Patch-level unsupervised planetary change detection

Sudipan Saha *Member, IEEE* and Xiao Xiang Zhu, *Fellow, IEEE*

***Abstract*—Change detection (CD) is critical for analyzing data collected by planetary exploration missions, e.g., for identification of new impact craters. However, CD is still a relatively new topic in the context of planetary exploration. Sheer variation of planetary data makes CD much more challenging than in the case of Earth observation (EO). Unlike CD for EO, patch-level decision is preferred in planetary exploration as it is difficult to obtain perfect pixelwise alignment/co-registration between the bi-temporal planetary images. Lack of labeled bi-temporal data impedes supervised CD. To overcome these challenges, we propose an unsupervised CD method that exploits a pre-trained feature extractor to obtain bi-temporal deep features that are further processed using global max-pooling to obtain patch-level feature description. Bi-temporal patch-level features are further analyzed based on difference to determine whether a patch is changed. Additionally, a self-supervised method is proposed to estimate the decision boundary between the changed and unchanged patches. Experimental results on three planetary CD datasets from two different planetary bodies (Mars and Moon) demonstrate that the proposed method often outperforms supervised planetary CD methods. Code is available at https://gitlab.lrz.de/ai4eo/cd/-/tree/main/planetaryCDUnsup.**

***Index Terms*—Change detection, Planetary exploration, Transfer learning, Pooling, Unsupervised learning.**

I. Introduction

Interest in planetary exploration missions has increased significantly in the last decade [1] as such missions enrich our knowledge about the solar system [2]. A crucial role in most such missions is played by the scientific imaging instruments that are used for various purposes, including planetary surface characterization and spectral mapping for mineralogy. Large volume of images are currently being captured by ongoing planetary imaging investigations, e.g., High Resolution Imaging Science Experiment (HiRISE) and ConTeXtCamera (CTX) on the MarsReconnaissance Orbiter [3], [4].

Change detection (CD) is one of the most studied topics in Earth observation (EO). CD plays crucial role in several EO tasks, e.g., disaster management [5], urban monitoring [6], and military applications. Despite its established significance in EO, CD has not been explored much in context of planetary exploration. However just like EO, CD may play a significant role in planetary explorations. As detailed by Kerner *et. al.* [3], one such application of CD in planetary exploration is to monitor the changes induced by meteorite impact. Such impacts strongly alter the landscape of the planets. Another such application is monitoring of recurring slope lineae (RSL) that appear/disappear on surface of Mars on timescales close to a year.

Kerner *et. al.* [3] studied several supervised methods in context of planetary CD. However, when talking about CD for EO, unsupervised methods are more popular than the supervised ones [6]. This is because collection of labeled multi-temporal training data is difficult in context of CD. Moreover, even if training data is collected for one particular application or geography, supervised methods do not generalize well for other applications or geography. While variation of geography limits applicability of supervised CD methods for EO (i.e., one planet), it certainly limits their applicability when considering many planets and their hundreds of moons. This is confirmed by the work of Kerner *et. al.* [3], where accuracy significantly drops when a model trained on HiRISE RSL dataset is applied on CTX meteorite impacts dataset. This shows the necessity of moving beyond supervised methods for planetary CD. Another difference between planetary CD and CD for EO is that the latter assumes near-perfect pixelwise alignment between bi-temporal input while such alignment is difficult to obtain for the former. Due to this, previous work on planetary CD [3] focuses on patch-level CD instead of pixelwise CD.

Deep transfer learning methods that exploit a pre-trained model for bi-temporal feature extraction and comparison, have shown excellent performance in different CD applications [6], [5]. Inspired by this we propose a deep transfer learning based CD method that ingests bi-temporal patches and processes it through a set of convolution layers (pre-trained network). The bi-temporal feature maps are processed using a global max-pooling to summarize the content of both patches and account for possible presence of misalignment. Finally the difference of features after global max-pooling are taken, thresholded using a decision boundary obtained with a self-supervised method, to determine whether the considered patch is changed/unchanged.

