

Attention Mechanism for Remote Sensing

Subjects: Computer Science, Artificial Intelligence | Remote Sensing

Contributor: Saman Ghaffarian

Machine learning, particularly deep learning (DL), has become a central and state-of-the-art method for several computer vision applications and remote sensing (RS) image processing. Researchers are continually trying to improve the performance of the DL methods by developing new architectural designs of the networks and/or developing new techniques, such as attention mechanisms. Since the attention mechanism has been proposed, regardless of its type, it has been increasingly used for diverse RS applications to improve the performances of the existing DL methods.

Keywords: remote sensing ; image processing ; attention mechanism ; spatial attention ; channel attention ; deep learning ; CNN ; image classification ; change detection ; object detection ; image fusion ; image segmentation ; meta analysis ; review

1. Introduction

Remotely sensed images have been employed as the main data sources in many fields such as agriculture ^{[1][2][3][4]}, urban planning ^{[5][6][7]} and disaster risk management ^{[8][9][10]}, and have been shown as an effective and critical tool to provide information. Accordingly, processing remote sensing (RS) images is crucial to extract the useful information from them for such applications. RS image processing tasks include image classification, object detection, change detection, and image fusion ^[11]. Different processing methods were developed to address them, and they aimed to improve the performance and accuracy of the methods to address RS image processing. Machine learning methods such as support vector machines and ensemble classifiers (e.g., random forest and gradient boosting) obtained fairly high accuracies for different RS processing tasks ^{[12][13]}. In particular, deep learning (DL) methods have recently become state-of-the-art methods in RS image processing and automatically extracting the required information from RS images ^{[14][15]}. Since DL has entered this field, researchers try to improve the performance and increase its accuracy by developing new techniques and different architectural designs, e.g., various convolutional neural networks (CNN) ^{[16][17]}, generative adversarial networks (GAN) ^[18], graph neural networks (GNN) ^[19]. Recently, the attention mechanism was proposed by Bahdanau, et al. ^[20] initially for machine translation application, which aims to guide deep neural network methods by providing focus points and highlighting the important features while minimizing the others. Thereafter, it was used in different applications, including computer vision ^[21] and RS image processing ^{[22][23][24]}. Accordingly, most of the studies reported an increase in the performance of the DL methods when guided with attention mechanism ^{[25][26][27]}.

In recent years, researchers reviewed the developed/used DL methods in RS literature mostly from a general perspective ^{[11][28]} or focusing on one application, e.g., image classification ^[15]. Zhang, et al. ^[14] reviewed the DL methods in RS big data processing and provided a technical tutorial on the state-of-the-art methods. Zhu, et al. ^[28] reviewed the DL methods applied to RS data analysis and investigated the challenges of DL in RS applications. They also provided a comprehensive list of resources for DL-RS data analysis. Li, et al. ^[15] conducted a survey study on the developed DL methods for RS image classification. They also analyzed and compared the performance of the different DL methods. In addition, the recent advances in DL for pixel-level image fusion were reviewed by Li, et al. ^[29]. Ma, et al. ^[11] conducted a systematic literature review on applications of the DL on RS and they comprehensively reviewed and categorized DL methods. In addition, Niu, et al. ^[21] reviewed the different architectural designs of the attention mechanism used in conjunction with DL from a general perspective and provided some application domains. However, the effect of such a mechanism for DL methods in RS image processing has not yet been reviewed and investigated. Accordingly, a systematic literature review is conducted in this study by following a structured review on the DL methods with an embedded attention mechanism for RS image processing applications. Thus, the literature is reviewed systematically to respond to the predefined research questions rather than summarizing the papers. The main objective of this study is to extract the effect of attention mechanism in the performance of deep learning-based RS (DL-RS) image processing. In addition, the current trends, achievements and applications in publications, using attention mechanism-based DL (At-DL) methods and RS image processing applications are extracted to provide insights and guidelines for future studies.

The rest of the paper is organized as follows. Background information regarding the attention mechanism, its different types, and how it is being used in DL methods are provided in Section 2. Section 3 presents and describes the integration of attention mechanisms with different deep neural network architectures to address RS image processing tasks. The steps of the executed systematic literature review are explained in Section 4. Then, Section 5 presents and visualizes the quantitative results and discusses them according to the defined research questions, and reveals the effect of attention mechanism in the performance of the DL-RS image processing. Finally, Section 6 concludes the paper.

