FUSING MULTIPLE UNTRAINED NETWORKS FOR HYPERSPECTRAL CHANGE DETECTION

Sudipan Saha[1,*], Jakob Gawlikowski[1,2], Xiao Xiang Zhu[1,3]

[1] Data Science in Earth Observation, Technical University of Munich, Ottobrunn, Germany - sudipan.saha@tum.de
[2] Institute of Data Science, German Aerospace Center (DLR), Jena, Germany - jakob.gawlikowski@dlr.de
[3] Remote Sensing Technology Institute, German Aerospace Center (DLR), Weßling, Germany - xiaoxiang.zhu@dlr.de

KEY WORDS: Hyperspectral Images, Change Detection, Untrained Networks, Decision Fusion.

ABSTRACT:

Change detection in hyperspectral images is challenging due to the presence of a large number of spectral bands. Due to the differences in band composition, a deep model trained on one hyperspectral sensor cannot be reused on another hyperspectral sensor. This challenge can be tackled by using untrained models as feature extractor for change detection in hyperspectral images. However, results produced by such a strategy may show variance if the untrained model is slightly perturbed. Different change detection maps are produced from different versions of the untrained model. We propose a decision fusion based strategy that can combine such different results and produce a final change detection map that retains the change information from all change maps. This approach improves the change detection performance and also improves reliability of the result. Experimental results on two publicly available hyperspectral datasets show the effectiveness of the proposed approach.

1. INTRODUCTION

Hyperspectral sensors have become popular in remote sensing due to their capability to characterize fine-grained spectral information. They analyze a wide electromagnetic spectrum unlike other optical technologies that only capture the primary colors. Hyperspectral remote sensing is well accepted in many applications, including mineral characterization (Kruse et al., 2003), water quality monitoring (Cao et al., 2021), ship monitoring (Li et al., 2021), and vegetation mapping (Hirano et al., 2003). In the hyperspectral images, the spectral information is added as additional value to the two-dimensional spatial data, thus forming a data hypercube. While the fine-grained spectral information is the main motivation behind using hyperspectral images, handling the large number of bands is a key challenge in processing them (Mou et al., 2021).

In addition to mere classification, change detection (CD) plays a key role in many hyperspectral remote sensing applications (Chevrel et al., 2012). Deep transfer learning has recently emerged as a popular method for change detection in multispectral (Prexl et al., 2021) and Synthetic Aperture Radar (Saha et al., 2020) images. However, transfer learning based methods require pre-trained feature extractor (Saha and Zhu, 2021), which is not always available for hyperspectral sensors, given their large variation in spectral coverage and band composition. This challenge is tackled in (Saha et al., 2021) by using untrained deep models, merely initialized with a weight initialization strategy (He et al., 2015), as feature extractor for hyperspectral images. This approach is motivated from deep image prior in computer vision (Ulyanov et al., 2018). Being untrained, such models can be initialized to ingest as many number of bands as desired for processing the specific hyperspectral input. While such strategy has produced good result outperforming other unsupervised hyperspectral change detection methods, their change detection accuracy may vary if the weights of the untrained model are perturbed. Standard devi-

ation of up to 1% (over 5 different runs) is observed in some datasets in (Saha et al., 2021). Such variance/uncertainty may limit its trustworthiness in practical applications. Instead of considering such variance as a pitfall, we propose a strategy that can exploit the differences in the different CD maps to generate a more accurate CD map. Our proposed unsupervised strategy produces different change detection maps by slightly perturbing weights of the feature extractor and subsequently fuses the different change detection maps in a way that retains the relevant change information from all change maps. Fusion is accomplished by splitting the analyzed scene into small patches and then finding out the optimum CD map for each patch. The proposed approach not only reduces the uncertainty of the CD maps produced in (Saha et al., 2021), it also improves the CD performance by combining meaningful information from different maps.

The main contributions of this work are as follows:

1. We propose a fusion-based strategy that can combine different hyperspectral CD maps at decision level. The fusion strategy is unsupervised that does not make any assumption about the original models that have produced the CD maps.

2. We propose a strategy that circumnavigates the variance in change detection map produced by untrained models in (Saha et al., 2021).