The key contributions of this paper are as follows:

1) In context of planetary CD, this paper proposes an unsupervised deep transfer learning based method that can determine whether a pair of bi-temporal patches are changed, even if they are not perfectly co-registered.
2) This paper further proposes a method using pseudo unchanged pairs to determine threshold for distinguishing

The work is jointly supported by by the German Federal Ministry of Education and Research (BMBF) in the framework of the international future AI lab "AI4EO – Artificial Intelligence for Earth Observation: Reasoning, Uncertainties, Ethics and Beyond" (grant number: 01DD20001), the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. [ERC-2016-StG-714087], Acronym: *So2Sat*), by the Helmholtz Association through the Framework of Helmholtz AI (grant number: ZT-I-PF-5-01) - Local Unit "Munich Unit @Aeronautics, Space and Transport (MASTr)" and Helmholtz Excellent Professorship "Data Science in Earth Observation - Big Data Fusion for Urban Research"(grant number: W2-W3-100) (Corresponding author: Xiao Xiang Zhu.)

Sudipan Saha is with Department of Aerospace and Geodesy, Data Science in Earth Observation, Technical University of Munich, 85521 Ottobrunn, Germany. E-mail: sudipan.saha@tum.de

Xiao Xiang Zhu is with Remote Sensing Technology Institute, German Aerospace Center (DLR), 82234 Weßling, Germany and also with Department of Aerospace and Geodesy, Data Science in Earth Observation, Technical University of Munich, 85521 Ottobrunn, Germany. E-mail: xiaoxiang.zhu@dlr.de

changed and unchanged patches.

3) The paper validates the proposed method on three diverse planetary CD datasets, two of which show misalignment between pre-temporal and post-temporal images.

We organize the rest of the paper as follows. Related works are discussed in Section II. Section III outlines the proposed method. Datasets and experimental results are detailed in Section IV. Finally, we conclude the paper in Section V.

II. Related Work

Considering relevance to our work, in this Section we briefly discuss unsupervised CD and planetary CD.

A. Unsupervised CD in EO

Prior to the emergence of deep learning, most unsupervised CD methods for EO used the concept of pixelwise image differencing, i.e., change vector analysis (CVA) [7]. Many variants of CVA, e.g., Parcel CVA (PCVA) [8] and Robust CVA (RCVA) [9], incorporated the notion of spatial context in CVA. Deep learning based unsupervised CD methods are generally based on transfer learning [6]. Deep CVA (DCVA) [6] is one such framework that incorporates CVA with pre-trained deep network based feature extraction. DCVA has shown excellent performance in many tasks, e.g., building CD [5] and agricultural monitoring [10]. Another class of unsupervised CD methods pre-classifies some samples with high confidence as changed/unchanged using some traditional approach and further uses those confident samples for training a CD model [11]. Such methods have limited applicability in planetary CD as error introduced by sensor or unseen geographic characteristics may significantly impact the choice of confident samples.

B. CD in planetary exploration

Kerner *et. al.* [3] presented a detailed study of the supervised methods for planetary CD. To the best of our knowledge, this is the only work on planetary CD. They specifically promoted the use of convolutional autoencoder along with different supervised classifiers for planetary CD. For the scenario where training data is not available, they reused the supervised CD model trained on another dataset. However, their results clearly show that such straightforward reuse of supervised model does not perform satisfactorily when the source and target domains are significantly different, e.g., when source domain is HiRISE RSL dataset and the target domain is CTX meteorite impacts dataset.

III. Proposed method

Let us consider two sets of I unlabeled patches $\mathbf{X_1} = \{\mathbf{x_{1i}}, \forall i = 1, ..., I\}$ and $\mathbf{X_2} = \{\mathbf{x_{2i}}, \forall i = 1, ..., I\}$, captured over same planetary surface at time t_1 and t_2. Spatial dimension of the patches are $R \times C$. We assume that $\mathbf{x_{1i}}$ and $\mathbf{x_{2i}}$ represent same area/object, however may not be properly aligned, which is in stark contrast to near-perfect alignment assumption made in most CD methods [9], [6]. Instead of pixel-level CD, we are interested to assign each patch pair a label: changed or unchanged.

