ISPRS Journal of Photogrammetry and Remote Sensing 146 (2018) 38–51

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Iterative feature mapping network for detecting multiple changes in multi-source remote sensing images

Tao Zhan, Maoguo Gong*, Jia Liu, Puzhao Zhang

Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, International Research Center for Intelligent Perception and Computation, Xidian University, Xi'an, Shaanxi Province 710071, China

ARTICLE INFO

Keywords:
Change detection
Iterative feature mapping network
Hierarchical clustering analysis
Multiple changes
Multi-source images

ABSTRACT

Owing to the rapid development of remote sensing technology, various types of data can be easily acquired at present. However, it has become an important but more challenging task for effectively highlighting changes occurring on the land surface from these available data. In this paper, we propose an iterative feature mapping network learning framework for identifying multiple changes with focus on multi-source images, which are often obtained from sensors with different imaging modalities. Firstly, high-level and robust feature representations are extracted from multi-source images via unsupervised feature learning. Then, on this basis, an iterative feature mapping network is established to transform these features into a common high-dimensional feature space. It aims to learn more discriminative features by shrinking the difference between the paired features of unchanged positions while enlarging that of changed ones. Note that the network parameters are learned by optimizing a well-designed objective function, and the whole learning process is fully unsupervised. Finally, based on a hierarchical tree for clustering analysis, all possible change classes can be detected accurately. In addition, the proposed framework is found to be also suitable for change detection in homogeneous images. The impressive experimental results obtained over different types of remote sensing images demonstrate the effectiveness and robustness of the proposed model.

1. Introduction

Over the past decades, a comprehensive understanding of changes occurring on the Earth's surface has become increasingly important for environmental monitoring. With the rapid development of remote sensing technology, various types of data can be easily acquired now. Based on change detection techniques (Lu et al., 2004), we can observe the global changes more precisely by fully exploiting these data. Change detection is the process for detecting changes taking place on the same geographical area by analyzing two (or more) images that are acquired at different times (Singh, 1989). It has been widely used in many fields, such as land-use and land-cover evaluation (Gil-Yepes et al., 2016), urban growth monitoring (Taubenböck et al., 2012), natural disaster assessment (Brunner et al., 2010), etc. Due to the huge volume of remote sensing images and different imaging modalities between them, it is promising but challenging to highlight changes from multi-source images. Therefore, it is essential to develop novel change detection techniques that can automatically identify changes for different types of remote sensing images.

Generally, change detection techniques can be divided into two categories based on homogeneous and multi-source images according to imaging modalities. Here, homogeneous images refer to images coming from the satellite sensor(s) with the same imaging modality, e.g., radar or optical sensors. Among the existing change detection methods, most of them are designed for homogeneous data, and they are often composed of the following three steps: (1) data preprocessing, the two captured images are coregistered and corrected to make their spatial positions consistent and eliminate the influence of atmosphere; (2) generating a difference image (DI), aiming to extract change information from the considered images; and (3) identifying changes, threshold and clustering methods are widely used to segment the DI into changed and unchanged classes. The aim of threshold-based methods is to find an optimal threshold to distinguish changed pixels from unchanged ones, e.g., the expectation-maximization (EM) algorithm (Bruzzone and Prieto, 2000) and the generalized Kittler and Illingworth thresholding (GKIT) algorithm (Moser and Serpico, 2006). However, these methods are often very sensitive to noise that leads to bad performance, especially when synthetic aperture radar (SAR) images are used to perform the change detection task. In contrast, clustering methods (Gong et al., 2012; Krinidis and Chatzis, 2010) are

* Corresponding author.
E-mail address: gong@ieee.org (M. Gong).

https://doi.org/10.1016/j.isprsjprs.2018.09.002
Received 27 January 2018; Received in revised form 31 August 2018; Accepted 1 September 2018
0924-2716/

more attractive because of the use of context information, which makes them robust to noise and easy to highlight changes. In addition, some advanced techniques have been proposed, such as genetic algorithms (Celik, 2010), level set-based methods (Bazi et al., 2010), and neural network-based approaches (Ghosh et al., 2013; Gong et al., 2016, 2017b).

With the advance of remote sensing technology, multi-source data are available now. Different from homogeneous images, they are often acquired from sensors with disparate imaging modalities (Prendes et al., 2015b; Touati and Mignotte, 2018). Nevertheless, traditional approaches (e.g., those methods mentioned above) are incapable to perform change detection tasks in multi-source images, due to the great difference in their appearances and statistical properties. Therefore, there exists an urgent demand for processing multitemporal and multi-source images. Specially, it is of great practical significance under some practical scenarios. For a suddenly occurring disaster (e.g., earthquake or flood), the pre-event optical image can be obtained from archived data of remote sensing platforms, whereas maybe only the post-event SAR image can be captured for the same area because of the constraints of weather and illumination. In fact, SAR and optical images are complementary to some extent. Owing to the active imaging characteristic, SAR sensors can record the scene information in all weather and at all times. By contrast, optical sensors can acquire high-quality images including abundant spectral and textural information of ground objects, but are often susceptible to weather conditions, such as cloud or sunlight.