3. We experimentally validate the proposed method for two different hyperspectral datasets.

We organize the rest of the paper as following. Related works are presented in Section 2. Proposed decision fusion based method is presented in Section 3. Results are presented in Section 4. The paper is concluded in Section 5.

---

* Corresponding author

2. RELATED WORKS

There are only few deep learning based methods for change detection in the hyperspectral images (Wang et al., 2018, Chen and Zhou, 2019, Ou et al., 2022, Shi et al., 2022). A pre-classification based method is proposed in (Wang et al., 2018). Such pre-classification strategy is also employed in (Ou et al., 2022) where a feature fusion grouping is used to generate a more discriminative feature group. A supervised CD method based on a joint affinity tensor is proposed in (Chen and Zhou, 2019). However, supervised CD methods cannot be adapted conveniently from one hyperspectral sensor to another because of differences in the band composition. Popular unsupervised CD methods like Deep Change Vector Analysis (DCVA) (Saha et al., 2019) can apply a model trained using multi-spectral images on the subset of bands of hyperspectral images. In computer vision, Ulyanov *et. al.* (Ulyanov et al., 2018) showed that significant portion of the image statistics are captured by the structure of the network itself. Based on this, untrained models are used in conjunction with DCVA in (Saha et al., 2021), which is called Deep Prior. While this method makes it possible to effectively use unsupervised CD on hyperspectral images, it may produce different results for slightly different initialization of the weights of the untrained model. Such variation in result potentially reduces its reliability. Furthermore, the question arises whether results from different combinations can be combined to produce a more accurate change detection map.

Decision fusion has long been used in remote sensing (Zhang, 2010). This topic is often presented in the context of multi-sensor data fusion. As an example, a deep learning decision fusion approach is proposed in (Abdi et al., 2018) for multi-sensor urban remote sensing data classification. Another context for applying decision fusion is when combining result from multiple models (Ma et al., 2019).

Our work takes forward the hyperspectral change detection method in (Saha et al., 2021) by reducing the uncertainty in its result and further improving its change detection map. Our method can be considered as a fusion of multiple models, where networks produced by different weight initialization can be considered to be different models.

3. METHOD

Given X_1 and X_2, a pair of co-registered hyperspectral images, our task is to segregate the changed pixels (Ω_c) in the image-pair from the unchanged ones (ω_{nc}). This task is popularly called binary CD in the literature. In this work, we do not focus on multiple/multi-class CD.

We choose a set of deep models with such an architecture that it can ingest the number of bands of the hyperspectral image pair. The weights of these deep models are initialized with a suitable technique (He et al., 2015). Each of these networks are separately used to obtain a CD map. Following this, a decision fusion scheme is used to integrate the results obtained using different models.

3.1 Deep models

Our hyperspectral inputs have B_0 channels, much larger than usual (3 or 4) bands in the multi-spectral images. This makes the models trained on multi-spectral input not suitable to be applied on the hyperspectral input. Towards this, we choose deep

model architectures in such a way that it is capable to ingest input with B_0 channels. The first convolution layer takes this input and projects it to a kernel size of $\beta_0 * B_0$. We simply use value of 4 for β_0, however any other value can be used. The following convolution layers preserve the number of kernels. Non-linearity is introduced by employing Rectified Linear Unit (ReLU) between the convolution layers. We do not introduce any such layer that may lead to downsampling (e.g., pooling). To summarize, our models follow the model architecture in (Saha et al., 2021). However, where only one model is used in (Saha et al., 2021), we use a number of models with the same architecture. The number of models (M) is chosen as 5, however any other value can be used. Weights of each model are initialized with a He initialization strategy (He et al., 2015), however with different random seeds in PyTorch (Paszke et al., 2019). Any other mechanism to introduce little differences in the weights of each model could be used. In this way, M different versions of an uninitialized deep network is obtained.