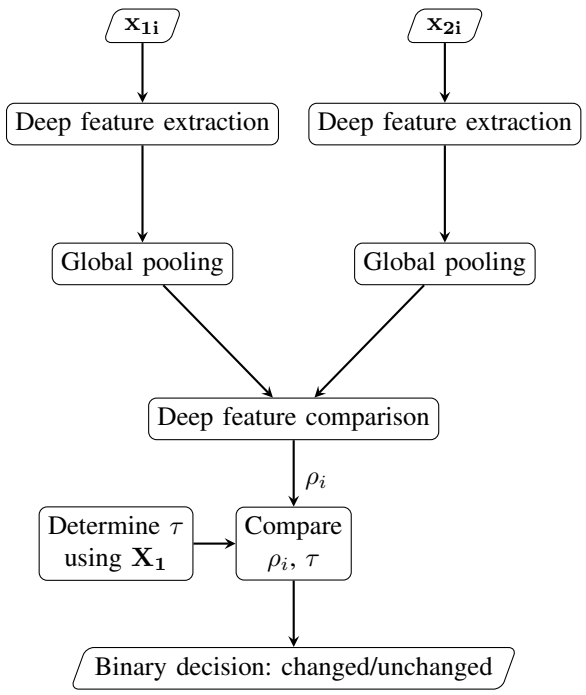


Fig. 1. Proposed patch-level planetary unsupervised CD technique.

Both $\mathbf{x_{1i}}$ and $\mathbf{x_{2i}}$ are separately processed through a multi-layered pre-trained CNN feature extractor, originally trained for some other task (see Section III-A). Pixelwise features are summarized using global max-pooling to obtain $\mathbf{f_{1i}}$ and $\mathbf{f_{2i}}$, corresponding to $\mathbf{x_{1i}}$ and $\mathbf{x_{2i}}$, respectively (see Section III-B). They are compared to obtain a deep feature difference ρ_i, larger value of which indicates the patch-pair $\mathbf{x_{1i}}$ and $\mathbf{x_{2i}}$ is more likely to be changed. We can obtain a binary changed/unchanged decision by comparing ρ_i to a threshold τ (see Section III-C). We determine the optimum value of τ by using a self-supervised mechanism that does not require any label information and uses only the pre-temporal sets of patches $\mathbf{X_1}$ (see Section III-D). The proposed patch-level CD method is shown in Figure 1.

While the proposed method is intended for patch-level CD, pixelwise CD map can also be obtained, which is briefly discussed in Section III-E.

A. Deep features extraction

Similar to the DCVA framework [6], bi-temporal patches $\mathbf{x_{1i}}$ and $\mathbf{x_{2i}}$ are separately processed using a pre-trained CNN to obtain features corresponding to both $\mathbf{x_{1i}}$ and $\mathbf{x_{2i}}$. In this work, we use VGG-16 model [12] trained on natural image dataset, however any other suitable pre-trained model can also be used [13]. By applying this network on our target planetary images, we reuse the trained CNN model to transfer the visual descriptors learned by the CNN for its original training task to solve the planetary CD problem. As previous works have shown [6], the intermediate layers are more suitable for transfer learning on targets that are semantically different from

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LGRS.2021.3130862, IEEE Geoscience and Remote Sensing Letters

original training data, as the lower layers of the CNN capture primitive features like edges and the very last layers capture features specific to the original training dataset. Following this, we use the sixth convolution layer of VGG-16 [12] for feature extraction. The dimension of the feature maps are $R' \times C' \times D$ where D is the number of features. R' may or may not be equal to R depending on whether there are downsampling operations in the pre-trained feature extractor. Unlike [6], due to possible misalignment between $\mathbf{x_{1i}}$ and $\mathbf{x_{2i}}$, straightforward pixelwise comparison of $\mathbf{x_{1i}}$ and $\mathbf{x_{2i}}$ may lead to errors.

Please note that the above-mentioned process of feature extraction does not require prior availability of any planetary data, labeled or unlabeled.

B. Patch summarization

The output feature maps obtained in the previous step are sensitive to the location of the features in the input patches. While this property is useful to obtain pixelwise CD map if $\mathbf{x_{1i}}$ and $\mathbf{x_{2i}}$ are perfectly aligned, this may cause error in our case, as $\mathbf{x_{1i}}$ and $\mathbf{x_{2i}}$ are possibly misaligned. For patch-level CD, we need to summarize/downsample the feature maps in a way such that summarization process is robust to the shift in the position of the feature in the image.

