

Journal Pre-proofs

An Object Based Framework for Building Change Analysis Using 2D and 3D Information of High Resolution Satellite Images

Hamid Mohammadi, Farhad Samadzadegan

PII: S0273-1177(20)30374-4

DOI: <https://doi.org/10.1016/j.asr.2020.05.041>

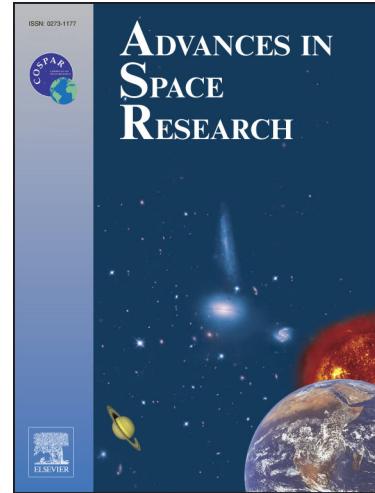
Reference: JASR 14813

To appear in: *Advances in Space Research*

Received Date: 29 October 2019

Revised Date: 22 May 2020

Accepted Date: 26 May 2020



Please cite this article as: Mohammadi, H., Samadzadegan, F., An Object Based Framework for Building Change Analysis Using 2D and 3D Information of High Resolution Satellite Images, *Advances in Space Research* (2020), doi: <https://doi.org/10.1016/j.asr.2020.05.041>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier Ltd on behalf of COSPAR.

An Object Based Framework for Building Change Analysis Using 2D and 3D Information of High Resolution Satellite Images

Hamid Mohammadi¹, Farhad Samadzadegan^{1,*}

¹ University of Tehran, College of Engineering, School of Surveying and Geospatial Engineering, North Kargar Ave., Jalal Al. Ahmad Crossing, Tehran, Iran, 1439957131.

*Corresponding Author, E-mail: samadz@ut.ac.ir

Abstract. With the rapid change of urban areas in developing countries, construction areas are constantly appearing in different parts of cities. Those changed areas require timely monitoring to provide up-to-date information for of urban information systems. As a result, it is a challenge to develop an effective change analysis of different objects, especially buildings in cities. This paper presents an object-based framework for analyzing building changes from high-resolution satellite stereo images (HRSSI). The disparity information extracted from stereo images and spectral information including visible vegetation index (VVI) help extract the buildings in a hierarchical approach. Evaluations show the accuracy of higher than 98% and F1-Score of higher than 87% for the building extraction step. Then, each building object is classified into three main categories including “Remained Building”, “Removed Building” or “Added Building”. Also, the “Remained Building” objects are categorized into four change states including “Only 2D Change”, “Only 3D Change”, “2D with 3D Change” and “No Change”. This is done by utilizing the object-based similarity analysis of the spectral information as well as the similarity analysis of the disparity information using CNN and their integration. Evaluations demonstrate the accuracy of higher than 97% and F1-Score of higher than 90% for this step.

Keywords: building extraction, building change analysis, high resolution satellite stereo images, 2D and 3D information, object-based image analysis, convolutional neural network.

1 Introduction

Building change detection and analysis is one of the most interesting topics in remote sensing. After change detection, differences in the state of objects are identified. This is done by observing objects in different times (Singh, 1989).

The rapid expansion of urban objects leads to some problems in urban management including increased traffic, air pollution and etc. (Pacifici et al., 2007). Therefore, city administrators need to be informed about the changes and type of them that have been made especially for building objects. Any information about 2D change (like changes in texture or color of building roofs) or 3D change (like changes in height of buildings) for such structures helps to make the right decisions to reduce city problems.

Also, very high resolution (VHR) satellite images, have a great potential for monitoring and mapping surface changes (Pacifici et al., 2007). But some factors like sensor noise, mis-registration errors, many details in VHR images and seasonal and meteorological effects reduce the final accuracy of change detection. This reduction would be more severe if the only 2D information is used. Therefore 2D change detection approaches are limited in VHR satellite images (Tian et al., 2015).

Because of these challenges, the solution of 3D change detection has been of interest to researchers. Actually, 3D change detection methods were considered by researchers due to the use of additional data source like disparity, height or full 3D information. Sustainability to changes in illumination, free of perspective effect (also for VHR) and providing volumetric differences are the advantages of 3D change detection approaches. But studies have shown that 3D change detection methods, also, have some challenges including uncertainties in 3D

information, combining of multi-modal data and using for different spatial resolutions and in different viewing perspective (Qin et al., 2016).

Thus, it is proposed to use 2D and 3D information simultaneously for achieving optimum results (Qin et al., 2016). Actually post-refinement with post-classification approaches are considered as potential optimal solutions. One of the new solutions in this regard, monitors 3D building change and urban re-development patterns using multi-view satellite images. They use SGM for 3D information generation. Also, they use spectral information such as NDVI and HSV for updating building maps and, then, 3D change detection (Wen et al., 2019). In another work, edge-based simple linear iterative clustering (ESLIC) is used for image super-pixel segmentation using spectral and spatial information. Next, these segments are used for DSM post-refinement and 3D building change detection (Gharibafghi et al., 2019). In another investigation, a roof-cut approach is developed using graph-based model to locate rooftops and demolished buildings through the use of imagery and pre-existing building footprint for building change detection (Gong et al., 2019). The existence of different solutions indicates that, the method used to integrate 2D and 3D information is very important and can be very effective in the quality of the final output.

Another important point is that, since the disparity map from dense stereo matching or the digital surface model (DSM) or other height information that are employed in change detection, have uncertainties, all previous methods proposed by researchers including: height differencing (Sasagawa et al., 2013, Chaabouni-Chouayakh et al., 2011, Stal et al., 2013, Dini et al., 2012, Tian et al., 2010, Chaabouni-Chouayakh and Reinartz, 2011), Euclidean distance height differencing (Zavodny, 2012, Akca et al., 2010, Champion et al., 2010, Xiao et al., 2013, Qin et al., 2014), projection-based differences (Qin, 2014a, Qin and Gruen, 2014, Ulusoy and Mundy,

2014, Taneja et al., 2013, Taneja et al., 2011, Crispell et al., 2011, Schindler and Dellaert, 2010), post-refinement algorithms (Pang et al., 2014, Chaabouni-Chouayakh and Reinartz, 2011, Guerin et al., 2014, Choi et al., 2009, Qin, 2014b), direct feature fusion (Tian et al., 2014b, Tian, 2013, Trinder and Salah, 2012, Tian et al., 2014a, Qin, 2014a) and post-classification methods (Qin et al., 2015, Nebiker et al., 2014, Champion et al., 2009, Matikainen et al., 2010, Tian, 2013), have limitations. The height differencing methods have advantage of easy implementation but sensitivity to misregistration is the limitation. Also, Euclidean distance height differencing methods require complicated implementation and are time consuming. The projection-based differences methods, also, may have missing detection in homogeneous area. For post-refinement algorithms, if some changes are missed, recovery is not available in subsequent steps. In direct feature fusion methods, it is needed to configure the parameters of fusion algorithms and any fault may lead to errors. Finally, in post-classification methods, the results highly depend to classification accuracy (Qin et al., 2016).

On the other hand, deep learning has been widely used in recent years, for image change detection (Wang et al., 2018, Gong et al., 2015, Mou et al., 2018, Reichstein et al., 2019, Peng et al., 2019, Zhang et al., 2019, Pang et al., 2019). Reichstein introduces convolutional neural networks (CNNs) as emergent or potential approaches for change detection applications (Reichstein et al., 2019). Wang applied a change detection technique based on faster R-CNN for high resolution satellite images. Peng proposes an end-to-end change detection method based on effective encoder-decoder architecture for semantic segmentation named UNet++, where change maps could be learned from scratch using available annotated datasets. These methods extract deep features from training data and increase the precision and accuracy of change detection. But, due to the need for training data, the level of automation is reduced (Wang et al., 2018). On

the other hand, in the method proposed by Zhang (Zhang et al., 2019), the change detection is based on a comparison of airborne laser scanning data and the DSM obtained from dense image matching. In the proposed method, a deep network is used to detect changes. In the first step, point cloud data and ortho images are combined to generate a raster data. In the second step, this raster data enters the deep network in the form of square patches to detect changes.

In different way, Pang (Pang et al., 2019) presented a building change detection method with a combination of co-segmentation and superpixel-based graph cuts segmentation. In the proposed method, the point clouds of two dates are combined and co-segmentation using SLIC algorithm is applied to obtain bi-temporal superpixels at first. Secondly, building areas are extracted using a semantic segmentation based on deep CNN. Finally, graph-cuts-based building change detection algorithm is proposed to extract the changed buildings. Actually, deep learning based methods, reduce the challenges of object-based methods such as need to setup segmentation parameters, but, it is clear that the geometric and contextual spatial features are lost. Therefore, it would be good if a solution could be introduced that can take advantage of both object and deep learning based methods simultaneously. It would also be good to be able to process only the patches of the image that contain the building object instead of processing the whole image.

Another aspect that has been neglected in change detection approaches is that they only identify changes. In short, no information about the type of change is available for the users. For example, a building may have 2D and 3D changes simultaneously, or may have 2D Change without any 3D change and such information can be valuable for an urban manager, but, only the change detection cannot provides such information.

In this paper, a new strategy for building change analysis from high resolution satellite stereo images is proposed. This strategy utilizes both the benefits of combining 2D and 3D information,

and the benefits of object and CNN-based methods for providing better results. The main innovations of this research are the following:

- Performing 2D and 3D information fusion to extract buildings in a hierarchical procedure;
- Proposing a two-step change analysis (first and second level change analyses) of building objects using HRSSI;
- Computing object-based structural dissimilarity of building objects using spectral information as 2D change indicator and generating 2D change map;
- Using CNN for feature extraction from 3D information on the first and second dates and measuring the dissimilarity of features as the 3D change indicator and generating 3D change map;
- Integrating 2D and 3D change maps for classifying “Remained Building” objects (after first level change analysis) into “Only 2D Change”, “Only 3D Change”, “2D with 3D Change” or “No Change” in the second level change analysis.

2 Methodology

Figure (1) shows the proposed method for building change analysis. Accordingly, buildings are extracted on both dates and, then, change analysis is done for each building object. Details of each step are given below.

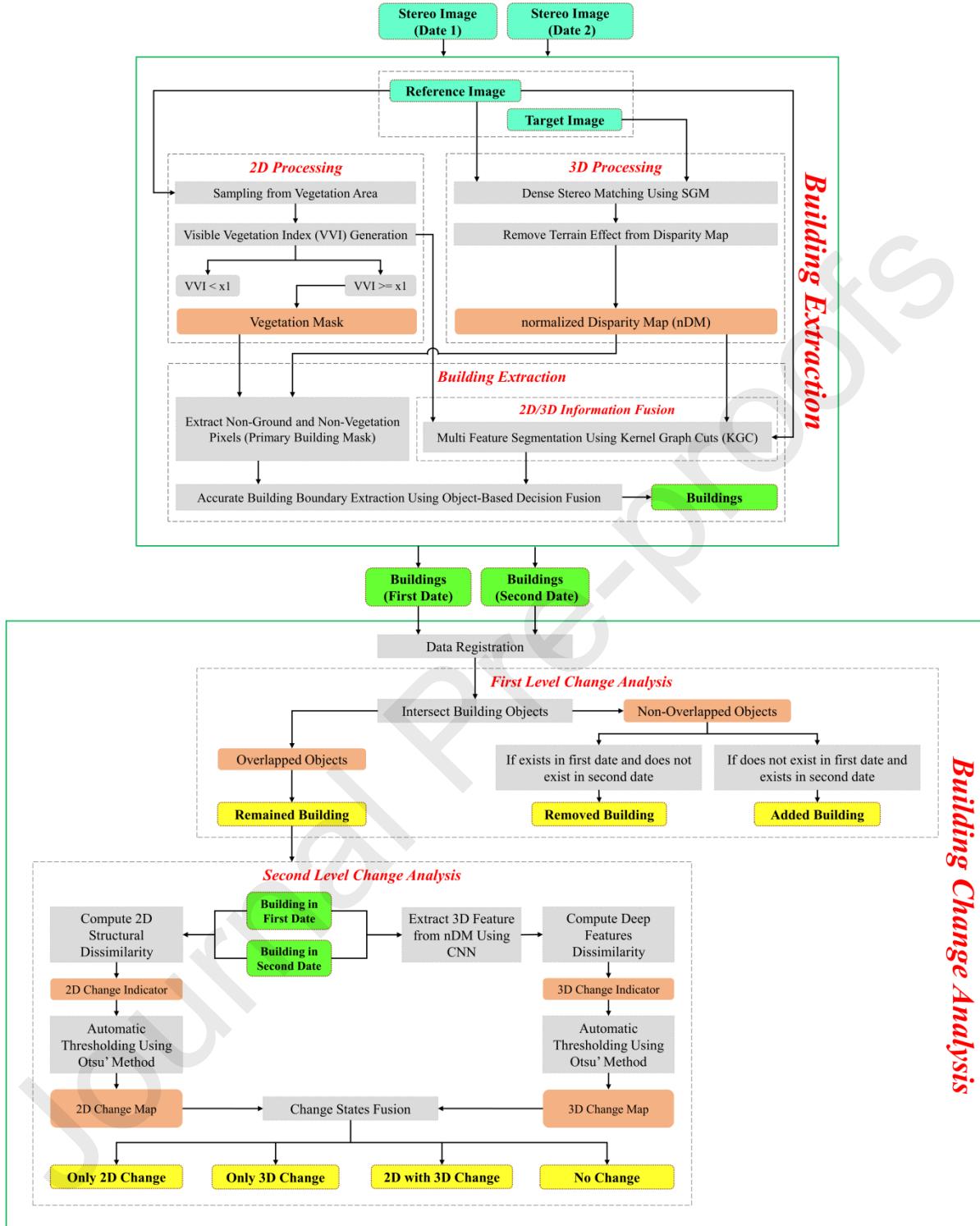


Figure 1: The proposed method for building change analysis using 2D and 3D information of high resolution satellite stereo images.

