end-to-end attention recursive convolutional network (ARCNet) for scene classification, which construct attention mechanism to focus on important locations for improvement on classification performance. Sun et al. [20] propose a new change detection method based on similarity measurement between heterogeneous images. The method constructs a graph for each patch based on the nonlocal patch similarity to establish a connection between heterogeneous data, and then measures the change level by measuring how much the graph structure of one image still conforms to that of the other image. Experiments demonstrate the effective performance of the proposed nonlocal patch similarity-based heterogeneous change detection method. Aiming at the real-time detection of multiple objects, a cascaded convolutional neural network real-time object-detection framework for remote sensing images is proposed by Hua et al. [21], which integrates visual perception and convolutional memory network reasoning. The experimental results show that the proposed algorithm significantly improves the efficiency of object detection while ensuring detection accuracy and has high adaptability.

Paying attention to the design of loss function, Cheng et al. [22] optimize a new discriminative objective function by combining metric learning loss and cross-entropy loss, which results in resemble feature description on similar scenes. To better classify remote sensing scenes, Zhang et al. [23] propose a simple but efficient DenseNet on the basis of full convolution network, which generates a large number of reusable feature maps through dense connections and less parameters. They prove that their proposed model could significantly improve classification performance on several public datasets like UCM, AID, OPTIMAL-31 and NWPU-RESISC45.

In fact, largely increasing volume of remote sensing image data requires researchers to design high efficiency and effective algorithms. In this work, we focus on designing a method to detect variations in the watershed by extracting and classifying distinctive information from remote sensing images.

2.2 Bag of visual words model

Bag of words model originates from applications for text classification [24], where texts are regarded as a set of words without order. Afterwards, it is widely used in domains of image processing and computer vision [25], where images are considered as a set of representations of local regions, being independent of their locations. In other words, features extracted from local regions for images are equivalent to words for text. Therefore, the most important feature for a bag of visual words model lies in the fact that it ignores location information extracted from input images.

With such property, distribution of extracted features for local areas could be used to classify images, which is especially helpful to deal with large-scale image such as classification on high-resolution remote sensing imagery, due to rich features extracted from ground object and less complexity in involving additional location information. Wang et al. [26] propose a visual attention based bag-of-words (VABOW) model for image classification task. This VABOW model combines shape, color and texture cues and uses L1 regularization logistic regression method to select the most relevant and most efficient features, and results show good performance on image classification tasks. Karakasis et al. [27] present an image retrieval framework that uses affine image moment invariants as descriptors of local image areas. Detailed feature vectors are generated by feeding the produced moments into a bag-of-visual-words representation. The results are promising compared with other widely used local descriptors. The improvement of high-resolution satellite images makes the spectrum and texture more rich and complex, which poses challenges for the automatic classification. Jinying et al. [28] combines active learning and bag of word model for image classifications. To verify the effectiveness and robustness, the high-resolution image in Shandong province was used as experimental data. The results show that the proposed method can effectively classify the study area into four types: water, ground, vegetation, and building, with the overall accuracy of over 90.6%.

The general process for a bag of visual words model could be concluded as follows [29]:

1. Extract local visual features under a settles scale from images in the training set.
2. Construct a visual dictionary by first clustering local visual features and then regarding each category as a visual word.
3. Match local visual features with visual words in each training image. Afterwards, compute frequency of each word to form a histogram, which is regarded as feature vectors to train classifier.
4. Build classifier by appropriate machine learning methods, which regards feature vectors collected from the former step as input.
5. Extract local visual features for an image in testing set and obtain the corresponding classification results by the constructed classifier.

Applying a bag of visual words model to process remote sensing imagery could help solve the classification problem [30], i.e., similar category of ground objects could show a large difference in spectral representation, where it adopts quantity of and multiple categories of features for accurately recognizing local regions. In fact, such phenomenon is similar to the difficulties brought by synonyms and polysemy for text analysis [31]. Inspired by a local binary pattern (LBP), the proposed method applies visual vocabulary on classification in pixel-level, where visual word patterns for pixel neighborhood are regarded as distinctive features in pixel-level. In summary, we have fused multiple features and a bag of visual words model together to achieve pixel-level classification of ground objects, which could better solve the problem of spectral differences in remote sensing images. Experiments on several datasets have proved the effectiveness of adopting fused features and bag of words for processing, compared with current methods.

3 Algorithm for water body variation detection

Since the core idea for water body variation detection is to first perform classification and then detection, we construct workflow of the proposed method with two stages as shown in Fig. 1. During the classification stage, we firstly extract multiple features from remote sensing images collected by GF-1 satellite, where we utilize these features to form 4B3I1T mixed feature space. In 4B3I1T feature space, we cluster feature vectors into a certain number of visual words. Afterwards, we extract visual word patterns for each pixel from input images. Finally, we use SVM classifier to perform supervised pixel-level classification, which computes category labels for each pixel.

During the detection stage, we use CVA algorithm to detect water body variations based on the pixel-level classified images. Afterwards, we adopt appropriate post-processing to eliminate misclassification results and obtain more accurate detection results.

3.1 4B3I1T mixed feature space

How to fuse multiple features to construct a task-specified feature space is a key and hot issue in detection methods [32]. Due to special requirements in water body variation

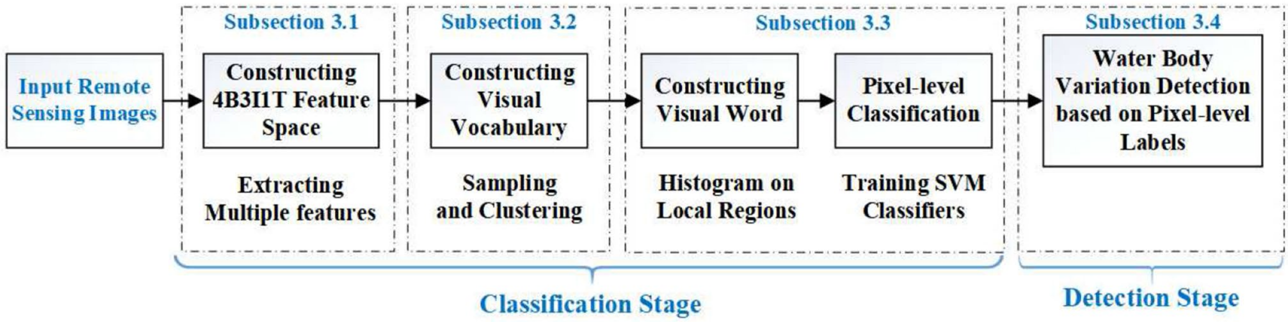


Fig. 1 Workflow of the proposed water body variation detection algorithm, which consists of two stages, i.e., classification and detection

detection, it's difficult to apply a common-used feature space for accurate detection. We thus design 4B3I1T mixed feature space for effective water body variation detection.