We use global pooling operation as a global image descriptor, i.e., to summarize the patch. In contrast to popularly used average pooling, we use max pooling. Average pooling assumes that the feature descriptors in a patch are independent and identically distributed. Thus, average pooling is sensitive to more frequently occurring descriptors, a phenomenon known as visual burstiness [14] that hinders average pooling from capturing feature relevant to distinguish different samples.

The use of spatial max pooling was introduced by LeCun et al. [15] and was later extended for global max pooling [16]. Remarkably, global max pooling achieves partial invariance to small translations because the max of a patch depends only on the single largest element in the given patch. If a small translation doesn't bring in a new largest element at the edge of the patch or does not remove the largest element by taking it outside of the patch, then the outcome of global max pooling does not change.

Global max pooling summarizes $R' \times C' \times D$ feature maps obtained from $\mathbf{x_{1i}}$ and $\mathbf{x_{2i}}$ to D-dimensional $\mathbf{f_{1i}}$ and $\mathbf{f_{2i}}$, respectively.

C. Determining whether patch is changed

$\mathbf{f_{1i}}$ is subtracted from $\mathbf{f_{2i}}$ and ℓ_1 norm is applied on difference to obtain a change indicator ρ_i. Denoting the steps from deep feature extraction to obtaining change indicator as g:

$$\rho_i = g(\mathbf{x_{1i}}, \mathbf{x_{2i}}) \tag{1}$$

Larger value of ρ_i indicates the patch-pair $\mathbf{x_{1i}}$ and $\mathbf{x_{2i}}$ are more likely to be changed. We can determine whether the patch-pair $\mathbf{x_{1i}}$ and $\mathbf{x_{2i}}$ are changed or not by comparing ρ_i to a threshold τ.

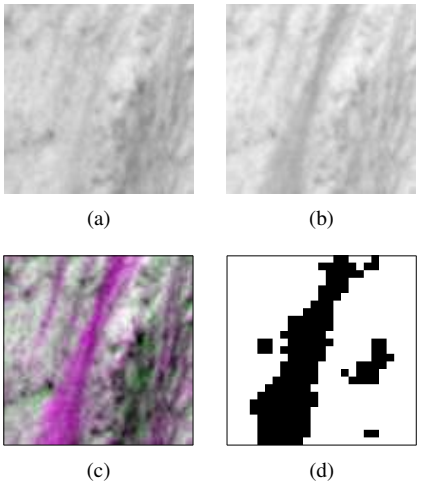


Fig. 2. An example from HiRISE RSL dataset: (a) pre-change patch, (b) post-change patch, (c) False color composition created using R: pre-change, G: post-change, B: pre-change patch, (d) pixelwise CD map, as detected by the proposed method.

D. Threshold determination

We further propose an automatic self-supervised method of determining τ used in Section III-C. Taking an unlabeled patch $\mathbf{x_{1i}}$ from $\mathbf{X_1}$, we obtain a noisy version $\mathbf{x'_{1i}}$ as:

$$\mathbf{x'_{1i}} = h(\mathbf{x_{1i}}) \tag{2}$$

where, $h(.)$ is equivalent to applying Gaussian noise and shifting the patch by few pixels. In practice, we applied shift of up to 10 pixels. $\mathbf{x_{1i}}$ and $\mathbf{x'_{1i}}$ can be treated as pseudo unchanged pair as they represent the same scene however with slight differences as commonly observed in multi-temporal planetary images. Thus, $\mathbf{x_{1i}}$ and $\mathbf{x'_{1i}}$ are processed following the steps in Sections III-B, III-C, III-D to obtain ρ'_i:

$$\rho'_i = g(\mathbf{x_{1i}}, \mathbf{x'_{1i}}) \tag{3}$$

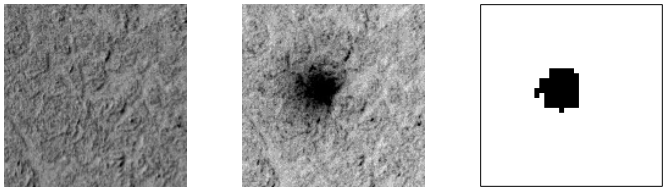
ρ'_i is a sample ρ value for an unchanged pair. Similarly, I such values can be generated for $i = 1, ..., I$, which provides us a distribution of ρ for the unchanged pairs. An upper bound for ρ for the unchanged pairs can be used as threshold τ to distinguish the unchanged patch pairs from the changed ones. τ can be obtained as:

$$\tau = max(\rho'_1, \rho'_2, ..., \rho'_I) \tag{4}$$

In practice, accounting for possible anomaly we exclude the top 5 percentile of ρ'_i while calculating τ.

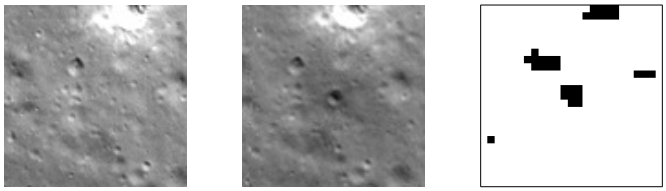
E. Pixelwise CD

If the proposed method determines a patch-pair $\mathbf{x_{1i}}$ and $\mathbf{x_{2i}}$ to be changed, pixelwise CD map can also be obtained if the patches are reasonably aligned, by pixelwise comparison of feature maps obtained in Section III-A using the DCVA framework [6].



(a) (b) (c)

Fig. 3. An example from CTX meteorite impact dataset: (a) pre-change patch; (b) post-change patch; (c) pixelwise CD map, as detected by the proposed method.



(a) (b) (c)

Fig. 4. An example from LROC (Moon) dataset: (a) pre-change patch, (b) post-change patch, and (c) pixelwise CD map as detected by the proposed method.

IV. Experimental Results

A. Datasets

Following planetary datasets are used for evaluation of the proposed method:

1) **HiRISE RSL**: RSL are dark and narrow features that are thought to be formed due to the shallow subsurface water flow. They incrementally fade and recur throughout the year. RSL dataset collected by [3] focuses on Garni crater on Mars. The images in the dataset are collected using the HiRISE camera which is onboard the Mars Reconnaissance Orbiter. The images are of approximately 30 cm/pixel. The dataset is formed using red channel [3]. A total of 254 pairs of 100×100 pixels size (examples in Figure 2) are used in the test set [3]. The dataset also has training and validation set that are not used by us since our proposed method is unsupervised and thus do not require training.
2) **CTX meteorite impacts**: Many planets including Mars are continuously impacted by meteorites. It is important to know the occurrence and location of such impact as this data helps scientists to estimate the current cratering rate in our solar system [17], [18]. The meterorite impact dataset collected by [3] is composed of 96 images pairs (example in Figure 3) that are collected by CTX onboard the Mars Reconnaissance Orbiter with a spatial resolution of 6 m/pixel and size of 150×150 pixels. Compared to HiRISE RSL, the dataset size is much smaller, 254 test cases versus only 96 test cases.
3) **LROC Moon dataset**: In this dataset, surface changes are the result of meteorite impacts and a spacecraft landing. The bi-temporal images in this dataset are misregistered by as many as 40 pixels, making it a challenging dataset. The test set consists of 5 changed pairs and 5 unchanged pairs of 100×100 pixels each (example in Figure 4).

B. Experiment objectives

The only existing work on planetary CD [3] experimented on a set of supervised learning paradigms (combination of classifier and input representation) to find out suitable supervised learning paradigm for planetary CD. This experiment was performed on HiRISE RSL dataset. They further experimented on CTX meterorite impacts dataset and LROC dataset to investigate if models trained on HiRISE RSL dataset can be transferred for CD in those datasets.

We compare our proposed unsupervised CD method to the following methods:

1) All supervised and transfer learning paradigms in [3].
2) Unsupervised RCVA [9]. While RCVA is proposed for pixel-based prediction, we designed a patch-based version of it by following similar strategies used for the proposed method.
3) A variant of the proposed method using global average pooling instead of max pooling.
4) A variant of the proposed method where instead of the proposed thresholding scheme, the ρ_i values corresponding to all patch pairs are clustered using k-means clustering with $k = 2$ to obtain two clusters, corresponding to the changed and unchanged patches, respectively.