2.1 Building Extraction

Three main steps are applied in this paper for building extraction: detecting tree pixels as 2D processing using VVI, generating disparity map using semi-global matching (SGM) and producing the normalized disparity map (nDM) as 3D processing and, finally, simultaneously using 2D and 3D information for accurate building boundary extraction. Details of these steps are given below. This strategy for building extraction is based on previous authors' works (Mohammadi et al., 2019). Details of the mentioned steps are discussed below.

2.1.1 2D processing: Detecting tree pixels

The trees are elevated objects as buildings. Conventionally, they grow nearby the buildings in the urban area and have the same altitude as buildings. So, it is necessary to detect and remove them from the building extraction process. If NIR spectral band is available, tree and vegetation pixels can be identified by producing and thresholding the normalized vegetation index (NDVI). But, in cases, in which only RGB spectral bands are available, it is proposed to use visible vegetation index (PHL, 2018). In this research, VVI is used for detecting vegetation pixels from high resolution satellite images. The VVI is computed based on Eq. (1).

$$\text{VVI} = \left[\left(1 - \left| \frac{R - R_0}{R + R_0} \right| \right) \times \left(1 - \left| \frac{G - G_0}{G + G_0} \right| \right) \times \left(1 - \left| \frac{B - B_0}{B + B_0} \right| \right) \right]^{1/w}, \quad (1)$$

The values of $R_0 = 35$, $G_0 = 50$ and $B_0 = 51$ (for the first date) and $R_0 = 38$, $G_0 = 55$ and $B_0 = 57$ (for the second date) are constant. The mentioned values are used in this paper for the test area. To assign the values for these parameters, some vegetation areas in the input images are employed for sampling and averaging. Furthermore, $w = 1$ is determined by trial-and-error. After computing the VVI, a vegetation mask is generated by thresholding the product. The Otsu's method (Otsu, 1979) is used to automatically determine the optimal threshold value.

2.1.2 3D processing: Generating normalized disparity map

The inputs are stereo pair images; so, SGM (Hirschmuller, 2007) is used to produce a disparity map as 3D feature layer. In order to prepare the disparity map for use in the building extraction process, it is necessary to eliminate the terrain slope effect. Therefore, a normalized disparity map (nDM) will be produced. The goal of nDM production is to convert a disparity map into a bare-earth disparity map. The method used for this, is based on the algorithm proposed in (Krauß, 2018) that generates DTM from DSM. The main advantage of this approach is the possibility to use it with incomplete DSMs containing no data values which can be, for example, occlusions in the calculated DSM. This study uses the proposed method for similarly no-filled disparity map. Subsequently, the bare-earth disparity map is removed from the disparity map to leave only non-ground pixels. This eliminates the effect of ground slope from disparity map for non-flat areas and nDM will be able to separate non-ground pixels from ground pixels in the building extraction process. Actually, this product is equivalent to the normalized digital surface model (nDSM) in similar works.

2.1.3 Extract accurate building boundary

By generating the VVI (2D feature) and nDM (3D feature), building objects could be coarsely identified. For this, vegetation mask intersected with nDM and all pixels that are labeled as vegetation are transformed into ground pixels in the nDM. This causes the trees to be removed from the nDM. Therefore, all the pixels with non-ground label in nDM produce a primary building mask.

This primary building mask is not used as the final result because of uncertainties in the VVI and nDM. In fact, in addition to defects in nDM and, so, in the primary building mask, this product

does not have any geometric constraints. Thus, in order to add this constraint, a segmentation using kernel graph cuts (KGC) algorithm is employed.

In fact, the KGC segmentation has the ability to produce building segments with accurate boundaries by simultaneously using 2D and 3D information. To this end, the VVI, the nDM and the image channels (RGB bands) are the inputs of KGC for producing better image segments. The main reason for using KGC is that the basis of graph-cuts is separating the background (non-building) from the foreground (building). Graph-cut is one of the categories of normalized cuts in which the segmentation is based on the discontinuities in the input data. When 2D and 3D feature layers are used as KGC inputs, discontinuities become specifically prominent at the boundaries of buildings. This method also aims at optimal splitting by reducing the number of regions. It is based on the graph theory, in which each pixel is a vertex and edges link nearby pixels (4-connectivity) (Zaitoun and Aqel, 2015).

The segmentation function can be defined as Eq. (2). Let $I : p \in \Omega \subset \mathbb{R}^2 \rightarrow I_p = I(p) \in \Gamma$ be an image function from a positional array Ω to a space Γ of photometric variables such as intensity, disparities, and color or texture vectors. Actually in this paper, I is a combined data consisting VVI, nDM and the RGB values. For transforming the input data, the KGC uses kernels (Salah et al., 2011). If parameters of kernel function are set accurately, better results can be produced (Hearst et al., 1998).

$$F_k(\{\mu_l\}, \lambda) = \sum_{l \in \Gamma} \sum_{p \in R_l} (K(I_p, I_p) + K(\mu_l, \mu_l) - 2 \times K(I_p, \mu_l)) + \sum_{\{p, q\} \in N} \min (const^2, |\mu_{\lambda(p)} - \mu_{\lambda(q)}|^2), \quad (2)$$

where λ is a mapping function that indicates the label of each pixel, K means the kernel function, μ_l is the representative of the region R_l (e.g. the average of the intensity values of the pixels in

cluster l), $const$ is constant and N is a neighbourhood set containing all pairs of neighbouring pixels p and q . The *RBF* kernel is used in this paper via setting $\sigma = 0.5$ (Eq. (3)).

$$K(y, z) = \exp\left(-\frac{\|y - z\|^2}{\sigma^2}\right), \quad (3)$$

The goal after KGC segmentation is to detect building objects and replace them in the primary building mask. Eq. (4) is used in this step between objects in the primary building mask (after regrouping pixels) and segments in KGC result.

$$\text{if } \left(\frac{\text{Area}_{\cap}}{\text{Area}_{\text{segment}}}\right) > T \Rightarrow \text{Label segment as building segment and replace}, \quad (4)$$

In Eq. (4), Area_{\cap} specifies the number of pixels overlapped between the segment and the corresponding object in the building's primary mask and $\text{Area}_{\text{segment}}$ determines the number of pixels for any segment. As mentioned, because of using 2D and 3D information simultaneously, the boundary of these segments is more accurate than the corresponding objects in the primary building mask. Finally, each detected building object in the segmentation output is replaced in the primary building mask.

The proposed building extraction methodology is applied for stereo pair images on the first and second dates. So, building objects are extracted and are ready to enter the change analysis process. Indeed, this can be categorized as “post-classification” methodology that is mentioned in the section (1).

2.2 Building Change Analysis

The goal of building change analysis is to recognize the change state of the detected buildings on the first or second dates. So, the building objects are labeled in three main categories: The “Removed Building”, the “Added Building” and the “Remained Building”. This step is considered as the first level change analysis. Each building detected on the first or second

dates will fall into one of these three categories. It is clear that “Removed Building” and “Added Building” labels detect building changes for the users, but no more information is achieved about their 2D or 3D changes. Thus, it is proposed to have the “Remained Building” label for analyzing the other aspect of change state.

The “Remained Building” objects are categorized into four change states including “Only 2D Change”, “Only 3D Change”, “2D with 3D Change” and “No Change”. This stage is considered as the second level change analysis. In this category, 2D change means change in pixel DN values such as change for roof color from the first to second dates. Also, 2D change can be found as change in textural information. Likewise, 3D change means change in disparity values or, in reality, height values from the first to second dates. It should be noted that, if at the end of the second level change analysis, an “Only 2D Change” label is given to a building; it means that no 3D change is identified for the object. Similarly, if an “Only 3D Change” label is given to a building, it means that no 2D change is detected. It is therefore clear that, for buildings labeled “2D with 3D Change”, both 2D change and 3D change information is identified. By doing this analysis, other aspects of changes will also be revealed, which can provide other valuable information for urban executives. It is obvious that, for “Removed Building” and “Added Building” objects, both 2D and 3D changes occur and more details about the type of changes are not necessary. Fig. (2) shows how different change states are defined in this paper.

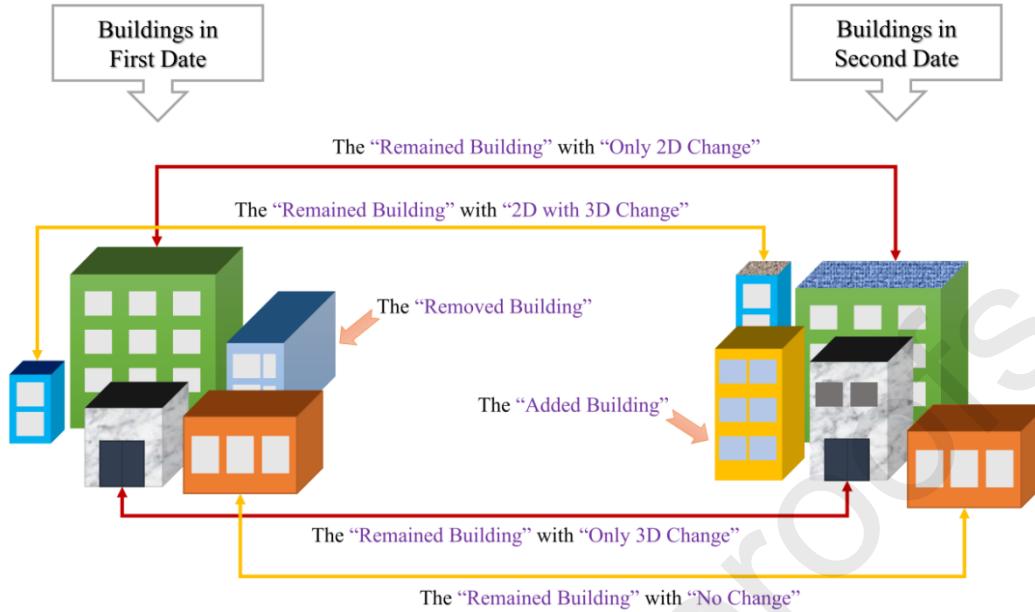


Figure 2: Different building change states in change analysis strategy.

To perform the change analysis, some steps must be applied. Details of these steps are discussed below. The following steps are shown in Fig. (1), in the section on building change analysis.

2.2.1 Data registration

It is necessary to first register the data of two dates. It is proposed to do data registration in image space, except doing it in the object space. Registering data in object space is not proposed due to reasons including: imposing more processing such as producing the DSM using space intersection techniques by rational polynomial coefficients (RPCs), need to generate normalized digital surface model (nDSM), instead of using the produced nDM, need for ground control points (GCPs) and, finally, low quality of the ortho image, which can reduce the accuracy of 2D change analysis. But, there is a challenge for registering data in the image space. Because of distortions caused by satellite tilt angle and relief displacements in the images and disparities, more powerful data registration algorithms should be used. So, the following solution is suggested for data registration in this paper.