2. Attention Mechanism in Deep Learning

The attention mechanism, like other neural network-based methods, tries to mimic the human brain/vision to process data. Human vision does not process the entire image at once; however, it only focuses on the specific parts. With this, the focused parts of the human view space are perceived in “high-resolution” while the surroundings are in “low-resolution”. In other words, it gives higher weight to the relevant parts while minimizing the irrelevant ones, giving them lower weights. This allows the brain to process and focus on the most important parts precisely and efficiently, rather than processing the entire view space. This characteristic of human vision inspired researchers to develop the attention mechanism. It was initially developed in 2014 for natural language processing applications [20], since then it has been widely used for different applications [30], in particular, computer vision tasks [21][31]. Its potential to enhance mostly CNN-based methods has been reported [32]. In addition, it has been used in conjunction with recurrent neural network models [33][34][35][36], and graph neural networks [37][38]. The main idea behind the attention mechanism is to give different weights to different information. Thus, giving higher weights to relevant information attracts the attention of the DL model to them [39]. Attention mechanism approaches can be grouped based on four criteria (Figure 1) [21]:

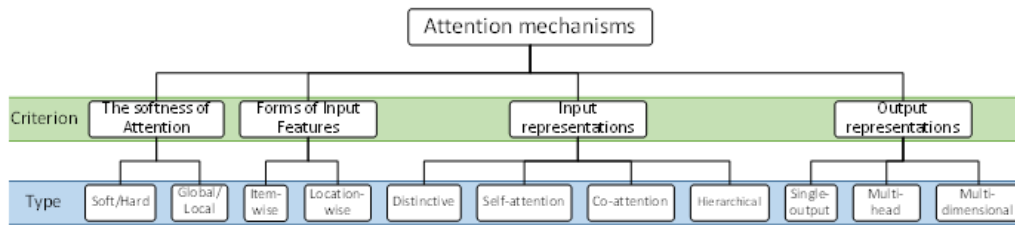


Figure 1. An overview of typical attention mechanism approaches [21].

(i) The softness of attention: the initial attention mechanism proposed by [20] is a soft version, which is also known as deterministic attention. This network considers all input elements (computes the average for each weight) to compute the final context vector. The context vector is the high-dimensional vector representation of the input elements or sequences of the input elements and in general the attention mechanism aims to add more contextual information to compute the final context vector. However, hard attention, which is also known as stochastic attention, randomly selects from the sample elements to compute the final context vector [40]. This, therefore, reduces the computational time. Furthermore, there is another categorization that is frequently used in computer vision tasks and RS image processing, i.e., global and local attentions [41][42]. Global attention is similar to soft attention since it also considers all input elements. However, global attention simplifies soft attention by using the output of the current time step rather than the prior one, while local attention is a combination of soft and hard attentions. This approach considers a subset of input elements at a time, and thus, overcomes the limitation of hard attention, i.e., being nondifferentiable, and in the meantime is less computationally expensive. (ii) Forms of input features: attention mechanisms can be grouped based on their input requirements: item-wise and location-wise. Item-wise attention requires inputs that are known to the model explicitly or produced with a preprocess [43][44][45]. However, location-wise attention does not necessarily require known inputs, in this case, the model needs to deal with input items that are difficult to distinguish. Due to the characteristics and features of the RS images and targeted tasks, location-wise attention is commonly used for RS image processing [42][46][47][48]. (iii) Input representations: there are single-input and multi-input attention models [49][50]. In addition, the general processing procedure of the inputs also varies between the developed models. Most of the current attention networks work with single-input, and the model processes them in two independent sequences (i.e., distinctive model). The co-attention model is a multi-input attention network that parallelly implements the attention mechanism on two different sources but finally merges them [50]. This makes it suitable for change detection from RS images [51]. A self-attention network computes attentions only based on the model inputs, and thus, it decreases the dependence on external information [52][53][54]. This allows the model to perform better in images with complex background by focusing more on targeted areas [55]. Hierarchical attention mechanism computes weights from the original input and different levels/scales of the inputs [56]. This attention mechanism is also known as fine-grained attention for image classification [57]. (iv) Output representations: single-output is the commonly used output representation in attention mechanisms. It processes a single feature at a time and computes weight scores. There are also two other multidimensional and multi-head attention mechanisms [21]. Multi-head

attention processes the inputs linearly in multiple subsets, and finally merges them to compute the final attention weights [58], and is especially useful when employing the attention mechanism in conjunction with CNN methods [59][60][61]. Multidimensional attention, which is mostly employed for natural language processing, computes weights based on matrix representation of the features instead of vectors [62][63].