C. Result analysis

HiRISE RSL dataset: Two supervised methods slightly outperform (by 0.4%) the proposed unsupervised method. However, surprisingly the proposed method outperforms as many as four supervised methods (Table I). The proposed method outperforms Inception-v3 based three different methods, Inception-v3 with signed difference, Inception v-3 with bottleneck representation, and Inception-v3 with composite grayscale. For more details of these supervised methods, refer to [3]. This result shows that proposed method's performance is almost comparable to the best performing supervised model. The proposed method also outperforms its average pooling based version, k-means clustering based version, and unsupervised RCVA. Moreover, Table I shows result of the proposed method using fifth and seventh convolution layer of VGG-16 for feature extraction. One example from HiRISE dataset is shown in Figure 2 that shows the proposed method is capable to indicate the location of change.

CTX meteorite impacts dataset: Here we investigate the proposed method to the transfer learning capability of the supervised methods in [3]. As tabulated in Table II, the proposed method outperforms all supervised methods on CTX meteorite impacts dataset. Notably this dataset has some pairs with significant misalignment. Superior result of the proposed method indicates that it is able to handle misalignment by effectively summarizing patch with global max-pooling. Moreover, this proves the superiority of the proposed unsupervised method in comparison to merely transferring supervised CD model from another dataset. An example from the CTX meteorite impacts dataset is shown in Figure 3.

LROC Moon dataset: In spite of strong misalignment error in this dataset, the proposed method is able to distinguish the changed patches from the unchanged ones. The proposed

TABLE I
COMPARISON OF THE PROPOSED UNSUPERVISED METHOD WITH THE SUPERVISED METHODS IN [3] FOR HIRISE RSL.

Method	Accuracy %
Proposed	94.1
Proposed (average pooling)	92.1
Proposed (k-Means clustering)	91.7
Proposed (Layer 7)	92.1
Proposed (Layer 5)	90.9
Inception-v3 absolute difference	**94.5**
Naive Bayes pixel difference	**94.5**
Inception-v3 signed difference	91.7
Inception-v3 composite grayscale	89.8
SVM absolute difference	87.8
Inception-v3 bottleneck representation	70.1
RCVA	64.6

TABLE II
COMPARISON OF THE PROPOSED UNSUPERVISED METHOD WITH THE TRANSFER LEARNING CAPABILITY OF THE SUPERVISED METHODS IN [3] FOR CTX METEORITE IMPACT DATASET.

Method	Accuracy %
Proposed	**81.3**
Proposed (average pooling)	76.0
Proposed (k-Means clustering)	75.0
Inception-v3 bottleneck representation	75.0
Inception-v3 absolute difference	71.9
Inception-v3 signed difference	54.2
Inception-v3 composite grayscale	55.2
Naive Bayes pixel difference	52.1
SVM absolute difference	67.7
RCVA	59.4

method successfully labels 9 out of 10 test patches, in comparison to the method in [3] that can only correctly label 8 patches. Proposed method also outperforms its average pooling variant, RCVA (both correctly label 7 patches), and k-means clustering based variant (correctly labels 8 patches). This shows that global max-pooling successfully summarizes the content of bi-temporal feature and subsequent comparison effectively identifies the changed patches. Example of a prediction on this dataset is shown in Figure 4 that shows the proposed method is capable to indicate the location of impact despite misalignment in dataset.

V. CONCLUSIONS

To the best of our knowledge, this paper introduces deep transfer learning based unsupervised CD for first time in the context of planetary exploration. Humankind has reached far beyond Earth and missions to new destinations are launched periodically. This poses us with challenge of processing multitemporal planetary data with huge variations. The proposed unsupervised method enables us to process varieties of unlabeled multitemporal planetary data without using any label. Towards this, the proposed method cleverly exploits deep transfer learning along with automatic threshold determination. Though unsupervised, proposed method outperforms most of the existing supervised methods. This shows the proposed method as a suitable option when labeled multi-temporal planetary data is not available. Considering the varieties of planetary mission and applications, it is impossible to always have abundant labeled multi-temporal data. Though proposed for multispectral input, the proposed method can be easily modified for hyperspectral input by choosing a different pre-trained network. Though proposed in the context of planetary CD, proposed method can be extended for any CD applications that require patch-level decision. Our future work will experiment on more planetary data. To conclude, our work is one step further in better understanding the temporal evolution of space and other planets.