The interest points are detected in the two images using SIFT and RANSAC algorithms. The point extraction process is done in grid base strategy to extract key points from almost all regions of images. These key points are then matched by comparing the extracted SIFT descriptors. In the next step, first, the nDMs on both dates are co-registered in the vertical dimension by comparing the nDM values at the matched points. To this end, a linear equation is used (Eq. (5)).

$$nDM_2 = a_0 + a_1 \times nDM_1, \quad (5)$$

where nDM_1 is nDM value on the first date and nDM_2 is the corresponding nDM value on the second date. a_0 and a_1 are the coefficients that are calculated using least square adjustment method (Wolf and Ghilani, 1997). By doing so, new nDM values are computed and replaced in the first date data.

In the next step, B-Spline registration (Zufeng et al., 2016) algorithm is used for the co-registration of images and nDMs in the horizontal dimension. A grid of B-spline control points is constructed which controls the transformation of the input data. An error measure is used to measure the registration error between the moving and static images. The quasi Newton optimizer is used to move the control points to achieve the optimal registration between both data with minimal registration error.

2.2.2 First level change analysis

For the first level change analysis, the existence or absence of a building is investigated based on object-based analysis. Three rules are defined and used in this step according to what is given in Table (1). To find the nearest building object, distances between center of mass of the target building on the one date with center of mass of buildings on another date are compared. By applying these rules to any building object on the first or second dates, the first level change analysis is finished.

Table 1: The defined rules for determining the first level change state of each building object.

Rule	Antecedence	Consequence
1	If the building on the first date does not intersect with the nearest building object on the second date	Then, the building is labeled as “Removed Building”
2	If the building on the second date does not intersect with the nearest building object on the first date	Then, the building is labeled as “Added Building”
3	If the building on the first date intersects with the nearest building object on the second date	Then, the building is labeled as “Remained Building”

2.2.3 Second level change analysis

For the second level change analysis, 2D and 3D change indicators are generated. Finally, these change indicators are combined to determine the second level change state of each “Remained Building” object. The following paragraphs describe the details of this step.

2.2.3.1 2D change map generation using structural similarity

The structural similarity (SSIM) index (Wang et al., 2004) is calculated based on the three terms. These three terms are referred to as the luminance term, the structural term and the contrast term. The index is a combination of the three terms (Eqs. (6) to (9)).

$$SIM(I_1, I_2) = [l(I_1, I_2)]^\alpha \times [c(I_1, I_2)]^\beta \times [s(I_1, I_2)]^\gamma, \quad (6)$$

$$l(I_1, I_2) = \frac{2\mu_{I_1}\mu_{I_2} + C_1}{\mu_{I_1}^2 + \mu_{I_2}^2 + C_1}, \quad (7)$$

$$c(I_1, I_2) = \frac{2\sigma_{I_1}\sigma_{I_2} + C_2}{\sigma_{I_1}^2 + \sigma_{I_2}^2 + C_2}, \quad (8)$$

$$s(I_1, I_2) = \frac{\sigma_{I_1I_2} + C_3}{\sigma_{I_1}\sigma_{I_2} + C_3}, \quad (9)$$

where $l(I_1, I_2)$ is luminance term, $c(I_1, I_2)$ is contrast term and $s(I_1, I_2)$ is the structural term.

Also, μ_{I_1} , μ_{I_2} are the local means, σ_{I_1} , σ_{I_2} are standard deviations and $\sigma_{I_1I_2}$ is cross-covariance for images I_1 and I_2 . If the defaults for exponents are used ($\alpha = \beta = \gamma = 1$) and $C_3 = C_2/2$ (default selection of C_3), the index is simplified as Eq. (10).

$$SSIM(I_1, I_2) = \frac{(2\mu_{I_1}\mu_{I_2} + C_1)(2\sigma_{I_1I_2} + C_2)}{(\mu_{I_1}^2 + \mu_{I_2}^2 + C_1)(\sigma_{I_1}^2 + \sigma_{I_2}^2 + C_2)}, \quad (10)$$

In this paper, $C_1 = (0.01 \times DR)^2$ and $C_2 = (0.03 \times DR)^2$ are used. The DR is dynamic range of the input image, specified as a positive scalar. For example, the default dynamic range is 255 for images of data type uint8. It is necessary to mention image patches limited to the bounding box of each “Remained Building” on the first and second dates are the inputs of SSIM. The higher value for SSIM equals lower 2D change for the input image patch so, “1 - SSIM” is used as the 2D change indicator. Finally, the Otsu’s method is employed for the automatic thresholding of 2D change indicator and a 2D change map is generated.

2.2.3.2 3D change map generation using CNN

In addition to the points outlined in section (1), there are several main reasons that why CNN is employed for feature extraction from 3D data in this paper. It is worth mentioning that the nDM is not used directly in the building extraction step because the disparity and nDM that are produced from HRSSI and dense stereo matching have uncertainties. Actually, the nDM is used only for primary building mask generation and, also, for helping the KGC generate precise building segments in the building extraction step. Furthermore, the nDM is used in 3D change analysis, but not the data itself. However, the features are extracted from it. The existence of blunder and miss-matched areas in the produced disparity and, then, the nDM might cause problems in the accurate comparison of data and may present false 3D changes. But, CNN can reduce these challenges by summarizing 3D data in the form of 3D features and allowing for the correct comparison of 3D information.

The easiest and maybe fastest way to use the representational power of the pre-trained CNN is feature extraction. Since feature extraction only requires one data pass, if you do not have a GPU

to accelerate network training with a lot of training data to learn it, this is a good starting point.

In this way, limitations of automation disappear and no supervised methods are needed.

Inception-ResNet-v2 (Szegedy et al., 2017) is a CNN which has been trained in more than one million images from the ImageNet database (2019). It can classify images into 1000 categories of objects, such as a mouse, keyboard, pencil, etc. As a result, the network learns rich feature representations for a wide range of images. The Inception-ResNet-v2 has more accurate architecture than the previous state of the art CNNs, like Inception V3 and ResNet 152. It has better Top-1 and Top-5 validation accuracies on the ILSVRC 2012 image classification benchmark. These validations are based on single crop of the input data (Posted by Alex Alemi, 2019). This CNN model can be thought as a combination of two components: feature extraction part and classification part. The convolution plus pooling layers perform feature extraction. Researchers are dealing with very powerful and efficient tools which perform automatic feature extraction to achieve better accuracy and, in this paper, this capability is used for feature extraction from 3D data that have some uncertainties; it cannot be used directly for 3D comparison and needs to be summarized. Fig. (3) shows the schematic diagram and the idea of using Inception-ResNet-v2 for feature extraction. The network creates a hierarchical representation of the input data. Deeper layers contain higher-level features. These higher-level features are constructed by the lower-level features in earlier layers of network. To get the data features, activations on the global pooling layer, “AvgPool” are used at the end of the network. The global pooling layers pool the input features over all spatial locations, giving 1536 features finally.

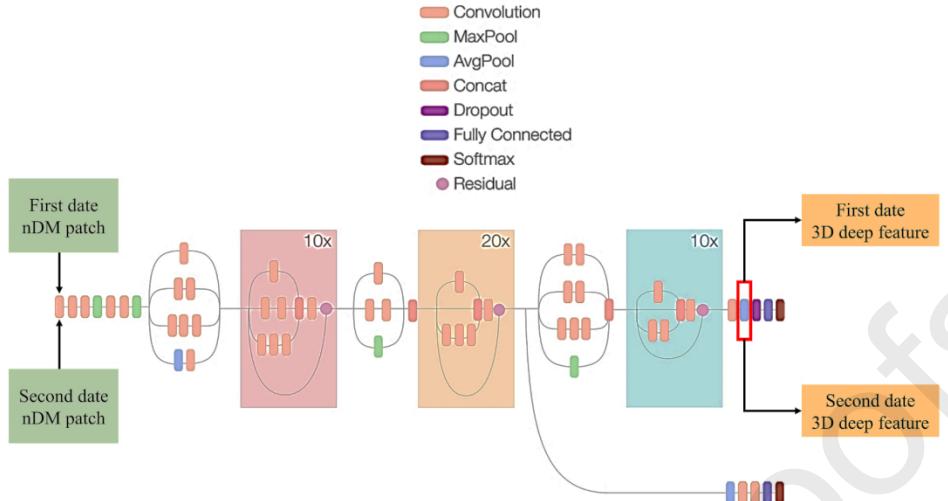


Figure 3: The compressed schematic diagram of Inception-ResNet-v2 (Posted by Alex Alemi, 2019) and idea for using network to extract deep features. The output of “AvgPool” layer (Red bounding box at the end of the network) is used as 3D feature.

In this paper, nDM patches limited to the bounding box of each “Remained Building” on the first and second dates are the inputs of Inception-ResNet-v2 network. Input data size to enter the network is 299-by-299, but the building objects have different sizes. To automatically resize the data before being the input to the network, the augmented image is created. In order to prepare the nDM data to resemble an image and allow it to enter the CNN model, the values of nDM data available at each patch are mapped from 0 into 255 as 8bit image data, linearly. It should be noted that, in order to avoid creating inaccurate features and so false 3D change alarms, it is necessary to calculate the minimum and maximum nDM values on both dates, first, and then, calculate the final minimum and maximum values and use them to map the two date patches from 0 into 255. Thus, a feature vector with 1536 members is calculated for each nDM patch on the first date. Also, another feature vector with 1536 members for the same nDM patch on the second date is computed. It is necessary to mention, all nDM patches are resized to 299-by-299-by-3 to enter the Inception-ResNet-v2 (nDM patch is replaced in second and third channels of input data to generate semi-RGB image).

The deep features extracted from the first and second dates for each building object must be compared to determine the 3D change of each object. In this case, normalized cross correlation (NCC) is employed to calculate feature similarity and can be used in the 3D change indicator. Therefore, a value between 0 and 1 is computed for each "Remained Building". Indeed, this comparison is a 3D comparison because of the type of input data for CNN. Clearly, the higher value for the NCC in this case equals the lower 3D change aspect and so "1-NCC" is utilized as the 3D change indicator. Finally, the Otsu's method is used for automatic thresholding of 3D change indicator and a 3D change map is generated.

2.2.3.3 2D/3D change map fusion for the second level change analysis

The 2D and 3D change map results in this stage are combined to determine the final change state of each "Remained Building" object. Therefore, each "Remained Building" object can be categorized in one of the "Only 2D Change", "Only 3D Change", "2D with 3D Change" and "No Change" labels by applying the defined rules in Table (2).

Table 2: The defined rules for determining the final change state of each "Remained Building" object as the second level change analysis.

Rule	Antecedence	Consequence
1	If the "Remained Building" exists in 2D change map and does not exist in 3D change map	Then, the change state is "Only 2D Change"
2	If the "Remained Building" does not exist in 2D change map and exists in 3D change map	Then, the change state is "Only 3D Change"
3	If the "Remained Building" exists in 2D change map and also exists in 3D change map	Then, the change state is "2D with 3D Change"
4	If the "Remained Building" does not exist in 2D change map and also does not exist in 3D change map	Then, the change state is "No Change"

This change information is useful for city administrators to plan for reducing city problems.

Finally, a change analysis map is created that includes building objects with the labels listed in Fig. (2).

3 Results and Discussions

For evaluating the results of the methodology, a test area in Tehran, Iran, with two datasets of HRSSI is selected.

3.1 Dataset

It is clear that, the evaluation data must be complete, complex, and applicable. In Fig. (4), it can be observed that there are buildings with different geometrical structures and very different heights in the study area, which can indicate the complexity of the study area. There are also, all the change classes mentioned in the change analysis steps, which shows the completeness of the study area. The study area is also, one of the areas where urban managers are interested to monitor because of high rate changes and so the applicability of the data is determined.