3.2 Change detection from a single model

In this Section, we describe the change detection process assuming just one model (out of M). The input X_1 and X_2 are pre-processed to have values between 0 and 1. An untrained model is applied on them (separately) to obtain deep features for each pixel. As the same model is applied on both input, similar deep features are obtained for a pixel unchanged in X_1 and X_2. On the other hand, a changed pixel tend to produce dissimilar deep features. By taking differences of deep features we obtain a hypervector (G) for each pixel. Following this, a magnitude (or one-dimensional index) ρ is computed from G. This is accomplished simply using Euclidean norm. Following this, we use Otsu's thresholding (Otsu, 1979) to group the pixels into two groups, one with comparatively higher ρ value, which corresponds to the changed pixels (Ω_c). The other group has comparatively smaller ρ value and corresponds to the unchanged pixels (ω_{nc}). The process of obtaining CD map using a single model is shown in Figure 1.

3.3 Decision fusion

A CD map is obtained for each untrained model. In other words, we have M different CD maps $C_1, ...,C_m,...,C_M$, each of size $R \times C$. Our task is to fuse them to obtain a single CD map. One particular CD map may not be the most accurate for the entire analyzed scene. Based on this assumption, we divide the analyzed scene of spatial scene $R \times C$ into smaller patches of size $R' \times C'$. We postulate that some of the M CD maps may be more accurate in some patches, while others may be more accurate for other patches.

For each patch, our task it to find the best CD map out of the M options ($C_1, ...,C_m,...,C_M$). The task is challenging as we need to find the best/optimum CD map in an unsupervised way. In contrary to the supervised tasks, we do not possess any well-defined oracle that can tell us about the suitability of the CD maps. Variance computation has previously been used to find suitable features in change detection (Saha et al., 2019). Features with the higher variances are shown to possess more relevant information for change detection. Variance has also been used for uncertainty computation in semantic segmentation (Rottmann and Schubert, 2019). Inspired by these, we postulate that the CD map with the highest variance score is the optimum choice for a particular patch. The variance is computed on the ρ values. An ideal model will produce ρ values
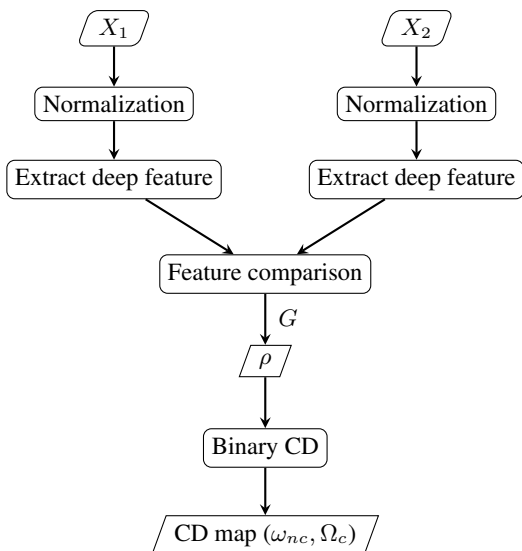
Figure 1. Method for obtaining CD map using a single untrained model. The $\rho, \omega_{nc}, \Omega_c$ are used for decision fusion while combining multiple-trained models.
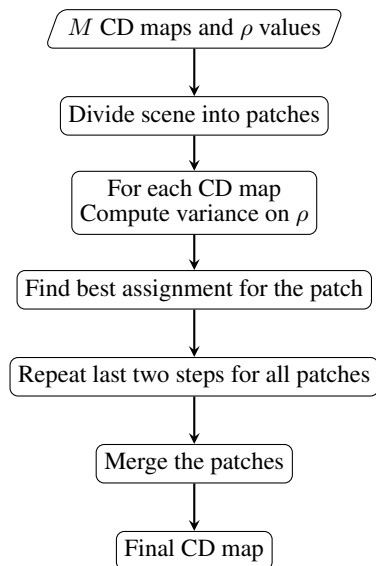

Figure 2. Method for combining results from multiple models.

which are either concentrated on the very small values (for unchanged pixels) or very large values (for changed pixels), thus producing higher variance score. Thus the optimum model for a particular patch can be determined by maximizing variance computed on ρ values.

The CD map with optimum variance score is assigned to the final CD map for the particular patch. In this way, optimum assignments are determined for all patches and the CD map for the entire scene is obtained. This process of combining CD maps from multiple models is shown in Figure 2.