Among multiple categories of features extracted from remote sensing images, features such as color, spectral, and textural, are usually adopted to form feature space. Specifically, color information plays an important role in detecting surface objects. However, the effect of color feature channel in classifying objects could be largely affected by imaging conditions and resolution of ground objects. Meanwhile, spectral features have the same advantage as color features, i.e., easy to access, and could avoid the influence of imaging conditions, which are the main factors to adopt it as one feature channel. However, it could bring unstable performance in classification, since information carried by spectral feature could be dominated by several specific pixels. Texture features are more diverse by making full use of local intensity information of pixels. However, their calculation is too complex to bring difficulty for instant calculation. Besides, several useful index is usually adopted by researchers, which could help analyze spectral characteristics to emphasize on features of targets and suppress features of background objects.

Based on the above discussions, we use four bands of spectral features, three categories of index features and one LBP texture feature to form 4B3I1T mixed feature space (shorted for 4 bands, 3 indexes and 1 texture feature). Considering the uneven value range of the above features, we firstly re-scale feature values into a similar order of magnitude, and then fuse them to obtain the mixed feature space, each pixel corresponds to an eight-dimensional feature vector and the vector can describe the pixel as a point in the hybrid feature space. We offer descriptions on these feature channels as follows:

- Spectral feature channels: The post-processed image obtained from GF-1 still retains original characteristics of four spectral bands, namely blue band, green band, red band and near-infrared band.
- Normalized differential water idex (NDWI): NDWI is widely-used to locate water area in remote sensing images, which is built based on information of the green band and near-infrared band. NDWI is capable to suppress vegetation information and highlight waterbody information. Moreover, it applies normalization to avoid difference in calculation results caused by multiple data resources.
- Normalized differential vegetation index (NDVI): NDVI is constructed based on the near-infrared band and red band, which could effectively highlight vegetation information.
- Ratio build-up index (RBI): RBI is constructed with information from all four bands, which could help extract building information, especially under the condition that Gf-1 isn't equipped with Short infrared band.
- Local binary patterns (LBP): We adopt a uniform mode of LBP for feature construction, which has advantages of rotation and intensity invariance, and simpleness in the calculation.

3.2 Visual vocabulary and pixel-level classification

In this subsection, we describe steps to construct visual vocabulary based on 4B3I1T mixed feature space and perform pixel-level classification, which extracts visual word pattern around pixel neighborhood as a pixel-level feature.

Specifically, the corresponding feature for each pixel in the input remote sensing images can be extracted to an eight-dimensional feature vector. After clustering points in feature space into visual vocabularies based on the visual dictionary method, these points can be divided into a number of visual words, which completes the construction of mapping from feature vectors to their belonged visual words. It's noted that mapping could help eliminate influences caused by variations of feature values in a certain range. According to the ideal thought, ground objects belongs to the same category should own similar feature values. However, their feature values may change within a certain range in the real-world.

Therefore, feature space belongs to the same ground object category is scattered in a hypersphere with a certain size. Through the mapping from feature vectors to visual vocabulary, points in the same hypersphere can share the same attributes to a large extent, which are defined as their corresponding visual word.

Algorithm 1 Visual Word Label Assigning Algorithm

Input: Image pixel set $I = \{x_i, i = 1, 2, ..., N\}$.
Output: Vision word labels set $W = \{w_i, i = 1, 2, ..., N\}$.

Step 1 For each pixel x_i, achieve the corresponding feature vector v_i according to the mixed feature space 4B3I1T:

$$v_i = \{Blue_i, Green_i, Red_i, NIR_i, NDWI_i \\ NDVI_i, RBI_i, LBP_i, i = 1, 2, ..., N\} \tag{1}$$

Afterwards, define feature vector set $V = \{v_i, i = 1, 2, ..., N\}$.

Step 2 Perform sparse sampling with a certain distance d on V to construct the corresponding subset of feature points $subV$:

$$subV = f_{ss}(V, d), where\ subV = \{\tilde{v}_j, j = 1, 2, ..., M\} \tag{2}$$

where function $f_{ss}()$ represents operation of sparse sampling, $\tilde{v}_j$ refers to representation feature points after sparse sampling, and M is the total number for feature points after sampling.

Step 3 Use K-means method to cluster feature vectors in $subV$ into K different classes, where each cluster center corresponds to a visual word c_i. Such process is defined as

$$c_k = f_c(subV), where\ k = 1, 2, ..., K \tag{3}$$

where function $f_c()$ represents operation of clustering.

Step 4 For each v_i, calculate distance to each visual word and assign the nearest visual word, i.e., its corresponding visual word label w_i.

$$w_i = f_w(\arg\min_k(f_E(v_i, c_k))), \\ where\ i = 1, 2, ..., N, k = 1, 2, ..., K \tag{4}$$

where functions $f_w()$ and $f_E()$ represents operation of assigning visual words based on label k, and calculation of Euclidean distance, respectively.

We firstly construct a visual dictionary through clustering and then assign visual word label to each pixel, which is illustrated in Algorithm. 1. It's noted that we should pay special attention on the number of categories *K*, which could bring redundancy by defining a relatively large value. Due to the reason that feature vectors of neighboring pixels are similar, we perform sparse sampling to define sample points instead, which can be regarded as a representation for all neighbour pixels. In this way, we not only decrease complexity of constructing visual vocabulary but also get a small subset of feature vectors *subV*.

After performing Algorithm. 1, visual word label could be assigned to each pixel, thus obtaining visual word image corresponding to each input remote sensing image. We show samples of original image and its corresponding visual word image in Fig. 2. It's noted that we perform later processing on visual word image instead.

Algorithm 2 Pixel-level Classification algorithm based on Visual Vocabulary

Input: Vision word labels set $W = \{w_i, i = 1, 2, ..., N\}$, Predicted classification label set $L = l_i, i = 1, 2, ..., N$.
Output: Predicted classification label set $L = l_i, i = 1, 2, ..., N$

Step 1 We firstly construct the original training sample set $T_1 = \{g_i, i = 1, 2, ..., N\}$ based on input set W, where visual word feature around neighborhood g_i is defined as follows:

$$g_i = \{[w_1, ..., w_j], where\{x_1, ..., x_j\} \in Neigh(x_i, \delta)\} \tag{5}$$

where operator [] refers to feature concatenate operation, function $Neigh(x_i, \delta)$ refers to a window centered on pixel x_i with size δ.