There are two stereo pair images acquired by high resolution satellite GeoEye-1. This satellite has 0.5-meter resolution at nadir. The first stereo pair image was acquired in 2009. The second stereo pair image was acquired in 2013 (Fig. (4)). Figs. 4(a) and 4(b) show the input data. To evaluate the efficiency of the building extraction process, building objects are labeled by an expert on the first and second dates manually as the ground truth data (Figs. 4(d) and 4(e)). Moreover, for evaluating the change analysis step, the labels of “Removed Building”, “Added Building”, “Remained Building”, “Only 2D Change”, “Only 3D Change”, “2D with 3D Change” and “No Change” is assigned to the buildings manually (Fig. 4(c)). As can be observed, there are a number of high building that are constructed in the test area and can be good targets to evaluate the proposed method.

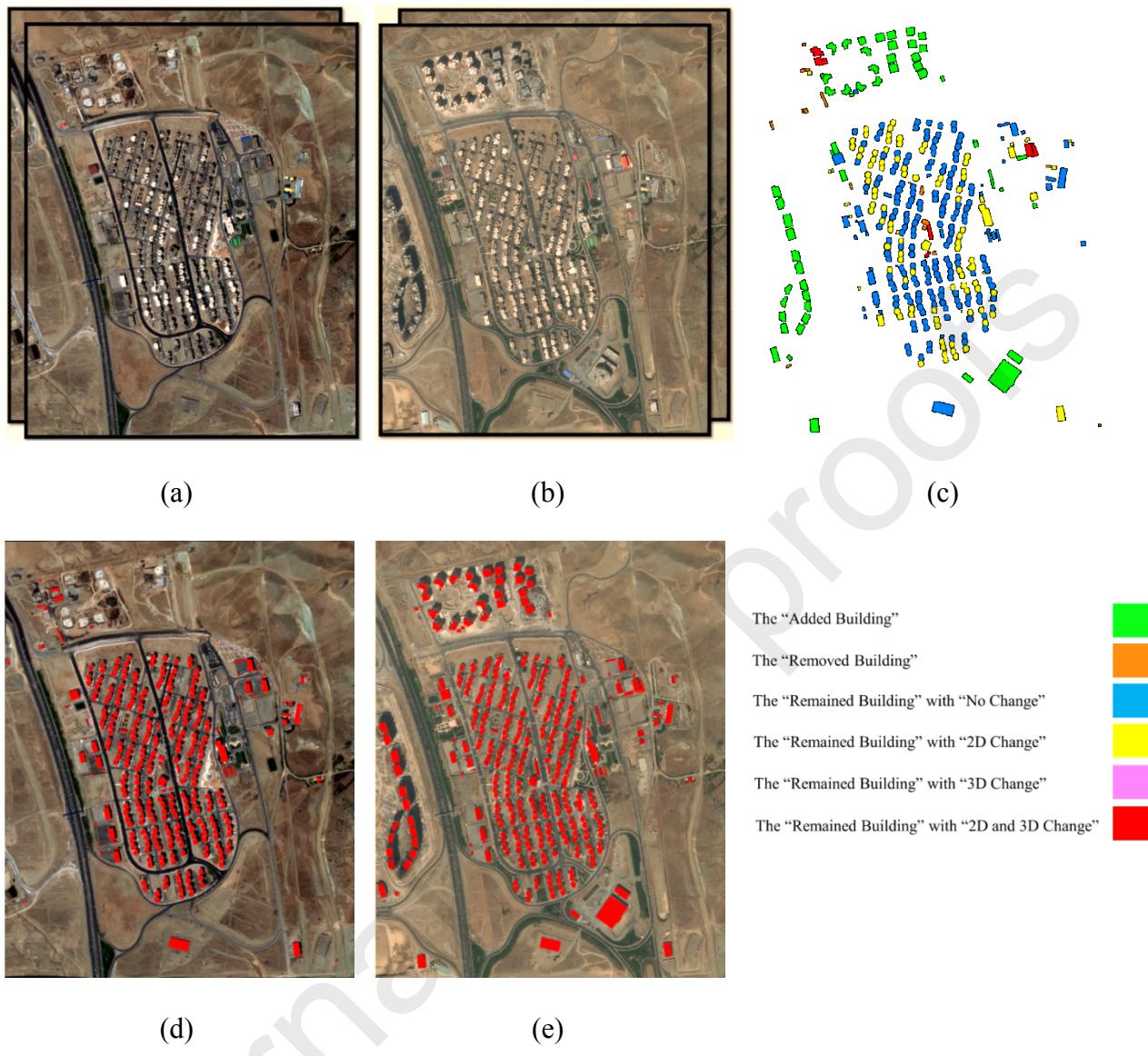


Figure 4: Input dataset: (a) Stereo pair images in 2009, (b) Stereo pair images in 2013, (c) Building change states that are labeled manually as ground truth data, (d) Ground truth data for building objects that is generated manually for 2009 and (e) Ground truth data for building objects that is created manually for 2013.

3.2 Implementation and evaluation

3.2.1 Results of 2D processing

In the building extraction step, VVI index and vegetation mask are generated for two dates as 2D processing (Fig. (5)). The VVI has a number of uncertainties: 1) higher value in the shadow area, 2) higher value in regions with color content the same as vegetation pixels and 3) lack of uniformity in vegetation areas. Regarding the first challenge, it should be noted that actually it will be good because buildings are the target objects not the shadows and removing shadows can also remove unnecessary data around buildings. For the second challenge, it should be mentioned that, there is no such regions in the test area, however, it is necessary to suggest other steps as one of the future works of this research if such color content exist. Finally, given the third challenge, it can be stated that, the vegetation mask is used for summarizing the nDM and the next steps of the proposed method will eliminate the shortcomings of this data.

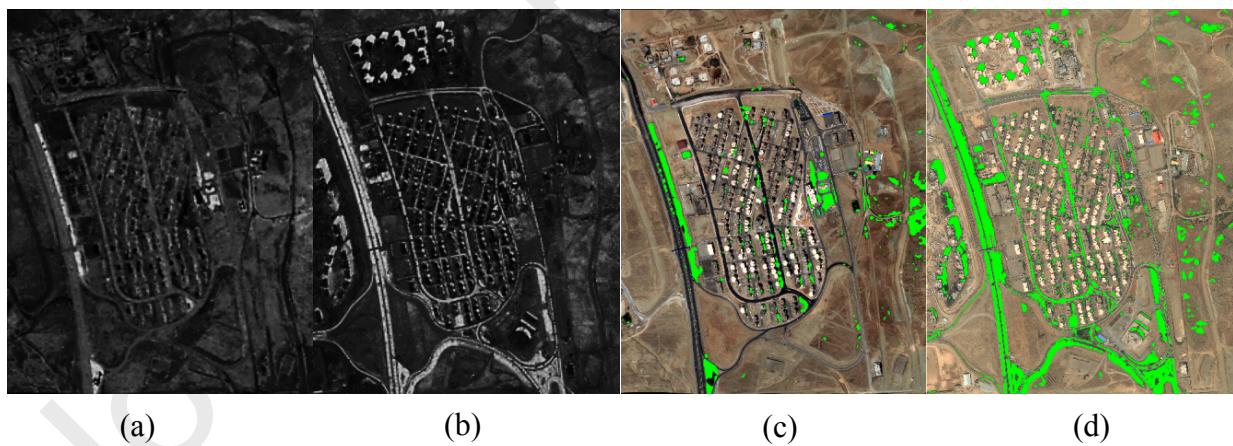


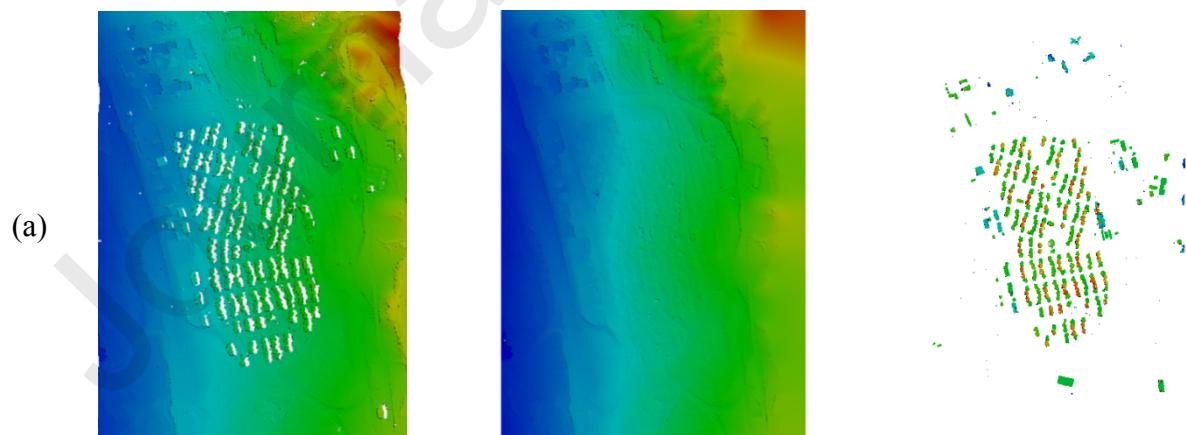
Figure 5: VVI index and vegetation mask for the first and second dates: (a) VVI index in 2009, (b) VVI index in 2013, (c) Vegetation mask after automatic thresholding of VVI in 2009 and (d) Vegetation mask after automatic thresholding of VVI in 2013.

Despite all these challenges, as shown in Fig. (5), if there is no NIR band, the VVI will be a good option for detecting vegetation in visible images. The importance of using such a solution is to

know that in order to detect buildings as elevation objects, other elevation objects must be somehow removed from the final result. The Otsu threshold value is calculated for first date is 0.43 and 0.46 for the second date.

3.2.2 Results of 3D processing

Fig. (6) represents, respectively, the disparity map, the bare-earth and the produced nDM for the first and second dates. As can be observed, some artifacts exist in the products. Also, 3D image objects do not have sharp boundaries in these products. These challenges make it impossible to use the disparity map alone to extract the buildings as accurate as possible. Also, methods that use the DSM differencing or disparity differencing to detect changes in building objects, will actually provide low accuracy. This important point has led to the use of a series of different solutions in this research to repair the produced disparity map in combination with the 2D information of the input images. The generation of nDM is also necessary because they are used for separating ground pixels from the non-grounds; so, they must be free from any terrain slope effects.