4. EXPERIMENTS AND ANALYSIS

4.1 Datasets

Following publicly available bi-temporal hyperspectral datasets are used for validation (López-Fandiño et al., 2018, López-

Fandiño et al., 2019) [1]:

1. The pre-change and post-change images in the Santa Barbara dataset are acquired on 2013 and 2014. The 984 × 740 pixel images show 224 spectral bands.

2. The pre-change and post-change images in the Bay Area dataset are acquired on 2013 and 2015. In this case, the spatial dimension of the analyzed scene is 500 × 500 pixels.

4.2 Compared methods

We compare the following methods:

1. Deep Prior, as in (Saha et al., 2021). Comparison to this method is essential as the proposed method extends this.

2. CVA, a popular unsupervised approach (Malila, 1980).

3. DCVA, a popular deep transfer learning based unsupervised approach (Saha et al., 2019).

4. Autoencoding of bi-temporal Hyperspectral Images for Change Vector Analysis (AICA) (Appice et al., 2019).

4.3 Results

In this work, we used $M = 5$, keeping consistency with (Saha et al., 2021) where results are shown as average of 5 runs. One important parameter is the patch length. For the Santa Barbara scene, we show the variation of the result of the proposed method versus patch length in Table 1. Best sensitivity (accuracy of changed pixels) and accuracy values are obtained for patch length 350 pixels. Specificity (accuracy of unchanged pixels) shows little variation as the patch size is varied.

For the Santa Barbara scene, quantitative comparison with state-of-the-art methods is shown in Table 2. The proposed method outperforms existing unsupervised paradigms like CVA, AICA (Appice et al., 2019), DCVA (Saha et al., 2019). It is evident that the proposed method can benefit from the multiple runs of randomly initialized networks and produces superior sensitivity, similar specificity, and superior accuracy than (Saha et al., 2021), without any uncertainty/variance. While the improvement in accuracy achieved by the proposed method over Deep Prior is only 0.68%, it is important to understand that the Deep Prior shows a deviation of 0.6% over 5 different runs and no mechanism is proposed in (Saha et al., 2021) to decide which run is more accurate than the others. On the other hand, our method can choose which run is more accurate for which patch and integrate them into a single result without any uncertainty. Results obtained using different methods are shown in Figure 3.

The variation of the result of the proposed method versus patch length for the Bay Area scene is shown in Table 3. Due to the smaller size of the Bay Area scene, we vary the patch length only up to 250 pixels. Best result in terms of sensitivity and accuracy is obtained using patch length 150 pixels. Quantitative comparison for the Bay Area scene is shown in Table 4. In this case also the proposed method yields superior sensitivity, slightly lower specificity, and superior accuracy than (Saha et al., 2021). The proposed method also outperforms CVA, AICA and DCVA. Reference map and the result obtained using the proposed method are visualized in Figure 4.