Step 2 For each g_i in T_1, compute frequency of visual words in each dimension, which form visual word pattern for each pixel:

$$h_i = f_h(g_i), i = 1, 2, ..., N \tag{6}$$

where function $f_h()$ represents histogram operation. We thus obtain final training set $T_2 = \{h_i, i = 1, 2...N.\}$

Step 3 Use T_2 to train SVM to achieve the corresponding classifier C:

$$C = f_S(T_2; \Delta), i = 1, 2, ..., N \tag{7}$$

where function $f_S()$ represents training process of SVM classifier and Δ represents the parameter settings.

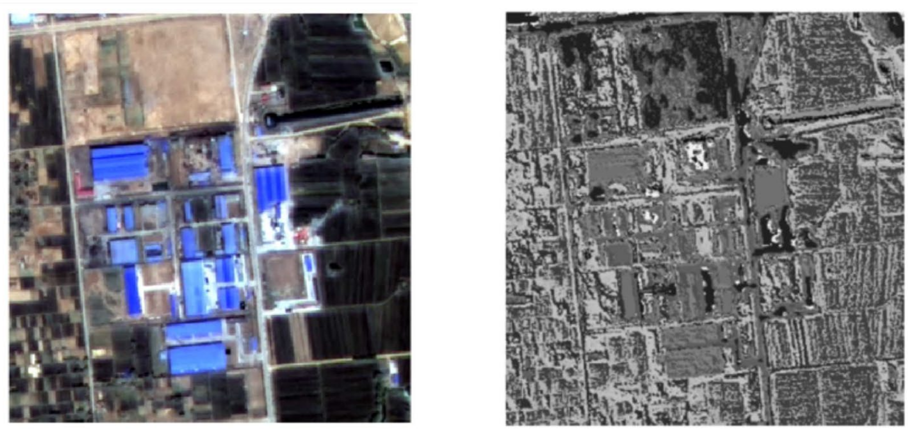
Step 4 For each pixel $\tilde{x}_i$ in image $\tilde{I}$ from testing set, we firstly compute its corresponding visual word pattern and then achieve its pixel-level classification label l_i, which could be represented as follows:

$$l_i = C(f_h(f_g(\tilde{x}_i))), i = 1, 2, ..., \tilde{N} \tag{8}$$

where function $f_g()$ refers to operations in Step 1, $\tilde{N}$ is the pixel number of $\tilde{I}$.

Specifically, we first build visual word histogram based on visual word frequency information around neighbourhood of each pixel, which involves local spatial information to improve feature distinctive ability. Afterwards, visual word histogram is regarded as a visual word pattern for pixel-level classification. Finally, we choose SVM to work as a supervised classifier, which are trained with visual word patterns. In fact, SVM is capable to construct complex and high-dimensional hyper-plane between different categories, which is suitable for classification tasks on remote sensing images due to its distinctive ability and high efficiency. Based on former discussions, we write down detailed steps of pixel-level classification in Algorithm 2.

It's noted that we define classification labels for pixel as three categories, namely water body, vegetation and



(a) Sample Remote Sensing Image (b) Corresponding Visual Word Image

Fig. 2 Samples of original remote sensing image and its corresponding visual word image

building. We show samples of original image and its corresponding pixel-level classification labels in Fig. 3.

3.3 Water body variation detection algorithm

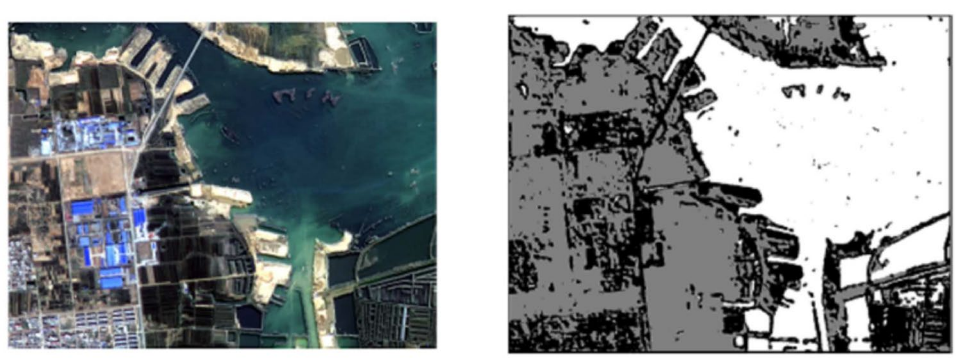
After the above processing on remote sensing images, we can obtain pixel-level classification results. In this subsection, we aim to perform water body variation detection based on locating pixels with varying classification labels.

Based on observations of classification results, we find pixels assigning with water body label are mostly accurate. However, there exist a small number of misclassification results on pixels with labels of building and vegetation. Since RBI feature is designed to perform classification on buildings and vegetation, we adopt RBI feature to modify pixel-level classification results before detection.

Considering input are two remote sensing images for the same location captured at two different time t_1 and t_2, we firstly obtain the corresponding classification labels L_{t_1} and L_{t_2} for both images. Afterwards, we use a simplified version of CVA method [33] to detect variations, where the core idea of CVA is to find locations of pixels with different labels by comparing L_{t_1} and L_{t_2}. Since there exist noise and several misclassification results after locating varying pixels, we propose to eliminate them by post-processing, such as image erosion and expansion operations to remove subtle variations in connection lines and small noisy points.

4 Distributed processing

Task of water body variation detection requires to deal with large amount of remote sensing image data, which leads us to apply the idea of distributed processing for instant feedback [34]. Moreover, the proposed method is fit with distributed processing, due to these two reasons: (1) most data processing units could perform the same algorithm, i.e., water body variation detection; (2) variables involved in the algorithm calculation are only extracted from the local neighborhood. Based on these two characteristics of the proposed method, we aim to apply MapReduce structure to perform distributed processing on the task of water body variation detection, which could largely improve running



(a) Sample Remote Sensing Image (b) Corresponding Pixel-level Classification Labels

Fig. 3 Samples of original remote sensing image and its corresponding pixel-level classification labels, where white, grey and black colors refers to water body, vegetation and building, respectively speed of the proposed method, especially when dealing with large amount input of remote sensing images.

We show workflow of distributed water body variation detection method in Fig. 4 based on MapReduce structure. Firstly, we transform the proposed water body variation detection method into two tasks: Map and Reduce. Afterwards, master node analyzes the input remote sensing image and distributes each block of the input image as a parameter of Map task to the corresponding slave nodes. After getting Map task from master node, slave nodes download input image from HDFS, execute Map task, i.e., water body variation detection illustrated in the last section, based on block information, and upload the corresponding results to HDFS. Finally, master node distributes Reduce task to recycle all the results on HDFS computed by slave nodes.