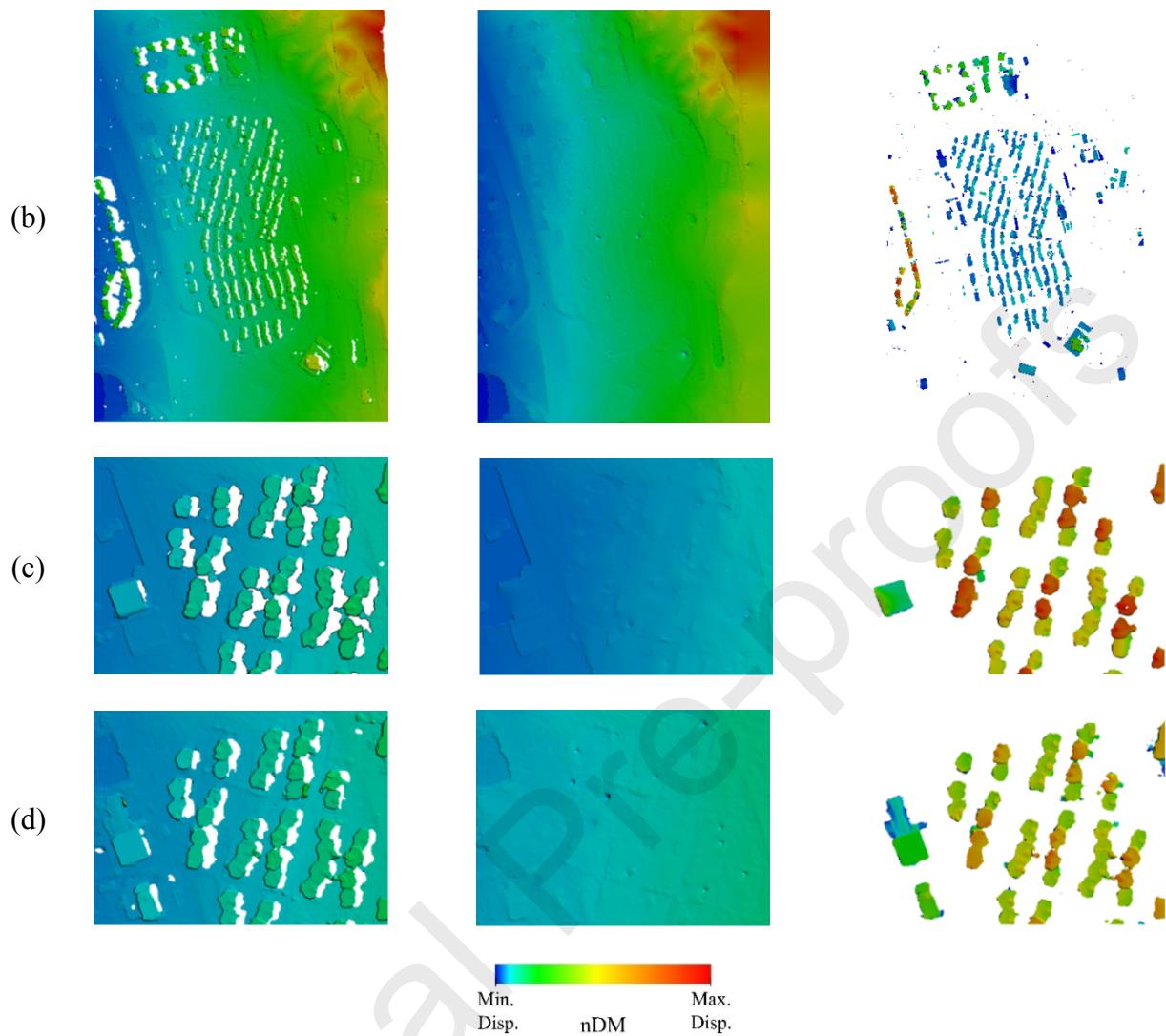
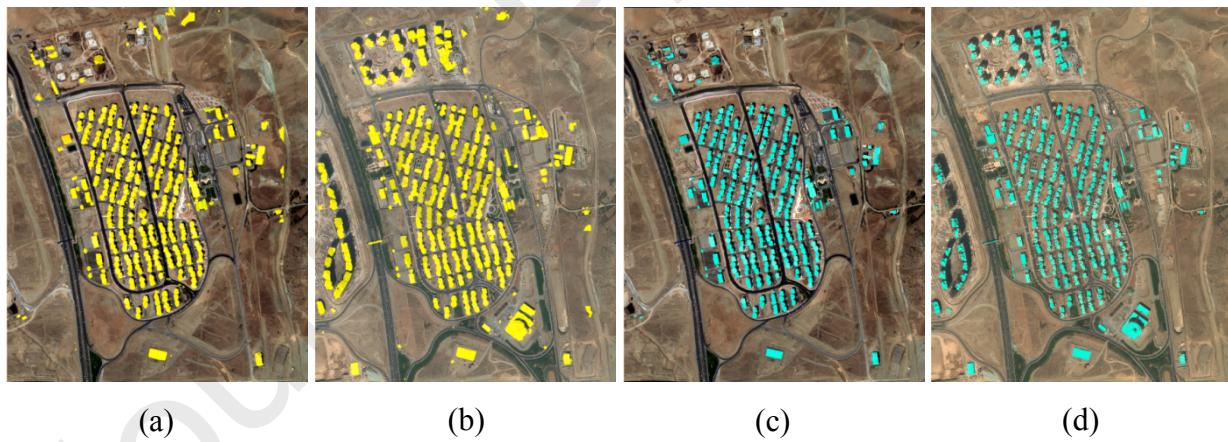


Figure 6: Disparity map and the produced nDM for the first and second dates: (a) Disparity map in 2009, Extracted bare-earth in 2009 and the produced nDM after removing bare-earth from disparity map in 2009 (left to right), (b) The disparity map, the bare-earth and the nDM in 2013 (c) Sample regions in 2009 and (d) Sample regions in 2013.

3.2.3 Results of accurate building boundary extraction

By generating nDM, pixels with nDM value of higher than 0 are labeled non-grounds. After removing vegetation and, actually, tree pixels (resulted in vegetation mask) from non-ground pixels, the primary building mask is produced. It is necessary to mention that using 0 values for nDM thresholding is acceptable with high probability in this research because the main purpose

of this step is to produce a primary building mask. Figs. 7(a) and 7(b) and also Figs. 7(e) and 7(f) show the outputs of primary building masks for the first and second dates. As can be observed, many undershoot and overshoot errors are visible in the primary building mask. Actually, the primary building mask is impossible to present the exact boundaries of the buildings; it is needed to use more sophisticated methods to extract accurate building boundaries. Once again, it should be stated that, given that the defects in the building extraction step can reduce the accuracy of the change analysis step, it is necessary to be sensitive to the exact extraction of the boundaries of buildings. Dissatisfaction with the results of the primary building mask caused the proposed method to take other steps to improve the boundaries of buildings. As mentioned in the methodology, it seems that combining with 2D information (which has precise boundary) can help to improve the quality of extracted building objects. Thus, a segmentation using KGC algorithm was employed.



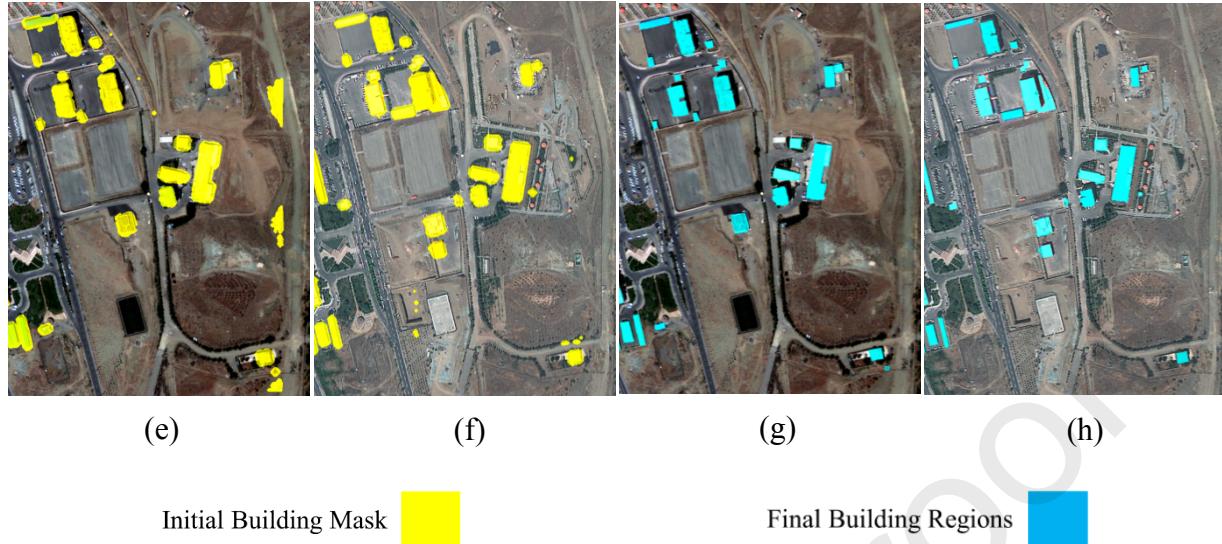


Figure 7: The primary building mask and the final building regions (after using KGC segmentation results): (a) The primary building mask in 2009, (b) The primary building mask in 2013, (c) Final building regions in 2009 and (d) Final building regions in 2013, (e) to (h) Sample regions from the first row.

The KGC is capable of producing good segments in building regions by using 2D and 3D features as the input data. These segments are intersected with primary building mask object after re-grouping it based on the defined rule in Eq. (3). A threshold equal to 0.8 (because the segments are not ideal) is used practically to detect building segments that correspond to the primary building mask objects. By replacing segmentation objects in the primary building mask, more sharp and accurate building boundaries are achieved (Figs. 7(c), 7(d), 7(g) and 7(h)). As can be seen in the Figs. 7(g) and 7(h), the boundaries of all the buildings have been precisely extracted. Also, the areas that were mistakenly entered in the primary building mask (Fig 7(e)) have been removed. Although the implementation of this step has increased the computational cost but it helps to reduce the uncertainties of building change detection step.

3.2.4 Evaluation of building extraction step

The results of different steps in the proposed method can be regarded as a special classification output. For the first evaluation, the output of building extraction step is compared with the manually labelled ground truth data. The building mask on the first and second dates, actually classify image pixels into ground and non-ground classes. The objective metrics including “Accuracy”, “Sensitivity”, “Specificity”, “Precision”, “F1-Score” and “False Positive Rate (FPR)” (Wikipedia, 2019) are used to provide quantitative evaluation of the results (Eq. (11)). In Eq. (11), TP (True Positive) is the number of predicted building pixels that are correctly classified, TN (True Negative) is the number of predicted non-building pixels that are correctly classified, FP (False Positive) is the number of predicted building pixels that are actually non-building pixels and FN (False Negative) is the number of predicted non-building pixels that are actually building pixels. Also, P is the number of building pixels and N is the number of non-building pixels. Table (3) shows the quantitative assessment of the building extraction step for the first and second dates.

$$\begin{aligned}
 \text{Accuracy} &= (TP + TN)/(P + N) \\
 \text{Sensitivity} &= TP/(TP + FN) \\
 \text{Specificity} &= TN/(FP + TN) \\
 \text{Precision} &= TP/(TP + FP) \\
 \text{F1 - Score} &= 2 \times TP/(2 \times TP + FP + FN) \\
 \text{False Positive Rate (FPR)} &= 2 \times TP/(2 \times TP + FP + FN)
 \end{aligned} \tag{11}$$

Table 3: Quantitative assessment of the building extraction step for the first and second dates.

Product	Accuracy	Sensitivity	Specificity	Precision	F1 Score	FPR
Buildings (First date)	0.9869	0.9571	0.9884	0.8061	0.8752	0.0116
Buildings (Second date)	0.9835	0.8293	0.9948	0.9208	0.8726	0.0052

The accuracy of higher than 98% and F1-Score of higher than 87% in both first and second dates confirm that the proposed method for building extraction is able to extract building objects accurately. It is obvious that any defects in the output of this step may cause low accuracy in the change analysis step. The most challenging issue is losing building objects on the first or second dates. Based on the low values for FPR in Table (3), it can be confirmed that the methodology is able to detect almost all buildings. The close values of parameters for the first and second dates indicate that the proposed method for building extraction is reproducible. Also it can be observed that the value of the sensitivity is reduced on the second date compared to the first date. The main reason for this decrease is the addition of high buildings on the second date (Fig. (4)). These high buildings have different geometric shapes on their roofs. The presences of these different geometric shapes causes the spectral and depth discontinuities and in the step of KGC segmentation, separate segments for one building object are produced. Therefore, some of these segments may be discarded during the building extraction process. This can increase the amount of FNR (False Negative Rate) and thus reduce the value of sensitivity parameter. Actually, the presence of high buildings in the study area with different geometric shapes on their roofs is one of the limitations for achieving the best results.

3.2.5 Results of the first level change analysis

In the data registration step, grid-based SIFT key points are detected and matched. The registration is first applied for nDMs. The values of $a_0 = 27$ and $a_1 = 1.0002$ are computed for registering nDM values. This registry is not pixel-by-pixel; but it is accurate enough because the purpose of this step is to match the data to compare correspondence objects in the change analysis step.

The results of the first level change analysis are shown in Fig. (8). The green regions are buildings with the “Added Building” label. The orange regions are buildings with the “Removed Building” label and the blue regions are buildings with the “Remained Building” label. This is actually building change detection result and will be a good map for the urban managers.