The above-explained attention mechanisms are the same in principle and are developed by researchers to adopt or improve the basic attention mechanism for their tasks. In addition, not all of them have been used for computer vision, and thus, RS image processing. In DL-based image processing, this mechanism is usually used to focus on specific features (feature layers) or a certain location or aspect of an image [64][65][66][67]. Accordingly, it can be classified into two major types: channel and spatial attentions.

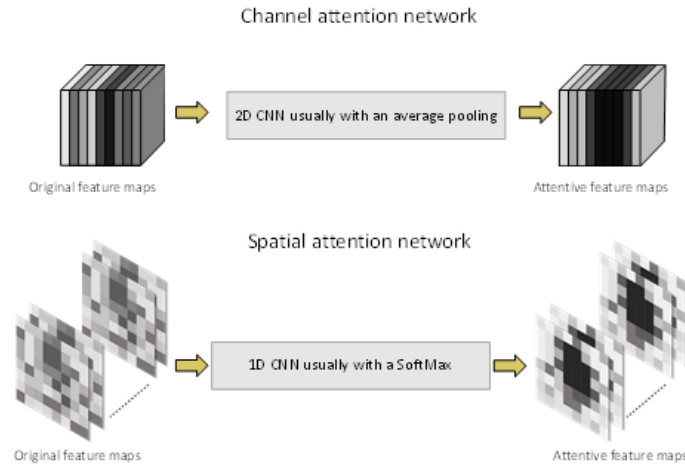


Figure 2. A simple illustration of the channel and spatial attention types/networks, and their effects on the feature maps.

Figure 2 illustrates simple channel and spatial attention types: (a) The channel attention network aims to boost the feature layers (channel) in the feature map that convey more important information and silence the other feature layers (channels) ; (b) the spatial attention network highlights regions of interest in the feature space and covers up the background regions. These two attention mechanisms can be used solely or combined within DL methods to provide attention to both important feature layers and the location of the region of interest. Papers in this review were classified according to these two types.

3. Deep Neural Network Architectures with Attention for RS Image Processing

In this section, we describe and provide examples of the four different deep neural network architectures (i.e., CNN, GAN, RNN, and GNN) that are improved using the attention mechanism to address RS image processing. CNN is the main method that has been used for image processing in general, as well as RS applications. Both spatial and channel attentions are embedded in CNN with different attention network designs. For CNNs the channel attention is typically implemented after each convolution but the spatial attention is mostly added to the end of the network [68][69][70][71]. However, in UNet-based networks, spatial attention is usually added to each layer of a decoding/upsampling section [72][73][74]. **Figure 3** shows an example of using spatial and channel attentions, in particular co-attention network, in a Siamese model for building-based change detection [51]. The proposed co-attention network is based on an initial correlation process with a final attention module. For GAN networks which are based on encoding and decoding modules, the process of adding attention networks is the same as of CNNs that can be used in both adversarial and/or discrimination networks depending on the targeted tasks [75] (**Figure 4**).