4.1 Block division strategy

Compared with text stream file processing with MapReduce, remote sensing images own two-dimensional matrix property with additional spatial information, which makes it not proper to directly apply MapReduce on the task of water body variation detection. We thus propose a localized and block-based processing strategy for remote sensing images, where the parallel running of water body variation detection is realized by firstly dividing the remote sensing image into blocks and then allocating the necessary information for each block. It is noted that such strategy could help obtain the global feature description for the whole image, meanwhile keeping necessary connection among blocks.

There exist a quantity of block division methods, such as rectangle block division, horizontal or vertical strip division methods [35]. After experiments, we adopt a rectangular block division method, i.e., divide image equally in both horizontal and vertical directions, which not only maintains the visual effects of blocks after segmentation, but also keeps similar characteristics with the original image.

It's noted that many methods adopt overlapping division method to offer additional information for edge pixels of blocks. In the proposed method, we download the whole image from the master node and process the corresponding block based on the allocated block information. Blocks thus remain close relations with others, where edge pixels for a block can easily access their local neighbourhood pixels. Moreover, overlapping brings complexity in integrating, since feature values for overlapping regions can vary after calculation in different blocks.

To sum up, we adopt a rectangle block division method without overlapping regions, resulting in block information for each slave node.

4.2 MapReduce programming structure

MapReduce is a programming model implemented in Hadoop for performing certain parallelizable processes on large amounts of data on a large number of computer nodes, often called clusters or grids. MapReduce makes it very easy for developers with no experience in distributed computing to write parallel distributed computing programs, without having to worry about machine communication, data distribution, or load balancing.

MapReduce encapsulates these complexities by simply providing an interface that automates the concurrent execution of large data-volume computations on unstructured file systems as well as on structured databases, and can effectively take advantage of the local nature of the data by processing it at a location near the data nodes to reduce the amount of data that can be processed in a single node. The MapReduce model is general enough to translate many different types of problems into specific forms of post-processing, so a big part of the developer's job is to analyze and translate the processing problem.

To support parallel running, MapReduce transforms processing into two phases, i.e., Map and Reduce. Developers are required to design their own specific Map and Reduce operations according to scenarios. The key issue in such design is thus to properly transform algorithm into a new version under the guidance of MapReduce framework. Specifically, the proposed water body variation algorithm

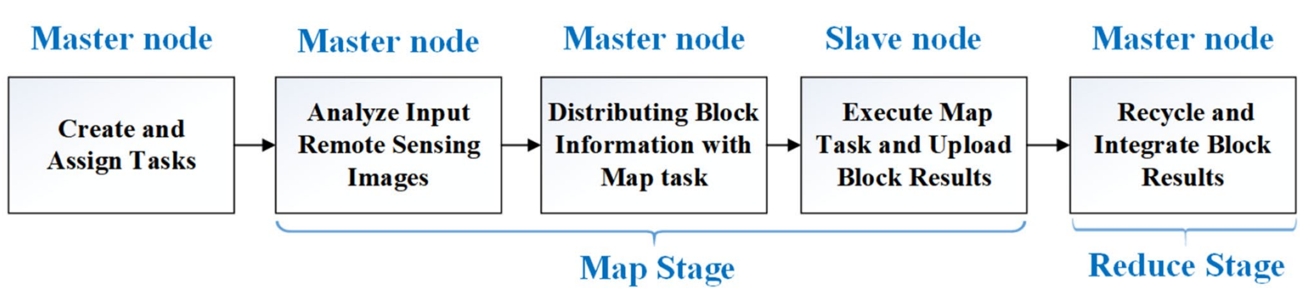


Fig. 4 Workflow of water body variation detection method with distributed processing, which utilizes MapReduce structure to facilitate parallel running

under restrictions of MapReduce programming could be represented as follows:

- Map: Master node first reads the input image and decide how to divide the image into blocks by analyzing the whole image. Afterwards, master node assigns block information and Map task as input key-value pair to slave nodes. It's noted each Map task corresponds to one specific block. Then, slave nodes read the whole image locally from HDFS and perform water body variation detection based on its corresponding block information. Finally, variation detection results computed by Map tasks in slave nodes are uploaded to HDFS.
- Reduce: Since Map phase has completed water body variation detection on blocks, Reduce phase is capable to integrate variation detection results. Specifically, once users require to achieve merged detection result, Reduce phase will obtain block detection results in HDFS and superimpose them according to block information, resulting in the merged detection result.

It's noted that block information in input parameters for slave nodes refers to a text line, which contains information about block number, starting offset of the block, and internal data offset of the block. Block information will be distributed along with Map tasks to slave nodes in the form of key-value pairs.

5 Experimental results and discussions

In this section, we first introduce our dataset. Then, we discuss measurements adopted during the experiments. Afterwards, we conduct experiments to show the effectiveness of the proposed water body variation detection algorithm. Finally, we perform experiments to prove the efficiency of the proposed distributed variation detection algorithm.

5.1 Experimental dataset

Since distributed variation detection algorithm requires large amount of remote sensing data to prove efficiency, its adopted big data will not be fit to conduct experiments with the original variation detection algorithm, due to large runtime cost. We thus perform two groups of experiments on different datasets.

We firstly adopt a relatively small number of Level-1A PMS remote sensing images captured by GF-1 satellite as a dataset for original variation detection algorithm, where these images correspond to surrounding areas of Shiliang river reservoir (located in northeast Jiangsu Province, China) within three groups representing different time periods. These original images include PAN data with a spatial resolution of 2 m and MSS data with a spatial resolution of 8 m. After pre-processing, we successfully obtain fused multi-spectral data with a spatial resolution of 2 m. Remote sensing data of the same month in different years are chosen to eliminate pseudo-variation information caused by seasonal change. Above all, we choose several sets of images corresponds to the same area at different time for water body variation detection.

To prove the efficiency of the proposed distributed variation detection algorithm, we adopt a large number of Level-1A PMS remote sensing images captured by GF-1 satellite as a dataset, where these images correspond to surrounding areas of Luoma Lake (located in northern Jiangsu Province, China) in December 2014 and December 2015. After pre-processing with the ENVI 5.1 tool, PAN data and MSS data with different spatial resolution are fused to obtain the multi-spectral data with a spatial resolution of 2 m.

We show three groups of experimental remote sensing data in Fig. 5, where each group contains images of the same area captured at different time. From Fig. 5, we can notice they contain multiple types of ground objects like buildings, vegetation and water body. After observation, we can find large visual differences of objects with the same category in one specific image, which brings great difficulties to pixel-level classification and variation detection of the water body.