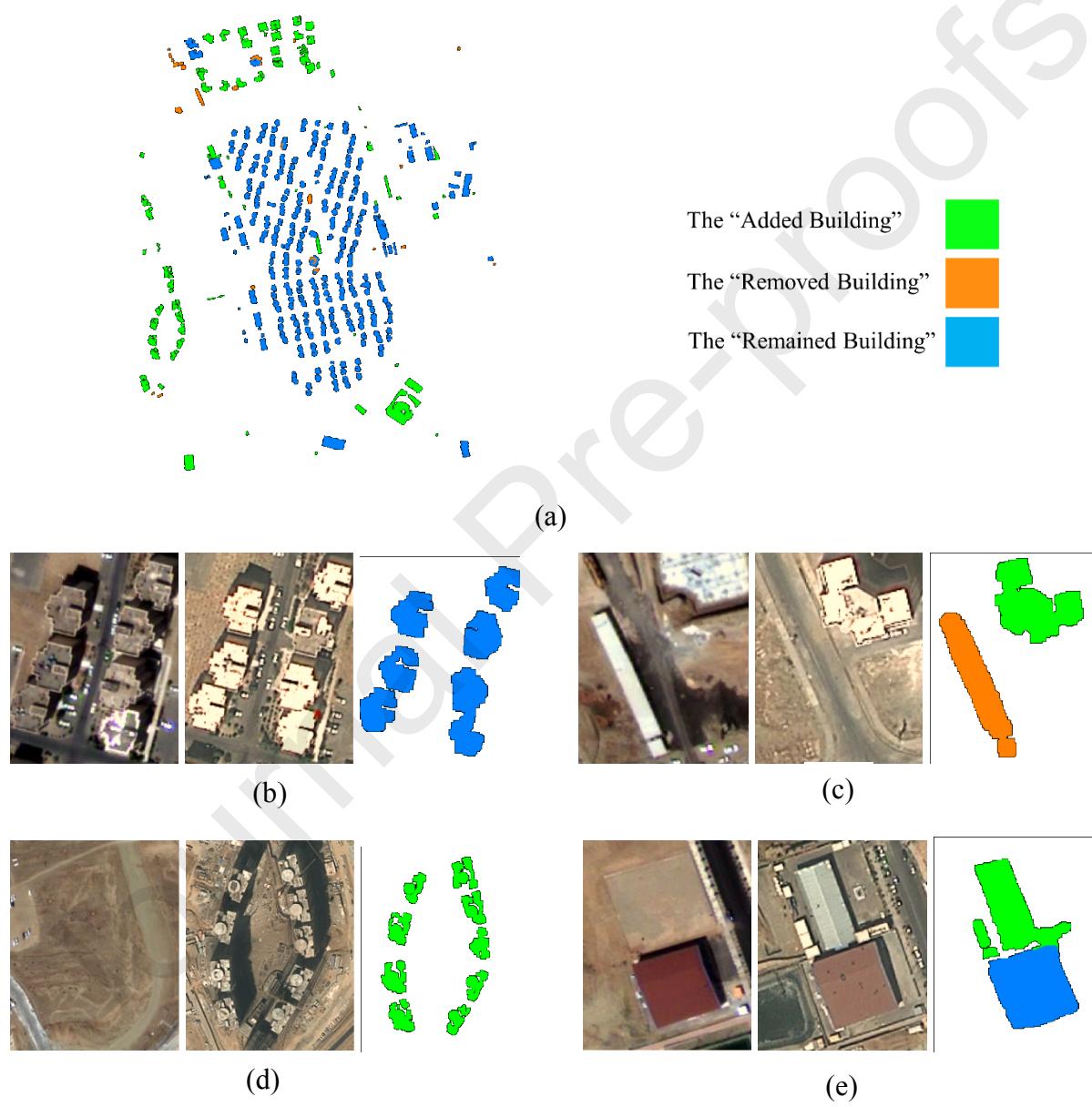


Figure 8: Outputs of the first level change analysis: (a) First level change states for entire test area and (b) to (e) some samples of test area including: the image patch in 2009, the image patch in 2013 and the first level change states (left to right).

The main obstacle for achieving high precision of this stage is the existence of defects in the process of extracting the exact boundaries of buildings. For this reason, attempts are made to extract the boundaries of buildings very precisely in the earlier steps. Since good results have been obtained in the building extraction step, achieving acceptable accuracy in the results of this step has not been unexpected. Due to the use of object-based analysis, the radiometric differences observed on the first and second data were not effective in the first level change analysis (Fig. 8(b)).

3.2.6 Evaluation of the first level change analysis

For the quantitative assessment, the output of the first level change analysis is also compared with the ground truth data. In the first level change analysis map, three class of objects exist which include the “Removed Building”, the “Added Building” and the “Remained Building”. So this change map can be considered as the classification output. The contribution of each class in the study area for the ground truth and output data of the first level change analysis is given in Table (4).

Table 4: The contribution of “Removed Building”, “Added Building” and “Remained Building” in the study area for the ground truth and the output data of the first level change analysis.

Product	Removed Building	Added Building	Remained Building
First Level Ground Truth	4.4%	16.5%	79.1%
First Level Change Analysis	5.4%	20.0%	74.6%

Table (4) shows that the first level change analysis has acceptable results generally but some defects exist in the results. Actually, the percent of “Remained Building” objects is lower than the reality and this affects the output of the second level change analysis; but it can be acceptable. Table (5) shows the evaluation of the first level change analysis. The F1-Score of higher than 90% also shows the high accuracy for the first level change analysis.

Table 5: Quantitative assessment of the first level change analysis step for the first and second dates for each building label and overall.

Product	Accuracy	Sensitivity	Specificity	Precision	F1 Score	FPR
Removed building	0.851	0.851	0.994	0.683	0.758	0.006
Added building	0.980	0.980	0.990	0.975	0.977	0.01
Remained building	0.982	0.982	0.974	0.989	0.985	0.026
Overall	0.979	0.937	0.986	0.882	0.907	0.014

As can be seen in Table (5), the value of the precision metric for the "Removed Building" class is low. This means that the FP rate is high in this class. So, some non-building pixels in this category are classified as a "Removed Building". Due to the details of the change in the study area, most of these removed buildings are adjacent to the high buildings that were being constructed in the test area from first to second date. Actually, they are temporary buildings that have been built as construction site. These buildings are mostly low in height and small in size, and due to the spatial resolution of the input image, the probability of error in this type of building is higher. However, due to the lower contribution of these buildings in the test area, the overall accuracy has been affected slightly.

3.2.7 Comparison with other methods

To verify the effectiveness of the methodology, a comparison between the proposed method and the other two methods is performed. The first method is a DSM differencing (Murakami et al., 1999) and the second is the post-refinement method (Pang et al., 2018). In this comparison the accuracy of building change detection is evaluated. So the changed buildings are labeled as 1 and the non-changed buildings are labeled as 0. Table (6) shows the quantitative comparison between these methods.

Table 6: Quantitative comparison between the proposed method and two other methods.

Product	Accuracy	Sensitivity	Specificity	Precision	F1 Score	FPR
DSM differencing (Murakami et al., 1999)	0.696	0.695	0.743	0.992	0.817	0.257
Post-refinement (Pang et al., 2018)	0.891	0.898	0.567	0.989	0.941	0.432
Proposed	0.990	0.996	0.712	0.994	0.995	0.288

It can be seen that the proposed method has been able to achieve higher values in accuracy, sensitivity, precision and F1-score. Compared with the DSM differencing and post-refinement methods there are some improvements. The main improvement is in simultaneous use of 2D and 3D information which can extract the building objects accurately. However, the slight increase in the FPR compared to DSM differencing method indicates that the overshoot error still remain in the accurate building boundary extraction step.

3.2.8 Results of the second level change analysis

The objects with the “Remained Building” label enter the second level change analysis. Fig. 9(a) shows the 2D change indicator output. Figs. 9(c) to 9(f) demonstrate the results of 2D change indicator for different image patches. As can be observed, 2D change indicator is able to recognize spectral distinctions. These spectral distinctions may be due to different reflections of building objects between the first and second dates. Indeed, this kind of 2D change is not real, but is a candidate for the final 2D change occurrence, and an expert must decide about it. As mentioned in the methodology, change indicator values are in the range of (0, 1) and a threshold must be applied to detect 2D changed regions. The Otsu threshold value for this case is equal to 0.29. Fig. 9(b) shows the final 2D change map.

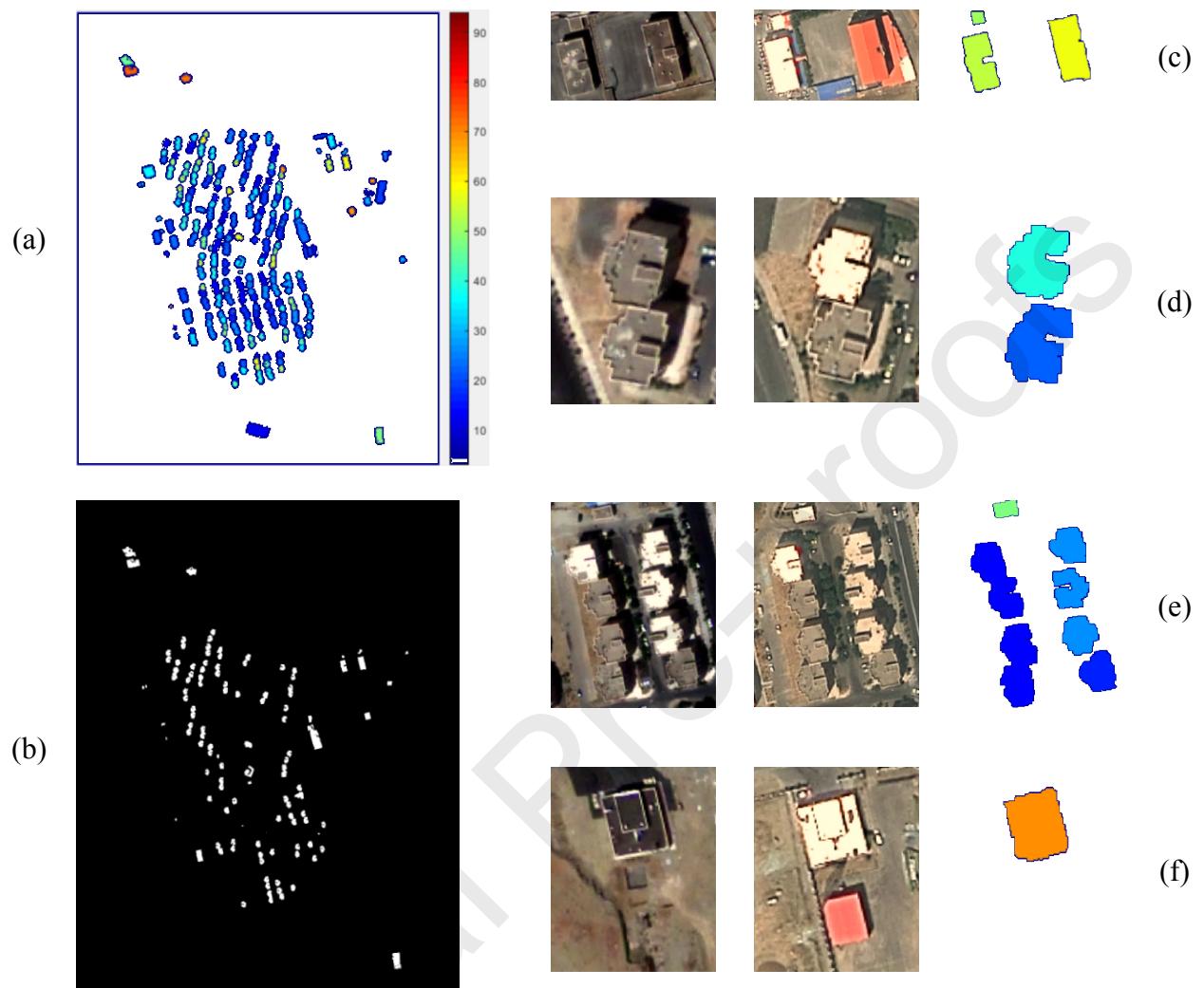


Figure 9: 2D change indicator output for the “Remained Building” objects: (a) 2D change indicator for entire image, (b) 2D changed objects after automatic thresholding of change indicator and (c) to (f) Some samples of test area including: the image patch in 2009, the image patch in 2013 and the 2D change indicator (left to right). In Fig. 10(a), the output of 3D change indicator is given for the entire test area. Fig. 10(b) shows the 3D changed object after automatic thresholding of 3D change indicator. Figs. 10(c) to 10(f) represent the output of 3D change indicator for some nDM patches on the first and second dates. It can be observed that 3D change indicator is able to perfectly compute 3D change aspects of building objects. Furthermore, no information about nDM differences is needed to separate the

changed from the not-changed objects. This is one of the advantages of the proposed method for detecting 3D changes. As mentioned earlier, due to the defects in the nDM, in case of using differential methods, the uncertainties in the results of 3D change detection can be predicted. It is also clear that changes in the shape and height of buildings, due to differences in disparity values, cause differences in the calculated similarity criterion based on the feature vector obtained from the deep neural network. As a result, the need for complex object-based and elevation-based analyzes is reduced, which means less human intervention and increased automation. The Otsu threshold value in this case equals 0.51 to separate buildings with 3D change from other buildings.