Due to the disparity of multi-source images in appearance, it is difficult to make a direct comparison by calculating the pixelwise difference between them, as well as to identify changed regions accurately. Even so, there are some works with respect to multi-source images. Mercier et al. (2008) proposed a copula-based quantile regression theory to model the dependence of unchanged regions, transforming the first image to achieve the similar statistic characteristics to the second one. In addition, the kernel canonical correlation analysis (Volpi et al., 2015) and the Bayesian nonparametric model coupled with a Markov random field (Prendes et al., 2015a) were also proposed. However, the performance of these methods strongly depends on handcrafted analysis of image properties, requiring enough labeled unchanged pixels for learning the inherent relationship between multi-source images. This process is usually costly and time-consuming, which further limits the application of such approaches. For multi-source images, actually, they share a common basis that both of them are the information representations of the same area, which indicates that they are comparable in some particular spaces. Inspired by this fact, it is natural to transform them into a high-dimensional feature space for change analysis. Liu et al. (2018) proposed an unsupervised symmetric coupling convolutional network (SCCN) for change detection based on heterogeneous radar and optical images. However, this method ignores the influence of changed pixels. To address this problem, Zhao et al. (2017) proposed an approximately symmetrical deep neural network (ASDNN) for multi-source images by considering the influence of unchanged and changed pixels simultaneously, and obtained good performance.

Note that most of the aforementioned methods are mainly designed for addressing the binary change detection problem, whether the given images are acquired from different sensors or not. Binary change detection methods aim to only distinguish changed pixels from unchanged ones, without considering any semantic information of changes. That is to say, all changes that occurred on the land surface are regarded as a single change class. By contrast, detecting different types of changes is an interesting but more challenging task. When the ground truth information is available, the most popular method is the supervised post-classification comparison (PCC) (Serra et al., 2003), which can identify multiple changes via a pixel-by-pixel comparison between the classification maps that are independently classified in advance. However, the performance of this approach relies heavily on the accuracy of classification algorithms. Thereby, the multidate classification (Meddens et al., 2013) method was proposed, which can classify two images (acquired at different dates) simultaneously, thus reducing the accumulated classification errors. For such methods, collecting the ground truth information is rarely feasible because of the constraints of practical conditions. In recent years, unsupervised change detection methods are becoming more attractive, because they do not need any prior knowledge of observed data. Bovolo et al. (2012) proposed a compressed change vector analysis (C^2VA) approach for detecting multiple changes in multitemporal multispectral images. Theoretically, it allows one to discriminate all kinds of changes by making full use of the spectral information of images. But some subtle changes within major changes are difficult to separate in a single operation using this method. To overcome this drawback, an unsupervised hierarchical change detection scheme was proposed to identify all potential changes in Prendes et al. (2015b). However, these methods are proposed for homogeneous images, which cannot be directly applied to detect multiple changes in multi-source images. Furthermore, these methods are operative only when changed regions are highlighted accurately, whereas this goal is difficult to achieve in real applications.

In this paper, we propose a novel iterative-feature-mapping-and-hierarchical-clustering-analysis-based (IFM-HCA) change detection framework for detecting different types of changes in multi-source remote sensing images. Firstly, high-level and robust features are extracted from multi-source images via unsupervised feature learning. Due to the fact that these feature pairs are generated automatically and they are inconsistent in appearance, it is useless to make a direct comparison between them. To make them comparable and learn more discriminative feature representations, an iterative feature mapping network (IFMN) is established. It aims to map these feature pairs into a common high-dimensional feature space by shrinking the feature difference of unchanged positions while enlarging that of changed ones. To achieve this, a novel objective function is designed for highlighting changes and suppressing unchanged regions simultaneously, and it can be optimized in an unsupervised way. Finally, based on a hierarchical tree for clustering analysis, we can detect all possible change classes accurately. In addition, as far as we know, this is the first attempt to employ deep neural networks to detect different types of changes based on multi-source images.

The rest of this paper is organized as follows: Section 2 briefly introduces the related background knowledge of this study. Section 3 describes the proposed model in detail. In Section 4, the experimental settings are presented. Section 5 shows the experimental results and discussions on different types of data sets. Finally, the conclusion of this work is drawn in Section 6.

2. Background

2.1. Feature learning models based on deep neural networks

Compared with traditional shallow neural networks, deep neural networks are with more hidden layers, which have the capacity of automatically extracting useful information representations from raw data in a hierarchical manner (LeCun et al., 2015). In recent years, many excellent deep neural networks have been proposed, such as deep convolutional neural networks (CNNs) (Krizhevsky et al., 2012) and deep belief networks (DBNs) (Hinton et al., 2006), which have been successfully applied in image processing (Hou et al., 2015; Stuhlsatz et al., 2012), speech recognition (Xue et al., 2014) and natural language understanding (Sarikaya et al., 2014). For the change detection task, higher-layer of representation can magnify aspects of input that are crucial for the identification of changes and restrain uncorrelated variations.

Generally, CNN is composed of several convolution and pooling layers, which are used to extract abstract feature representations from the raw data (e.g., image and video) and maintain the robustness to