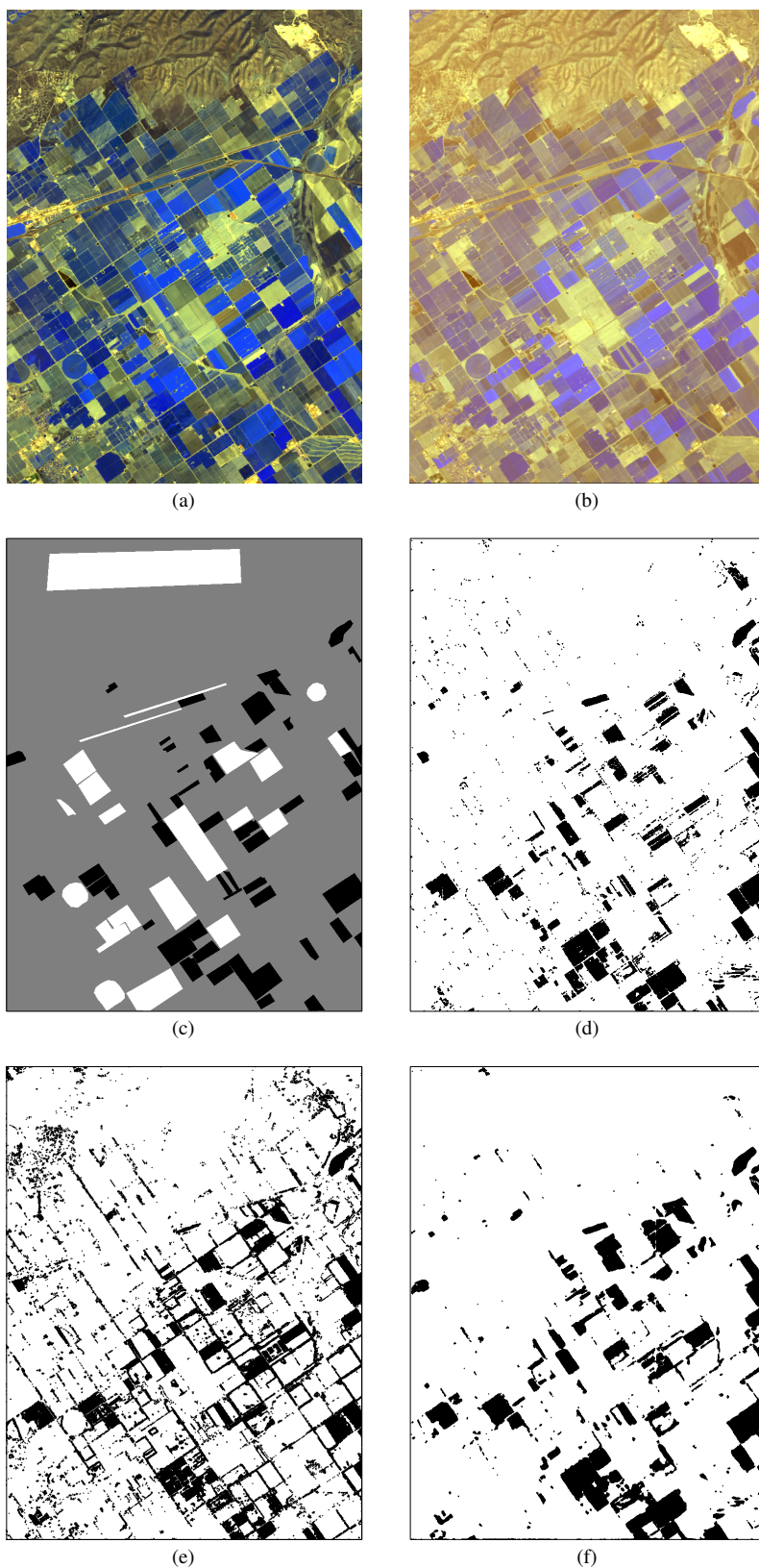[1] Datasets: https://citius.usc.es/investigacion/datasets/hyperspectral-change-detection-dataset

Figure 3. Visualization Of the Santa Barbara scene: (a) pre-change image, (b) post-change image, (c) reference image (black pixels are changed, white pixels are unchanged, and gray pixels are unknown), and CD maps: (d) CVA, (e) DCVA, (f) Proposed approach, i.e., fusion of the untrained models.
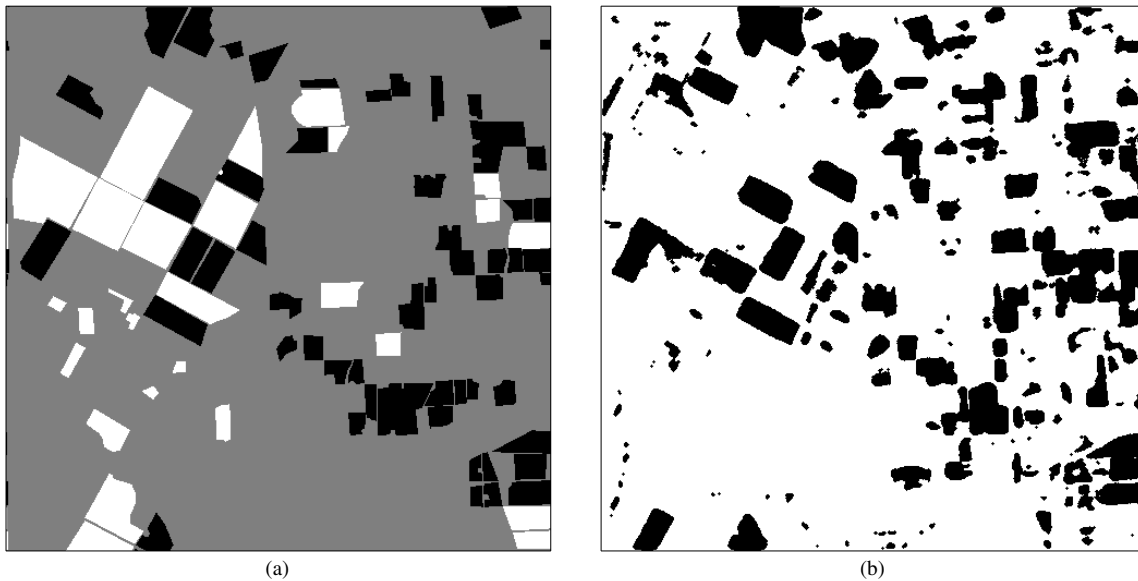
Figure 4. Bay Area scene: (a) reference image (white - unchanged, black - changed, gray - unknown), and (b) CD map obtained by the proposed method.