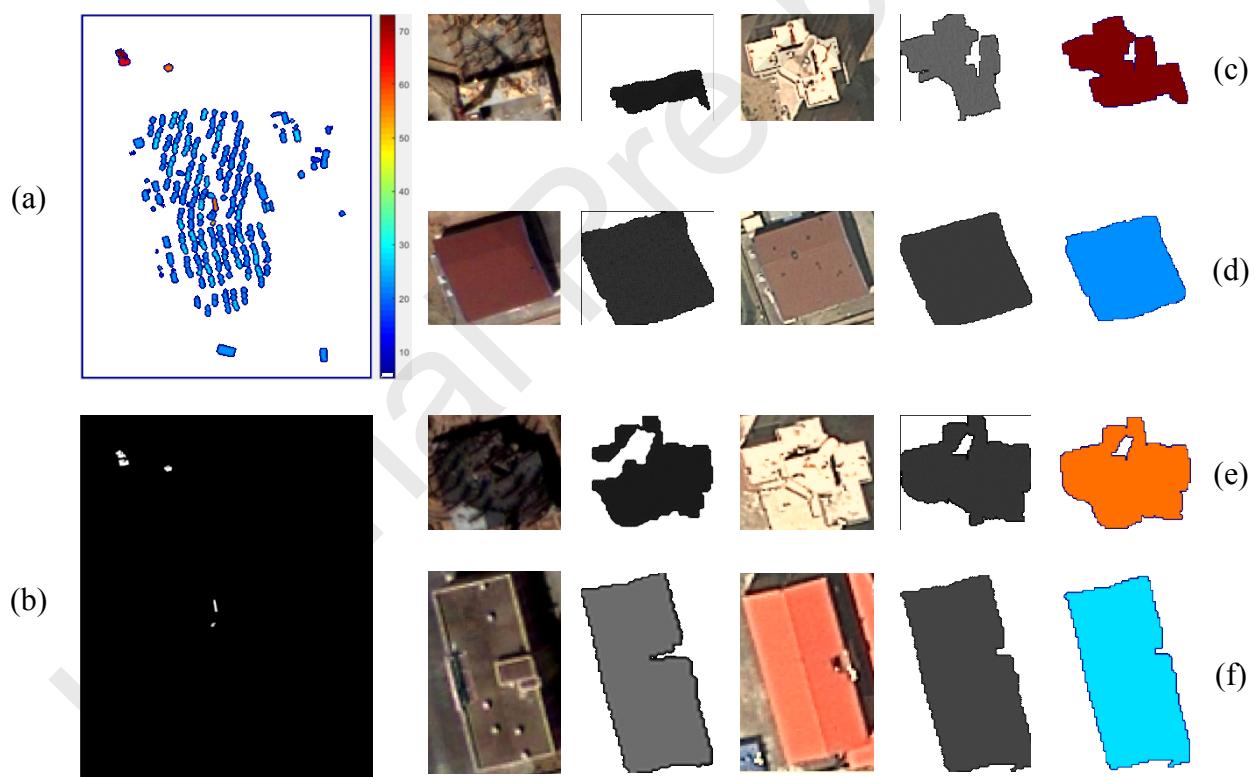
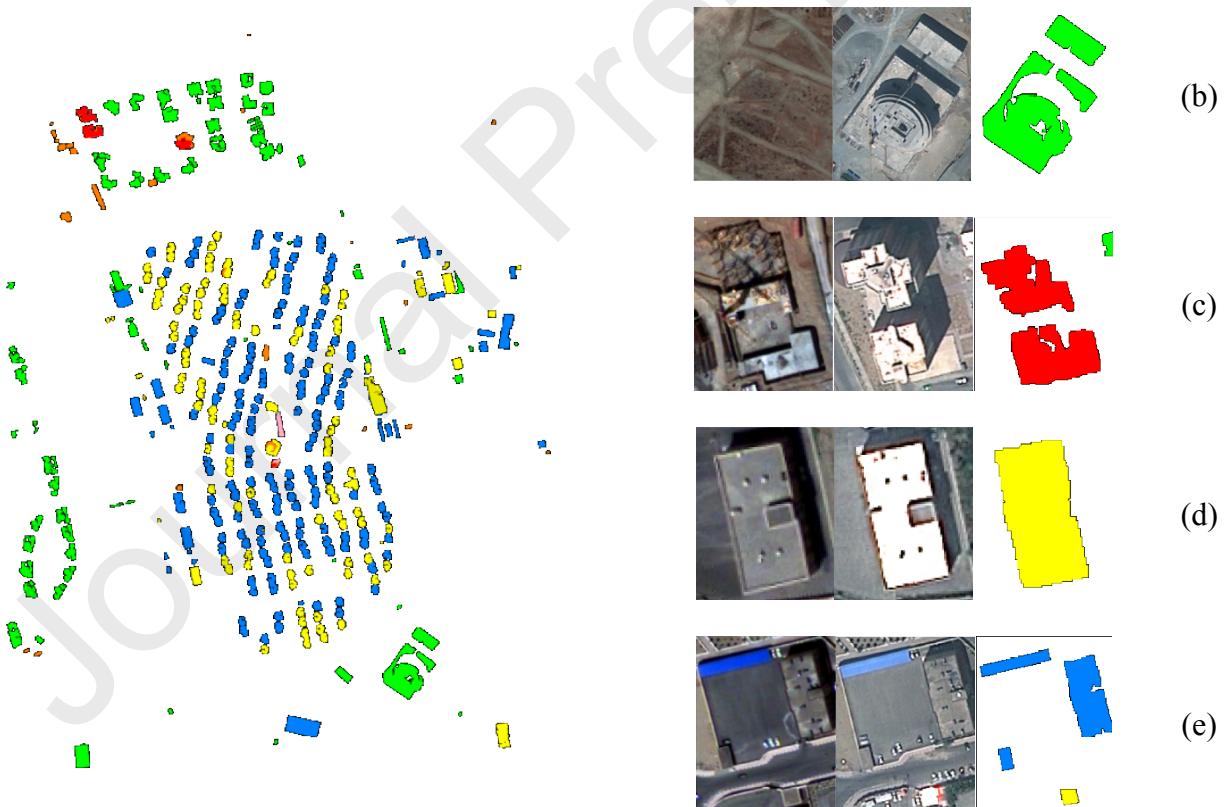


Figure 10: 3D change indicator output for the “Remained Building” objects: (a) 3D change indicator for entire image, (b) 3D changed objects after automatic thresholding of change indicator and (c) to (f) Some samples of test area including: the image patch in 2009, the nDM patch in 2009, the image patch in 2013, the nDM patch in 2013 and the 3D change indicator (left to right).

Now, Fig. (11) shows the result of fusing 2D and 3D change maps as the second level change analysis. This map provides more valuable information than simple change detection results and is more useful for urban managers. In the Figs. 11(b) to 11(g), the detected change class is given for a number of buildings in the study area. As can be seen, the accuracy of the change class determination is high. Although there are shortcomings in the boundaries of the extracted buildings (Fig. 11(b)), these defects have not been effective in determining the class of change. These defects are due to the shortcomings of the proposed building extraction method, which requires the use of more accurate methods in future works. Of course, it is clear that the building in Fig. 11(b) is a specially designed building, and examples of it are less commonly found in urban areas.



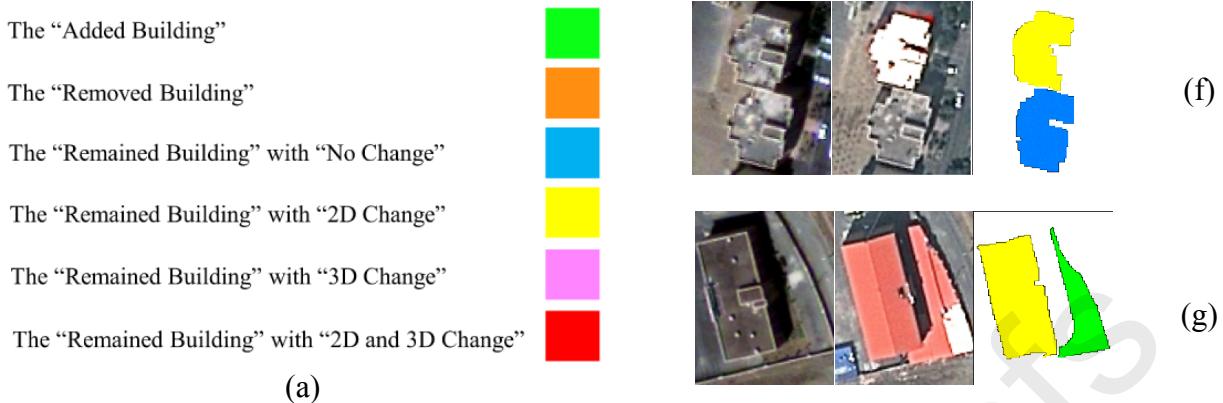


Figure 11: The second level change analysis results: (a) The result of entire test area and (b) to (g) some samples of test area including: the image patch in 2009, the image patch in 2013 and the second level change states (left to right).

3.2.9 Evaluation of the second level change analysis

For final quantitative evaluation, the output of the second level change analysis is compared with the ground truth data. The contribution of each class in the studied area for the ground truth and output data of this step is given in Table (7). In fact, Table (7) shows the contribution of each of the second level change analysis classes for the “Remained Building” objects. Also, Table (8) demonstrates the quantitative assessment of this result. The low F1-Score for the “Remained Building with Only 3D Change” is because of lack of this label in the ground truth data. It is acceptable because, in reality, almost always, a 3D change will have signs of a 2D change, and the likelihood of a 3D change occurring alone is low. Consequently, a building object is miss-labeled as 3D change only. However, accuracy of higher than 96% and F1-Score of higher than 72% for this step demonstrate the capability of the proposed method for analyzing building changes in high resolution satellite images.

Table 7: The contribution of “Only 2D Change”, “Only 3D Change”, “2D with 3D Change” and “No Change” buildings in the study area for the ground truth and the output data of the second level change analysis.

Product	Only 2D Change	Only 3D Change	2D with 3D Change	No Change
Second Level Ground Truth	32.2%	0.0%	1.8%	66.0%
Second Level Change Analysis	31.2%	8.0%	1.0%	59.8%

Table 8: Quantitative assessment of the second level change analysis step for “Remained Building” in the first and second dates.

Product	Accuracy	Sensitivity	Specificity	Precision	F1 Score	FPR
Only 2D Change	0.983	0.983	0.974	0.921	0.951	0.0262
Only 3D Change	1	1	0.995	0.0004	0.0008	0.0045
2D with 3D Change	0.547	0.547	0.997	0.819	0.656	0.0033
No Change	0.966	0.966	0.996	0.995	0.980	0.0039
Overall	0.961	0.888	0.992	0.739	0.725	0.0081

Finally, Table (9) shows the contribution of each class in the research area for the ground truth and the output data of change analysis steps. As can be seen, the results are acceptable and this shows the quality of the proposed method for satellite image change analysis.

Table 9: The contribution of each class of buildings in the study area for the ground truth and the output data of the first and second level change analysis simultaneously .

Product	Removed Building	Added Building	Remained Building			
			Only 2D Change	Only 3D Change	2D with 3D Change	No Change
Ground Truth	4.4%	16.5%	25.5%	0.0%	1.5%	52.1%
First and Second Level Change Analysis	5.4%	20.0%	23.2%	6.0%	0.8%	44.6%

4 Conclusions and Future Works

In this paper, an unsupervised procedure is proposed for building change analysis from high resolution satellite stereo images. The 2D and 3D information are fused in different steps of the proposed method to use the capability of each data type in the building change detection and analysis. Also, the capability of CNN is employed for 3D feature extraction to reduce the challenges of using height difference information in the building change detection. The proposed method also has made it possible to use the ability of deep networks without the need for

massive training data. Also, instead of analyzing the entire 3D data using the deep network, building objects are first extracted during an object-based analysis and, then, only the area limited to these objects enters the CNN network to reduce the processing time and the final accuracy increase. For future works, it is recommended to propose solutions for using 2D and 3D information simultaneously for building extraction using deep networks. This can reduce the uncertainties of the building extraction process that can lead to achieving more accurate results in the building change analysis step. Also, it is suggested for other researchers to apply the proposed method in different building structures and evaluate the results.