REFERENCES

[1] H. R. Kerner, K. L. Wagstaff, B. D. Bue, D. F. Wellington, S. Jacob, P. Horton, J. F. Bell, C. Kwan, and H. B. Amor, "Comparison of novelty detection methods for multispectral images in rover-based planetary exploration missions," *Data Mining and Knowledge Discovery*, vol. 34, no. 6, pp. 1642–1675, 2020.

[2] J. Vago, O. Witasse, H. Svedhem, P. Baglioni, A. Haldemann, G. Gianfiglio, T. Blancquaert, D. McCoy, and R. de Groot, "ESA exomars program: the next step in exploring mars," *Solar System Research*, vol. 49, no. 7, pp. 518–528, 2015.

[3] H. R. Kerner, K. L. Wagstaff, B. D. Bue, P. C. Gray, J. F. Bell, and H. B. Amor, "Toward generalized change detection on planetary surfaces with convolutional autoencoders and transfer learning," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 10, pp. 3900–3918, 2019.

[4] T. Wilhelm, M. Geis, J. Püttschneider, T. Sievernich, T. Weber, K. Wohlfarth, and C. Wöhler, "Domars16k: A diverse dataset for weakly supervised geomorphologic analysis on mars," *Remote Sensing*, vol. 12, no. 23, p. 3981, 2020.

[5] S. Saha, F. Bovolo, and L. Bruzzone, "Building change detection in VHR SAR images via unsupervised deep transcoding," *IEEE Transactions on Geoscience and Remote Sensing*, 2020.

[6] S. Saha, F. Bovolo, and L. Bruzzone, "Unsupervised deep change vector analysis for multiple-change detection in VHR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 6, pp. 3677–3693, 2019.

[7] W. A. Malila, "Change vector analysis: an approach for detecting forest changes with Landsat," in *LARS symposia*, p. 385, 1980.

[8] F. Bovolo, "A multilevel parcel-based approach to change detection in very high resolution multitemporal images," *IEEE GRSL*, vol. 6, no. 1, pp. 33–37, 2009.

[9] F. Thonfeld, H. Feilhauer, M. Braun, and G. Menz, "Robust change vector analysis (RCVA) for multi-sensor very high resolution optical satellite data," *International Journal of Applied Earth Observation and Geoinformation*, vol. 50, pp. 131–140, 2016.

[10] S. Saha, Y. T. Solano-Correa, F. Bovolo, and L. Bruzzone, "Unsupervised deep transfer learning-based change detection for HR multispectral images," *IEEE Geoscience and Remote Sensing Letters*, 2020.

[11] H. M. Keshk and X.-C. Yin, "Change detection in SAR images based on deep learning," *International Journal of Aeronautical and Space Sciences*, pp. 1–11, 2019.

[12] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[13] G. Sumbul, J. Kang, T. Kreuziger, F. Marcelino, H. Costa, P. Benevides, M. Caetano, and B. Demir, "Bigearthnet dataset with a new class-nomenclature for remote sensing image understanding," *arXiv preprint arXiv:2001.06372*, 2020.

[14] V. Christlein, L. Spranger, M. Seuret, A. Nicolaou, P. Král, and A. Maier, "Deep generalized max pooling," in *2019 International Conference on Document Analysis and Recognition (ICDAR)*, pp. 1090–1096, IEEE, 2019.

[15] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[16] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Learning and transferring mid-level image representations using convolutional neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1717–1724, 2014.

[17] I. Daubar, A. S. McEwen, S. Byrne, M. Kennedy, and B. Ivanov, "The current martian cratering rate," *Icarus*, vol. 225, no. 1, pp. 506–516, 2013.

[18] W. K. Hartmann and G. Neukum, "Cratering chronology and the evolution of mars," *Chronology and evolution of Mars*, pp. 165–194, 2001.