Patch length	Sensitivity	Specificity	Acc.
100	88.46	98.62	94.63
150	89.00	98.52	94.77
200	88.27	98.48	94.46
250	89.00	98.53	94.78
300	89.34	98.48	94.88
350	89.64	98.59	95.08
400	88.32	98.59	94.55

Table 1. CD results for the Santa Barbara scene using proposed method as the patch length is varied.

Method	Sensitivity	Specificity	Acc.
AICA	87.25	94.52	91.66
CVA	76.92	96.69	88.91
DCVA	78.01	93.60	87.47
Deep Prior	87.98	98.57	94.40±0.6
Proposed	89.64	98.59	95.08

Table 2. Quantitative comparison for the Santa Barbara scene.

Patch length	Sensitivity	Specificity	Acc.
100	79.23	97.49	88.08
150	79.92	97.41	88.40
200	79.60	97.47	88.26
250	79.53	97.40	88.20

Table 3. CD results for the Bay Area scene using proposed method as the patch length is varied.

Method	Sensitivity	Specificity	Acc.
AICA	69.18	97.26	82.80
CVA	74.44	97.54	85.64
DCVA	78.27	92.47	85.16
Deep Prior	78.51	97.86	87.89±1
Proposed	79.92	97.41	88.40

Table 4. Quantitative comparison for the Bay Area scene.

5. CONCLUSION

A model trained on one hyperspectral sensor cannot be reused on another sensor due to the differences in band composition. To alleviate this, previous works (Saha et al., 2019) showed that an untrained deep network, initialized with some weight initialization technique, can be used for hyperspectral CD. However, results may vary if untrained networks are slightly perturbed/modified. To circumnavigate this drawback and to exploit different CD maps that can be obtained using different versions of an untrained network, this work uses a decision fusion method by using patch-wise variance measure. The proposed method clearly improves the result, especially sensitivity to the changed pixels. In future, we will study the potential of replacing variance-based method by information theoretic approaches. We will also extend the method for multi-sensor hyperspectral change detection where pre-change and post-change images are acquired using different hyperspectral sensors.

ACKNOWLEDGEMENTS

REFERENCES

Abdi, G., Samadzadegan, F., Reinartz, P., 2018. Deep learning decision fusion for the classification of urban remote sensing data. *Journal of Applied Remote Sensing*, 12(1), 016038.

Appice, A., Di Mauro, N., Lomuscio, F., Malerba, D., 2019. Empowering change vector analysis with autoencoding in bitemporal hyperspectral images. *CEUR Workshop Proceedings*, 2466.

Cao, Q., Yu, G., Sun, S., Dou, Y., Li, H., Qiao, Z., 2021. Monitoring Water Quality of the Haihe River Based on Ground-Based Hyperspectral Remote Sensing. *Water*, 14(1), 22.

Chen, Z., Zhou, F., 2019. Multitemporal hyperspectral image change detection by joint affinity and convolutional neural networks. *2019 10th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp)*, IEEE, 1–4.