References

2019. *ImageNet* [Online]. © 2016 Stanford Vision Lab, Stanford University, Princeton University. Available: <http://www.image-net.org/> [Accessed 2019 2019].
- AKCA, D., FREEMAN, M., SARGENT, I. & GRUEN, A. 2010. Quality assessment of 3D building data. *The Photogrammetric Record*, 25, 339-355.
- CHAABOUNI-CHOUAYAKH, H., D'ANGELO, P., KRAUSS, T. & REINARTZ, P. Automatic urban area monitoring using digital surface models and shape features. 2011 Joint Urban Remote Sensing Event, 2011. IEEE, 85-88.
- CHAABOUNI-CHOUAYAKH, H. & REINARTZ, P. 2011. Towards automatic 3D change detection inside urban areas by combining height and shape information. *Photogrammetrie-Fernerkundung-Geoinformation*, 2011, 205-217 %@ 1432-8364.
- CHAMPION, N., BOLDO, D., PIERROT-DESEILLIGNY, M. & STAMON, G. 2010. 2D building change detection from high resolution satelliteimagery: A two-step hierarchical method based on 3D invariant primitives. *Pattern Recognition Letters*, 31, 1138-1147.
- CHAMPION, N., ROTTENSTEINER, F., MATIKAINEN, L., LIANG, X., HYYPÄ, J. & OLSEN, B. 2009. A test of automatic building change detection approaches. *Proceedings of CMRT09*, 03-04.
- CHOI, K., LEE, I. & KIM, S. 2009. A feature based approach to automatic change detection from LiDAR data in urban areas. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci*, 18, 259-264.
- CRISPELL, D., MUNDY, J. & TAUBIN, G. 2011. A variable-resolution probabilistic three-dimensional model for change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 489-500.
- DINI, G., JACOBSEN, K., ROTTENSTEINER, F., AL RAJHI, M. & HEIPKE, C. 2012. 3D building change detection using high resolution stereo images and a GIS database. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences-ISPRS Archives* 39 (2012), 39, 299-304.

- GHARIBBAFGHI, Z., TIAN, J. & REINARTZ, P. Superpixel-Based 3D Building Model Refinement and Change Detection, Using VHR Stereo Satellite Imagery. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2019. 0-0.
- GONG, J., HU, X., PANG, S. & WEI, Y. 2019. Roof-Cut Guided Localization for Building Change Detection from Imagery and Footprint Map. *Photogrammetric Engineering & Remote Sensing*, 85, 543-558.
- GONG, M., ZHAO, J., LIU, J., MIAO, Q. & JIAO, L. 2015. Change detection in synthetic aperture radar images based on deep neural networks. *IEEE transactions on neural networks and learning systems*, 27, 125-138.
- GUERIN, C., BINET, R. & PIERROT-DESEILLIGNY, M. 2014. Automatic detection of elevation changes by differential dsm analysis: Application to urban areas. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7, 4020-4037 %@ 1939-1404.
- HEARST, M. A., DUMAIS, S. T., OSUNA, E., PLATT, J. & SCHOLKOPF, B. 1998. Support vector machines. *IEEE Intelligent Systems and their applications*, 13, 18-28.
- HIRSCHMULLER, H. 2007. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on pattern analysis and machine intelligence*, 30, 328-341.
- KRAUß, T. 2018. A New Simplified DSM-to-DTM Algorithm—dsm-to-dtm-step.
- MATIKAINEN, L., HYYPÄÄ, J., AHOKAS, E., MARKELIN, L. & KAARTINEN, H. 2010. Automatic detection of buildings and changes in buildings for updating of maps. *Remote Sensing*, 2, 1217-1248.
- MOHAMMADI, H., SAMADZADEGAN, F. & REINARTZ, P. 2019. 2D/3D information fusion for building extraction from high-resolution satellite stereo images using kernel graph cuts. *International Journal of Remote Sensing*, 1-26.
- MOU, L., BRUZZONE, L. & ZHU, X. X. 2018. Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery. *IEEE Transactions on geoscience and remote sensing*, 57, 924-935.
- MURAKAMI, H., NAKAGAWA, K., HASEGAWA, H., SHIBATA, T. & IWANAMI, E. 1999. Change detection of buildings using an airborne laser scanner. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54, 148-152.
- NEBIKER, S., LACK, N. & DEUBER, M. 2014. Building change detection from historical aerial photographs using dense image matching and object-based image analysis. *Remote Sensing*, 6, 8310-8336.
- OTSU, N. 1979. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man and cybernetics*, 9, 62-66.
- PACIFICI, F., DEL FRATE, F., SOLIMINI, C. & EMERY, W. J. 2007. An innovative neural-net method to detect temporal changes in high-resolution optical satellite imagery. *IEEE Transactions on geoscience and remote sensing*, 45, 2940-2952.
- PANG, S., HU, X., CAI, Z., GONG, J. & ZHANG, M. 2018. Building change detection from bi-temporal dense-matching point clouds and aerial images. *Sensors*, 18, 966.
- PANG, S., HU, X., WANG, Z. & LU, Y. 2014. Object-based analysis of airborne LiDAR data for building change detection. *Remote Sensing*, 6, 10733-10749.
- PANG, S., HU, X., ZHANG, M., CAI, Z. & LIU, F. 2019. Co-Segmentation and Superpixel-Based Graph Cuts for Building Change Detection from Bi-Temporal Digital Surface Models and Aerial Images. *Remote Sensing*, 11, 729.

- PENG, D., ZHANG, Y. & GUAN, H. 2019. End-to-end change detection for high resolution satellite images using improved unet++. *Remote Sensing*, 11, 1382.
- PHL, P. B. 2018. *Visible Vegetation Index (VVI)* [Online]. Planetary Habitability Laboratory @ UPR Arecibo. Available: <http://phl.upr.edu/projects/visible-vegetation-index-vvi>.
- POSTED BY ALEX ALEMI, S. E. 2019. *Improving Inception and Image Classification in TensorFlow* [Online]. Google AI. Available: <https://ai.googleblog.com/2016/08/improving-inception-and-image.html> [Accessed 1 July 2019 2019].
- QIN, R. 2014a. Change detection on LOD 2 building models with very high resolution spaceborne stereo imagery. *ISPRS journal of photogrammetry and remote sensing*, 96, 179-192 %@ 0924-2716.
- QIN, R. 2014b. An object-based hierarchical method for change detection using unmanned aerial vehicle images. *Remote Sensing*, 6, 7911-7932.
- QIN, R. & GRUEN, A. 2014. 3D change detection at street level using mobile laser scanning point clouds and terrestrial images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 90, 23-35 %@ 0924-2716.
- QIN, R., GRUEN, A. & FRASER, C. Quality assessment of image matchers for DSM generation—a comparative study based on UAV images. Proceedings of the 35th Asian Conference on Remote Sensing (ACRS 2014). Nay Pyi Taw, Maynmar, 27–31 October 2014, 2014.
- QIN, R., HUANG, X., GRUEN, A. & SCHMITT, G. 2015. Object-based 3-D building change detection on multitemporal stereo images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8, 2125-2137.
- QIN, R., TIAN, J. & REINARTZ, P. 2016. 3D change detection—approaches and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 122, 41-56.
- REICHSTEIN, M., CAMPS-VALLS, G., STEVENS, B., JUNG, M., DENZLER, J. & CARVALHAIS, N. 2019. Deep learning and process understanding for data-driven Earth system science. *Nature*, 566, 195.
- SALAH, M. B., MITICHE, A. & AYED, I. B. 2011. Multiregion image segmentation by parametric kernel graph cuts. *IEEE Transactions on Image Processing*, 20, 545-557.
- SASAGAWA, A., BALTSAVIAS, E., AKSAKAL, S. K. & WEGNER, J. D. 2013. Investigation on automatic change detection using pixel-changes and DSM-changes with ALOS-PRISM triplet images. *International archives of the photogrammetry, remote sensing spatial information sciences*, 7, W2.
- SCHINDLER, G. & DELLAERT, F. Probabilistic temporal inference on reconstructed 3d scenes. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010. IEEE, 1410-1417.
- SINGH, A. 1989. Review Article Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing*, 10, 989-1003.
- STAL, C., TACK, F., DE MAEYER, P., DE WULF, A. & GOOSSENS, R. 2013. Airborne photogrammetry and lidar for DSM extraction and 3D change detection over an urban area—a comparative study. *International Journal of Remote Sensing*, 34, 1087-1110.
- SZEGEDY, C., IOFFE, S., VANHOUCKE, V. & ALEMI, A. A. Inception-v4, inception-resnet and the impact of residual connections on learning. Thirty-First AAAI Conference on Artificial Intelligence, 2017.

- TANEJA, A., BALLAN, L. & POLLEFEYS, M. Image based detection of geometric changes in urban environments. 2011 International Conference on Computer Vision, 2011. IEEE, 2336-2343.
- TANEJA, A., BALLAN, L. & POLLEFEYS, M. City-scale change detection in cadastral 3d models using images. Proceedings of the IEEE Conference on computer Vision and Pattern Recognition, 2013. 113-120.
- TIAN, J. 2013. *3D change detection from high and very high resolution satellite stereo imagery*. Ph. D. Thesis, Universität Osnabrück, Osnabrück, Germany.
- TIAN, J., CHAABOUNI-CHOUAYAKH, H., REINARTZ, P., KRAUß, T. & D'ANGELO, P. Automatic 3D change detection based on optical satellite stereo imagery. 2010. na.
- TIAN, J., CUI, S. & REINARTZ, P. 2014a. Building change detection based on satellite stereo imagery and digital surface models. *IEEE Transactions on Geoscience and Remote Sensing*, 52, 406-417.
- TIAN, J., DEZERT, J. & REINARTZ, P. Refined building change detection in satellite stereo imagery based on belief functions and reliabilities. 2015 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), 2015. IEEE, 160-165.
- TIAN, J., NIELSEN, A. A. & REINARTZ, P. 2014b. Improving change detection in forest areas based on stereo panchromatic imagery using kernel MNF. *IEEE Transactions on Geoscience and Remote Sensing*, 52, 7130-7139.
- TRINDER, J. & SALAH, M. 2012. Aerial images and LiDAR data fusion for disaster change detection. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci*, 1, 227-232.
- ULUSOY, A. O. & MUNDY, J. L. Image-based 4-d reconstruction using 3-d change detection. 2014. Springer, 31-45.
- WANG, Q., ZHANG, X., CHEN, G., DAI, F., GONG, Y. & ZHU, K. 2018. Change detection based on Faster R-CNN for high-resolution remote sensing images. *Remote sensing letters*, 9, 923-932.
- WANG, Z., BOVIK, A. C., SHEIKH, H. R. & SIMONCELLI, E. P. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13, 600-612.
- WEN, D., HUANG, X., ZHANG, A. & KE, X. 2019. Monitoring 3D Building Change and Urban Redevelopment Patterns in Inner City Areas of Chinese Megacities Using Multi-View Satellite Imagery. *Remote Sensing*, 11, 763.
- WIKIPEDIA. 2019. *Confusion matrix* [Online]. From Wikipedia, the free encyclopedia. Available: https://en.wikipedia.org/wiki/Confusion_matrix [Accessed 25 June 2019, at 20:13 (UTC). 2019].
- WOLF, P. R. & GHILANI, C. 1997. Adjustment computations-statistics and least. *squares in Surveying and GIS*.
- XIAO, W., VALLET, B. & PAPARODITIS, N. 2013. Change detection in 3d point clouds acquired by a mobile mapping system. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1, 331-336.
- ZAITOUN, N. M. & AQEL, M. J. 2015. Survey on image segmentation techniques. *Procedia Computer Science*, 65, 797-806.
- ZAVODNY, A. G. 2012. *Change detection in lidar scans of urban environments*, University of Notre Dame.

- ZHANG, Z., VOSSELMAN, G., GERKE, M., PERSELLO, C., TUIA, D. & YANG, M. 2019. CHANGE DETECTION BETWEEN DIGITAL SURFACE MODELS FROM AIRBORNE LASER SCANNING AND DENSE IMAGE MATCHING USING CONVOLUTIONAL NEURAL NETWORKS. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 4.
- ZUFENG, W., TIAN, L., JIANG, W., YI, D. & ZHIGUANG, Q. 2016. Medical Image Registration Using B-Spline Transform. *International Journal of Simulation: Systems, Science and Technology*, 17.