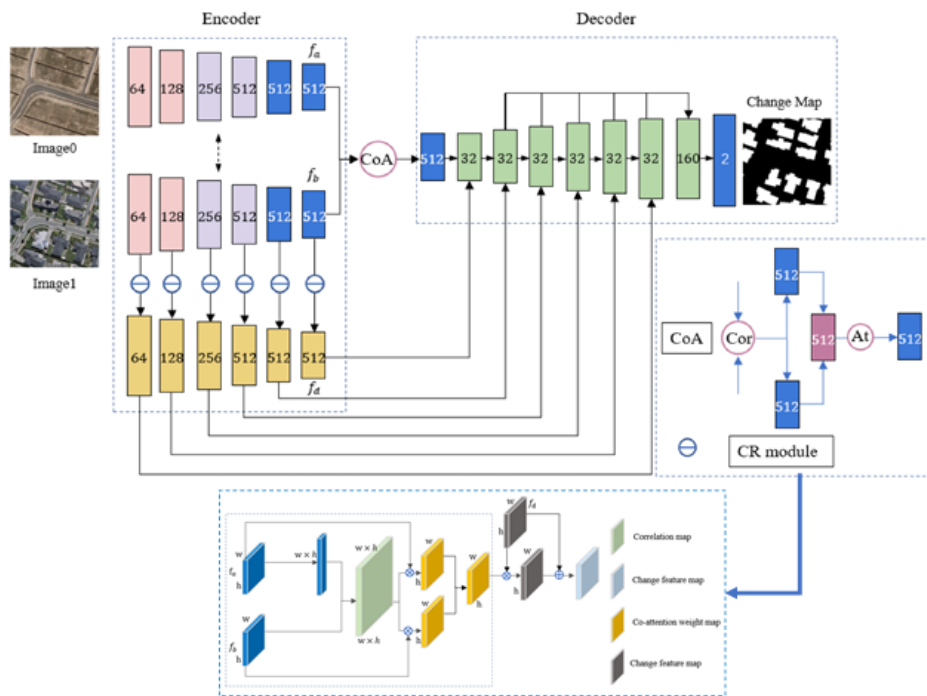


Figure 3. An example of adding attention network (i.e., Co-attention) to a CNN module (i.e., Siamese network) for building-based change detection [51]. CoA – Co-Attention module, At – Attention network, CR – Change Residual module.

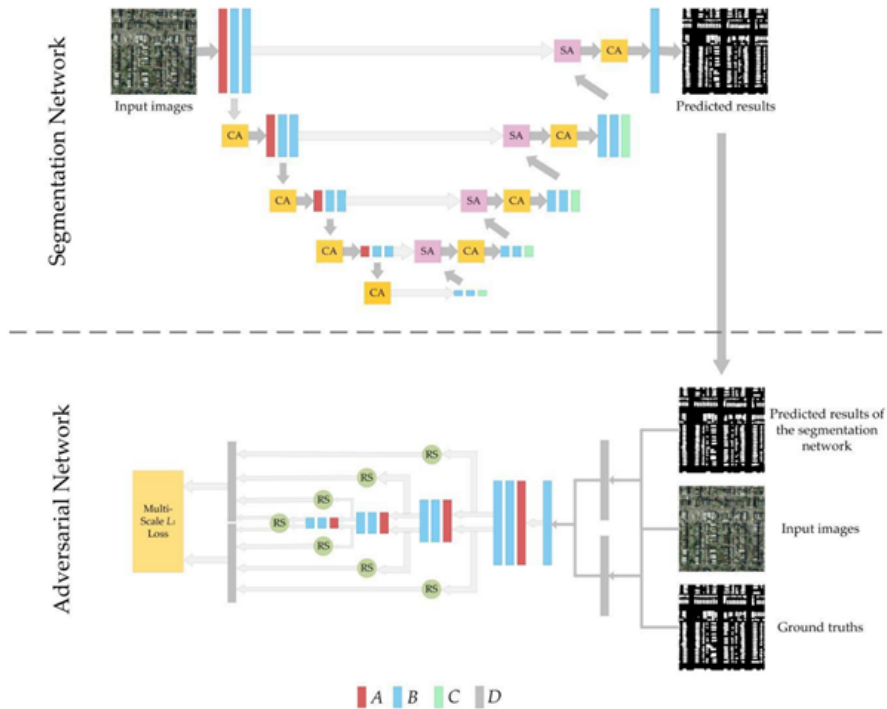


Figure 4. An example of adding spatial and channel attentions to a GAN module for building detection from aerial images [75]. A – max pooling layer; B – convolutional + batch normalization + rectified linear unit (ReLU) layers; C – upsampling layer; D – concatenation operation; SA – spatial attention mechanism; CA – channel attention mechanism; RS – reshape operation.

RNN is the first deep learning network that is improved by attention mechanism [20] for natural language processing tasks. RNNs are not as popular as CNNs for image processing due to the inherent characteristics of the images. However, RNN has been frequently used in conjunction with CNN for RS image processing [34][76][77][78]. This also allows the integration of the attention mechanism with RNN for RS applications. For example, Ref. [79] developed a bidirectional RNN module to provide channel attention and add the outcome weights to the CNN-based module which is supported with a spatial attention network for hyperspectral image classification (Figure 5).