Chevrel, S., Kopackova, V., Fischer, C., Dor, E. B., Adar, S., Shkolnisky, Y., Misurec, J., 2012. Mapping minerals, vegetation health and change detection over the sokolov lignite mine using multidate hyperspectral airborne imagery. *1st EAGE/GRSG Remote Sensing Workshop*, European Association of Geoscientists & Engineers, cp–305.

He, K., Zhang, X., Ren, S., Sun, J., 2015. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *Proceedings of the IEEE international conference on computer vision*, 1026–1034.

Hirano, A., Madden, M., Welch, R., 2003. Hyperspectral image data for mapping wetland vegetation. *Wetlands*, 23(2), 436–448.

Kruse, F. A., Boardman, J. W., Huntington, J. F., 2003. Comparison of airborne hyperspectral data and EO-1 Hyperion for mineral mapping. *IEEE transactions on Geoscience and Remote Sensing*, 41(6), 1388–1400.

Li, F., Song, M., Chi, J., Cheng, Y., 2021. Ship velocity estimation via images acquired by an unmanned aerial vehicle-based hyperspectral imaging sensor. *Journal of Applied Remote Sensing*, 15(3), 032206.

López-Fandiño, J., Garea, A. S., Heras, D. B., Argüello, F., 2018. Stacked autoencoders for multiclass change detection in hyperspectral images. *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, 1906–1909.

López-Fandiño, J., Heras, D. B., Argüello, F., Dalla Mura, M., 2019. GPU framework for change detection in multitemporal hyperspectral images. *International Journal of Parallel Programming*, 47(2), 272–292.

Ma, W., Guo, Q., Wu, Y., Zhao, W., Zhang, X., Jiao, L., 2019. A novel multi-model decision fusion network for object detection in remote sensing images. *Remote Sensing*, 11(7), 737.

Malila, W. A., 1980. Change vector analysis: an approach for detecting forest changes with Landsat. *LARS symposia*, 385.

Mou, L., Saha, S., Hua, Y., Bovolo, F., Bruzzone, L., Zhu, X. X., 2021. Deep reinforcement learning for band selection in hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–14.

Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1), 62–66.

Ou, X., Liu, L., Tu, B., Zhang, G., Xu, Z., 2022. A CNN Framework with Slow-Fast Band Selection and Feature Fusion Grouping for Hyperspectral Image Change Detection. *IEEE Transactions on Geoscience and Remote Sensing*.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L. et al., 2019. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32.

Prexl, J., Saha, S., Zhu, X. X., 2021. Mitigating spatial and spectral differences for change detection using super-resolution and unsupervised learning. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, IEEE, 3113–3116.

Rottmann, M., Schubert, M., 2019. Uncertainty measures and prediction quality rating for the semantic segmentation of nested multi resolution street scene images. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 0–0.

Saha, S., Bovolo, F., Bruzzone, L., 2019. Unsupervised deep change vector analysis for multiple-change detection in VHR images. *IEEE Transactions on Geoscience and Remote Sensing*, 57(6), 3677–3693.

Saha, S., Bovolo, F., Bruzzone, L., 2020. Building Change Detection in VHR SAR Images via Unsupervised Deep Transcoding. *IEEE Transactions on Geoscience and Remote Sensing*.

Saha, S., Kondmann, L., Song, Q., Zhu, X. X., 2021. Change Detection in Hyperdimensional Images Using Untrained Models. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 11029–11041.

Saha, S., Zhu, X. X., 2021. Patch-level unsupervised planetary change detection. *IEEE Geoscience and Remote Sensing Letters*, 19, 1–5.

Shi, J., Zhang, Z., Wu, T., Zhang, X., Lei, Y., 2022. Collaborative Self-Perception Network Architecture for Hyperspectral Image Change Detection. *IEEE Geoscience and Remote Sensing Letters*.

Ulyanov, D., Vedaldi, A., Lempitsky, V., 2018. Deep image prior. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 9446–9454.

Wang, Q., Yuan, Z., Du, Q., Li, X., 2018. GETNET: A general end-to-end 2-D CNN framework for hyperspectral image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 57(1), 3–13.

Zhang, J., 2010. Multi-source remote sensing data fusion: status and trends. *International Journal of Image and Data Fusion*, 1(1), 5–24.