5.2 Experimental measurements

To prove the effectiveness of the proposed water body variation detection algorithm, We adopt three measurements, i.e., false detection rate, missed detection rate and correct rate, to evaluate performance of water body variation detection algorithm. To better explain these three measurements, we describe how to calculate these measurements in Fig. 6.

Based on Fig. 6, we can notice false detection rate represents the proportion of pixels that are actually Changed (represented as UC) but not detected as Changed (represented as UC + CC). Meanwhile, missed detection rate represents the proportion of pixels that are actually Changed but detected as Unchanged (represented as CU) to pixels that are detected as Unchanged (represented as CU + UU). At last, the correct rate represents the proportion of pixels in which variation detection results are consistent with the real variation situation. The lower values of false detection rate and missed detection rate, the better the performance of water body variation detection algorithm. On the contrast, higher correct rate implies better performance achieved by the proposed algorithm.

Traditional measurements to measure the processing performance of MapReduce framework are acceleration ratio and execution efficiency. However, we design the proposed distributed variation detection algorithm to prove the improvement in efficiency by adopting block processing

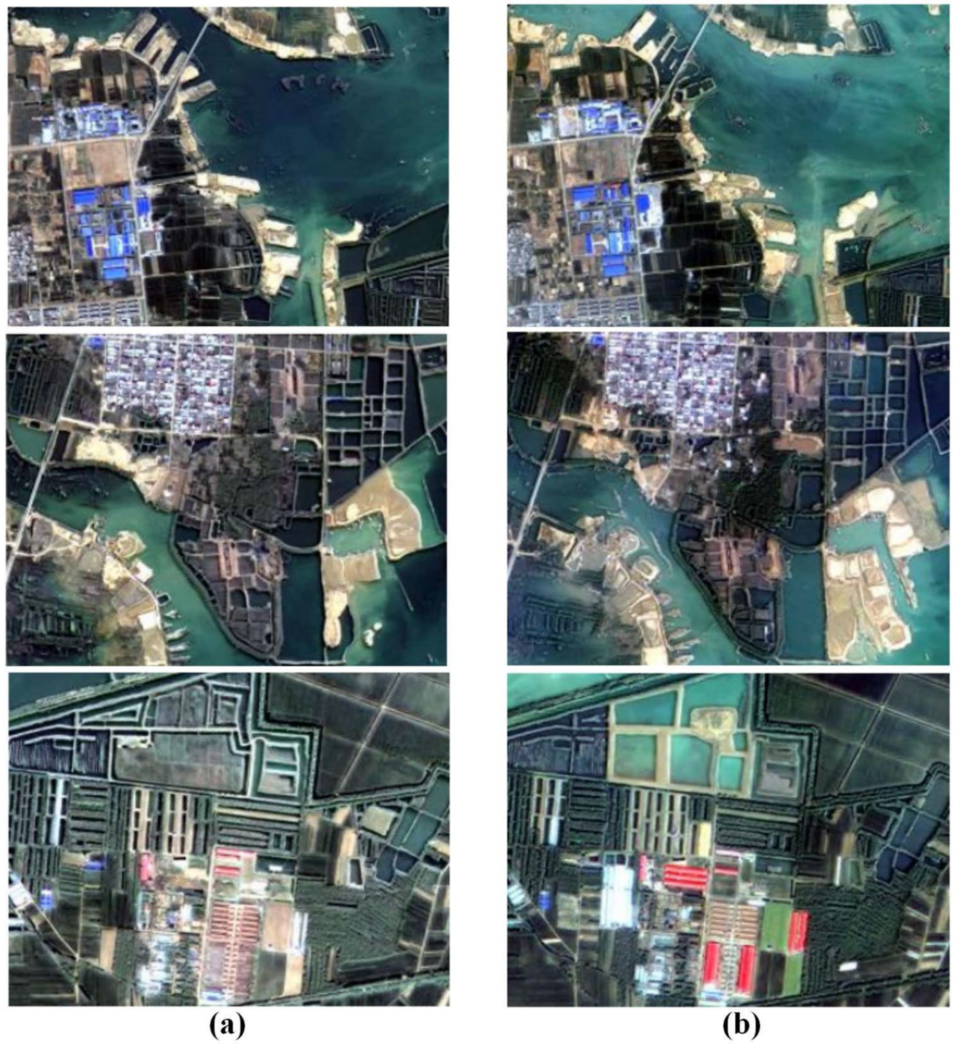


Fig. 5 Sample groups of experimental remote sensing images, where **a**, **b** correspond to the same area captured at different time, respectively

to deal with remote sensing images. Therefore, we build Hadoop environment on a single machine and adopt total time consumption of MapReduce jobs as the first measurement. To prove the improvement of run-time speed through block processing, we novelly design acceleration ratio S_{div} as another measurement:

$$S_{\mathrm{div}} = T_1/T_{\mathrm{p}}, \tag{9}$$

where T_1 represents cost time without block design, and T_{p} represents cost time in parallel processing with block design. Therefore, acceleration ratio describes the improvement in run-time speed of the proposed algorithm with block and parallel processing, compared with an algorithm without block design.

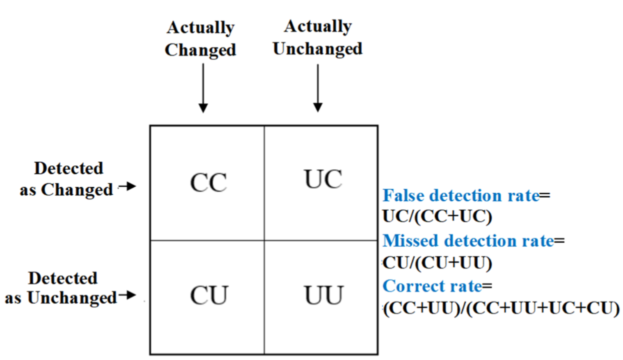


Fig. 6 Explanation on the adopted measurements to evaluate performance of water body variation detection

5.3 Water body variation detection results

In this subsection, we use pixel-level classification results to detect variations of the water body in experimental remote sensing data. We represent samples of water body variation detection results in Fig. 7, where we show comparisons

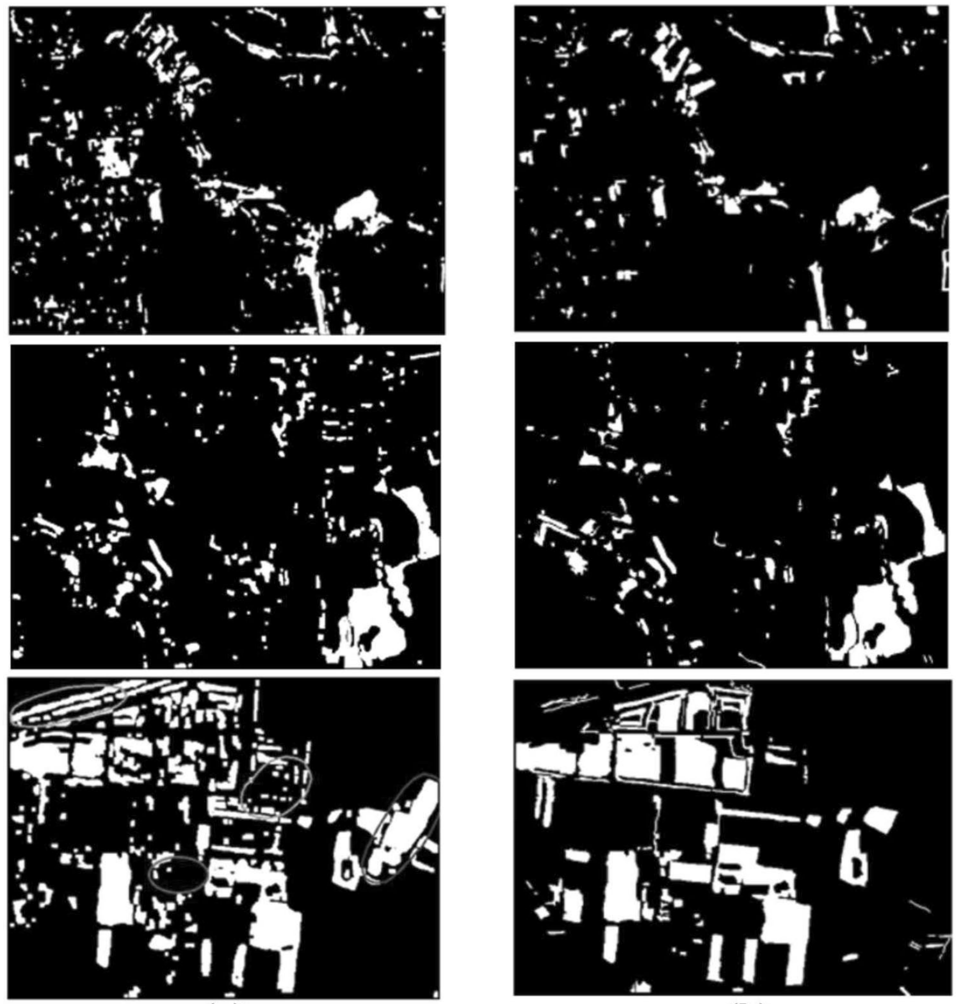


Fig. 7 Comparisons between detection results achieved by the proposed method and ground-truth images

between detection results achieved by the proposed method and ground-truth images. Circles labeled in the last row show situation of noise, which requires the proposed post-processing algorithms to deal with. It's noted that result samples in Fig. 7 correspond to images shown in Fig. 5.

Details of the comparative experiment is presented in Figs. 8, 9 and 10, where we adopt several comparison methods to perform variation detection, including object-based CVA method, variation detection method based on Chi-square transform (CST) [36], based on histogram (HIST) [37] and based on forward-backward heterogeneity (FB) [38].

In Table 1, we could observe that the proposed method can obtain better performance than comparative methods in water body variation detection, which could be proved by a higher correct rate, lower false detection rate and lower missed detection rate. Such performance is achieved due to proper usage of multiple features, including spectral, LBP texture and so on. Besides, Mapping from the fused feature to visual vocabulary successfully abstract characteristics

Table 1 Comparisons on effectiveness measurements between the proposed method and other comparative studies, where FDR, MDR and CR refer to false detection rate, missed detection rate and correct rate, respectively

Group	Method	FDR (%)	MDR (%)	CR (%)
Group1	CVA	73.5951	3.79041	87.9291
	CST	67.9402	3.84003	90.1792
	HIST	74.1219	3.82744	87.7388
	FB	58.8818	3.80054	92.2548
	Our method	36.5856	1.15717	95.8177
Group2	CVA	57.3363	4.70468	90.3878
	CST	54.5328	4.22390	90.8722
	HIST	58.6109	4.72451	90.1027
	FB	39.6742	4.83040	93.0362
	Our method	28.7140	0.89185	96.2023
Group3	CVA	35.8274	13.6124	83.9229
	CST	26.4467	12.3160	86.0905
	HIST	41.0264	12.8891	83.2444
	FB	18.3772	12.0749	87.2775
	Our method	38.2017	7.27053	85.9500

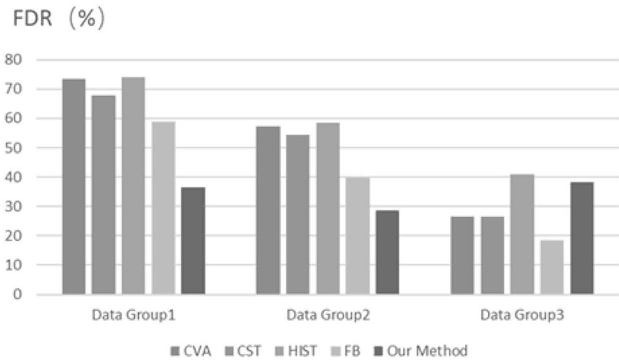


Fig. 8 Comparisons on false detection rate achieved by the proposed method and comparative methods

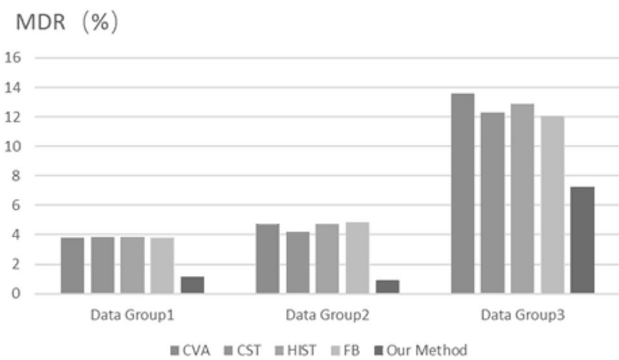


Fig. 9 Comparisons on missed detection rate achieved by the proposed method and comparative methods

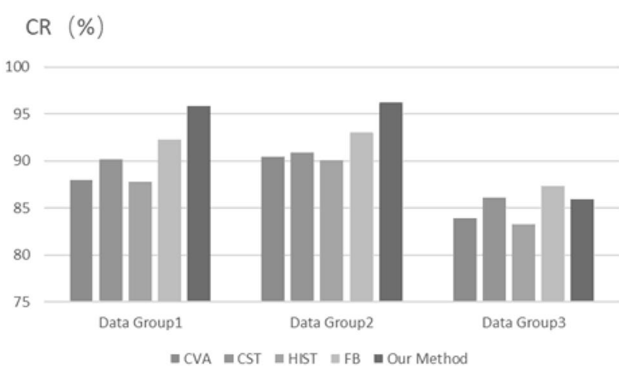


Fig. 10 Comparisons on correct rate achieved by the proposed method and comparative methods

of ground objects, meanwhile visual vocabulary histogram offers abundant information to help perform pixel-level classification.

Table 2 Comparisons on efficiency measurements with different parameters, i.e., block no.

Index	Block no.	Time (s)	Acceleration ratio
1	1	1545.50	1.00000
2	4	796.615	1.94252
3	9	537.798	2.87375
4	16	548.356	2.81842

5.4 Detection results with distributed processing

In this subsection, we mainly carry on experiments to prove improvements in efficiency brought by introducing distributed processing. Table 2 shows quantitative comparative results, where we conduct experiments with different settings of block number. It's noted block number refers to block division amount in both horizontal or vertical directions.

From Table 2, we can see the distributed method can improve the efficiency of water body variation detection. By defining block number with a larger value, the distributed method could obtain faster run-time speed, which could be proved by shorter total execution time and larger acceleration ratio value. However, we could observe total execution time and acceleration ratio reaches top and remain unchanged, when comparing between results achieved by defining block number as 3 and 4. This phenomenon is due to the limitation of computing resources in the pseudo-distributed platform for stand-alone machine. In fact, the total number of Map parallel tasks increases exponentially with the exponential growth of total number of blocks, resulting in large costs on task allocation, task scheduling, data storage and data migration. These costs quickly reach the processing bottleneck of single-computer computing resource, which makes the performance improvement brought by distributed processing quickly reach unchanged. We believe the proposed method can achieve real-time effects with more computer nodes for parallel computing. Further analysis shows that the distributed processing method does not affect the detection accuracy of the change detection algorithm because, on the one hand, the blocking strategy does not lead to edge problems between blocks, which makes the effective processing area of the image still remains consistent; on the other hand, the image data reading method makes each child node of the blocking process can still get the image unblocked. The global features and parameters of the former, i.e., the processing process can still retain the intrinsic relationship between the various segments of the image, instead of separating them from each other and working in isolation.

6 Conclusion

In this paper, we propose a novel detection method to help locates abnormal changes around the lake and reservoir caused by human activities. Furthermore, we successfully transform the original algorithm to parallel running version following the guidance of MapReduce structure. Specifically, we firstly construct mixed feature space to offer abundant descriptive information on remote sensing images. Afterwards, a bag of visual words are adopted to abstract neighboring feature information around pixels, which forms highly distinctive visual word patterns for pixel-level classification. After detecting label variations in images, the proposed method locates change areas to offer instant feedback. To improve the efficiency of the proposed method, we finally propose block division method under restrictions of MapReduce structure to help facilitate parallel running. Our proposed method for distributed variation detection based on a bag of visual words has been used on real-world rivers. Experiments on GF-1 PMS data have shown that we achieve better effects on the extraction of variation information. Therefore, the proposed model can be applied as a technique for real-time variance detection in remote sensing imagery.

In the future, this combination of technologies will allow us to monitor water bodies in real time and to detect "abnormal" changes around the lake and reservoir due to human activities, because of the rich feature details in high-resolution remote sensing images. And we will aim to adopt latest unsupervised methods to explore fast and robust algorithms to detect abnormal. Furthermore, big data domain is fast developing with quantity of new ideas and technologies, such as edge [39], cloud [40], fog [41] and so on. We believe the proposed distributed version can be further improved with these latest methods.

Acknowledgements Funding This work was supported by National Key R\&D Program of China under Grant 2018YFC0407901, the Fundamental Research Funds for the Central Universities under Grant B200202177, and the National Natural Science Foundation of China under Grant 61702160.

References

1. Xiaolong, X., Zhang, X., Gao, H., Xue, Y., Qi, L., Dou, W.: Become: blockchain-enabled computation offloading for IoT in mobile edge computing. IEEE Trans. Ind. Inform. **16**(6), 4187–4195 (2020)
2. Qi, L., Zhang, X., Li, S., Wan, S., Wen, Y., Gong, W.: Spatial-temporal data-driven service recommendation with privacy-preservation. Inf. Sci. **515**, 91–102 (2020)
3. Zhang, Y., Liu, X., Zhang, Y., Ling, X., Huang, X.: Automatic and unsupervised water body extraction based on spectral-spatial features using gf-1 satellite imagery. IEEE Geosci. Remote Sens. Lett. **16**(6), 927–931 (2019)
4. Kai, J., Jiang, W., Jing, L., Tang, Z.: Spectral matching based on discrete particle swarm optimization: a new method for terrestrial water body extraction using multi-temporal landsat 8 images. Remote Sens. Environ. **209**, 1–18 (2018)
5. Wang, N., Jing, W., Li, L.: An improved distributed storage model of remote sensing images based on the hdfs and pyramid structure. Int. J. Comput. Appl. Technol. **59**(2), 142 (2019)
6. Wang, P., Wang, J., Chen, Y., Ni, G.: Rapid processing of remote sensing images based on cloud computing. Future Gener. Comput. Syst. **29**(8), 1963–1968 (2013)
7. Luo, B., Jiang, S., Zhang, L.: Indexing of remote sensing images with different resolutions by multiple features. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. **6**(4), 1899–1912 (2013)
8. Han-Qiu, X.U.: A study on information extraction of water body with the modified normalized difference water index (MNDWI). J remote sens. **5**, 589–595 (2005)
9. Haipeng, L.: Prediction and analysis of chaotic time series on the basis of support vector. Electron. J. Syst. Eng. **19**(4), 806–811 (2008)
10. Ozturk, C.N., Bilgin, G.: A comparative study on manifold learning of hyperspectral data for land cover classification. In: Sixth International Conference on Graphic and Image Processing (ICGIP 2014), vol. 9443, p. 94431L. International Society for Optics and Photonics (2015)
11. Feng, X.U., Cheng, H.U., Jun, L.I., Plaza, A., Datcu, M.: Special focus on deep learning in remote sensing image processing. Sci. China **063**(004), P.1–P.2 (2020)
12. Mnih, V., Hinton, G.E.: Learning to detect roads in high-resolution aerial images. In: European Conference on Computer Vision, pp. 210–223. Springer (2010)
13. Chaib, S., Liu, H., Yanfeng, G., Yao, H.: Deep feature fusion for VHR remote sensing scene classification. IEEE Trans. Geosci. Remote Sens. **55**(8), 4775–4784 (2017)
14. Paisitkriangkrai, S., Sherrah, J., Janney, P., Hengel, V.-D., et al.: Effective semantic pixel labelling with convolutional networks and conditional random fields. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 36–43 (2015)
15. Hu, Y., Cahill, N.D., Messinger, D.W.: Low-dimensional representations of hyperspectral data for use in CRF-based classification. In: Image and Signal Processing for Remote Sensing XXI (2015)
16. Zhao, W., Du, S., Wang, Q., Emery, W.J.: Contextually guided very-high-resolution imagery classification with semantic segments. ISPRS J. Photogramm. Remote Sens. **132**, 48–60 (2017)
17. Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J.: Pyramid scene parsing network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2881–2890 (2017)
18. Sun, W., Wang, R.: Fully convolutional networks for semantic segmentation of very high resolution remotely sensed images combined with dsm. IEEE Geosci. Remote Sens. Lett. **15**(3), 474–478 (2018)
19. Wang, Q., Liu, S., Chanussot, J., Li, X.: Scene classification with recurrent attention of VHR remote sensing images. IEEE Trans. Geosci. Remote Sens. **57**(2), 1155–1167 (2018)
20. Sun, Y., Lei, L., Li, X., Sun, H., Kuang, G.: Nonlocal patch similarity based heterogeneous remote sensing change detection. Pattern Recognit. **109**, 107598 (2021)
21. Hua, X., Wang, X., Rui, T., Zhang, H., Wang, D.: A fast self-attention cascaded network for object detection in large scene remote sensing images. Appl. Soft Comput. **94**, 106495 (2020)
22. Cheng, G., Yang, C., Yao, X., Guo, L., Han, J.: When deep learning meets metric learning: remote sensing image scene

classification via learning discriminative CNNs. IEEE Trans. Geosci. Remote Sens. **56**(5), 2811–2821 (2018)
23. Zhang, J., Chaoquan, L., Li, X., Kim, H.-J., Wang, J.: A full convolutional network based on densenet for remote sensing scene classification. Math. Biosci. Eng **16**(5), 3345–3367 (2019)
24. Alahmadi, A., Joorabchi, A., Mahdi, A.E.: Combining bag-of-words and bag-of-concepts representations for arabic text classification. In: Irish Signals and Systems Conference and China–Ireland International Conference on Information and Communications Technologies (2014)
25. Zhao, L.-J., Tang, P., Huo, L.-Z.: Land-use scene classification using a concentric circle-structured multiscale bag-of-visual-words model. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. **7**(12), 4620–4631 (2014)
26. Wang, Q., Wan, S., Yue, L., Che, W.: Visual attention based bag-of-words model for image classification. In: Sixth International Conference on Digital Image Processing (2014)
27. Karakasis, E.G., Amanatiadis, A., Gasteratos, A., Chatzichristofis, S.A.: Image moment invariants as local features for content based image retrieval using the bag-of-visual-words model. Pattern Recognit. Lett. **55**, 22–27 (2015)
28. Jinying, Z.H.A.N.G., Guanghu, Y.A.O., Lin, L.I.N., Huaixuan, G.U.O.: Automatic classification of gf-2 remote sensing imagery based on active learning and bag of visual words model. Bull. Surv. Mapp. **2**, 103 (2019)
29. Peng, X., Wang, L., Wang, X., Yu, Q.: Bag of visual words and fusion methods for action recognition: comprehensive study and good practice. Comput. Vis. Image Underst. **150**, 109–125 (2016)
30. Huo, L.: Land-use scene classification using a concentric circle-structured multiscale bag-of-visual-words model. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. **7**(12), 4620–4631 (2014)
31. Zhang, J., Li, T., Lu, X., Cheng, Z.: Semantic classification of high-resolution remote-sensing images based on mid-level features. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. **9**(6), 2343–2353 (2016)
32. Zhang, L., Jing, Z., Zhang, D., Hou, X., Gang, Y.: Urban road extraction from high-resolution remote sensing images based on semantic model. In: The 18th International Conference on Geoinformatics: GIScience in Change, Geoinformatics 2010, 18–20 June, 2010. Peking University, Beijing, China (2010)
33. Lambin, E.F., Strahlers, A.H.: Change-vector analysis in multitemporal space: a tool to detect and categorize land-cover change processes using high temporal-resolution satellite data. Remote Sens. Environ. **48**(2), 231–244 (1994)
34. Qi, L., Chen, Y., Yuan, Y., Shucun, F., Zhang, X., Xiaolong, X.: A QoS-aware virtual machine scheduling method for energy conservation in cloud-based cyber-physical systems. World Wide Web **23**(2), 1275–1297 (2020)
35. Wang, R., Liu, Z., Chen, B.: Design and implementation of a decentralized selfcoordinating distributed remote sensing image processing system. In: Proceedings of Spie, 7146 (2009)
36. Wei, L., Li, P., Zhang, L., Zhong, Y.: An advanced change detection method based on object-oriented classification of multi-band remote sensing image. In: International Conference on Geoinformatics (2010)
37. Li, X.B., Zhou, Q.: A lossless data hiding transmission method for satellite remote sensing image based on histogram modification. J. Astronaut. **34**(5), 686–692 (2013)
38. Shijin, L.I., Wang, S., Huang, L.: Change detection with remote sensing images based on forward-backward heterogenicity. J. Shandong Univ. **48**(03), 1–9 (2018)
39. Xu, X., Shen, B., Yin, X., Khosravi, M.R., Wu, H., Qi, L., Wan, S.: Edge server quantification and placement for offloading social media services in industrial cognitive IoV. IEEE Trans. Ind. Inform. (2020)
40. Xu, X., Zhang, X., Liu, X., Jiang, J., Qi, L., Zakirul Alam Bhuiyan, Md.: Adaptive computation offloading with edge for 5g-envisioned internet of connected vehicles. IEEE Trans. Intell. Transp. Syst. 1–10 (2020)
41. Qi, L., He, Q., Chen, F., Zhang, X., Dou, W., Ni, Q.: Data-driven web APIs recommendation for building web applications. IEEE Trans. Big Data 1 (2020)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Yirui Wu received the B.Sc., M.Sc. and Ph.D. degrees in computer science from Nanjing University in 2011, 2013 and 2016, respectively. He served as an Assistant Professor with the Department of Computer and Information, Hohai University, from 2016 to 2019. He is currently an associate Professor with Hohai University. His current interests are in the areas of multimedia, edge computing, computer vision, and pattern recognition algorithms/systems. He has published over 30 papers in his area of interest.

Peifeng Han is now pursing master degree in computer science from Hohai University. His research interests include multimedia processing and computer vision.

Zhan Zheng received his master degree in College of Computer and Information from Hohai University, China, in 2017. His research interests include image processing, big data and pattern recognition.