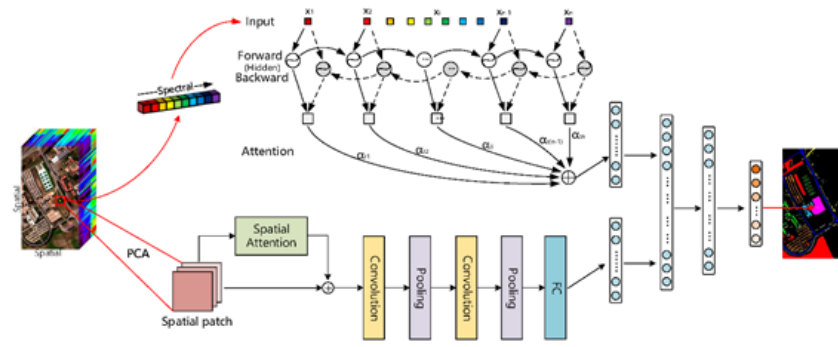


Figure 5. An example of adding attention networks (i.e., spatial and channel attentions) to a RNN+CNN module for hyperspectral image classification [79]. PCA – Principal Component Analysis.

GNN is another network architecture that has been employed in conjunction with CNN for RS image processing. Hence, this mechanism is used to focus on the most important graph nodes of the network. A typical integration of GNN with CNN is to implement a GNN after a CNN-based image segmentation to produce the final RS image classification results [80][81]. Accordingly, the attention network adjusts the weight for each graph node through the graph convolutional layers (**Figure 6**) [82].

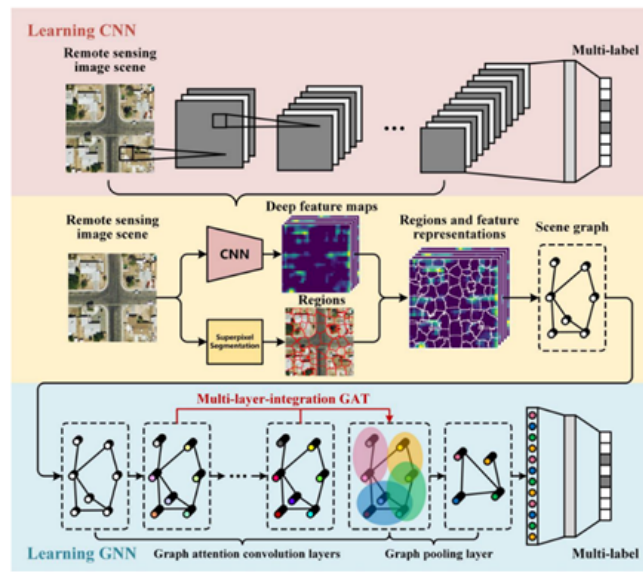


Figure 6. An example of adding a attention network to a GNN module for multi-label RS image classification [82].

4. Results and Discussion

At-DL methods entered RS image processing in 2018, while attention mechanism was developed in 2014 [20]. However, only since 2020, have most studies (i.e., 141 papers) employed this technique for different RS image processing applications, which reveals a significant interest in the technique in recent years (**Figure 7**). Just in 2021, 47 papers were published, knowing that the searches from the online databases were conducted in March 2021.

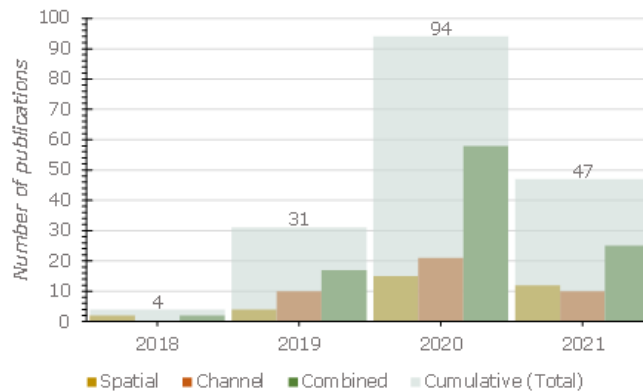


Figure 7. Year-wise classification of the papers and classified based on the used attention mechanism type.

Figure 8 shows the number of papers that employed the attention mechanism for each DL algorithm. Accordingly, the convolutional neural networks (CNN) algorithm is the predominant DL method that was enhanced with an attention mechanism to address RS image processing, which applied in 154 out of 176 reviewed papers [69][83][84][85][86][87]. This is an expected result since CNN is the most frequently used DL method in general computer vision and image processing. Recurrent neural networks (RNN), such as long-short term memories (LSTM) methods, were the second most frequently used DL method supported by attention mechanism for RS image processing with 18 papers [88][89][90], this algorithm is also the first DL method that was improved with attention mechanism [20]. In addition, it was observed that most of the RNN methods were used in combination with CNN methods [76][78][91]. Generative adversarial networks (GAN) [53][92][93], Graph Neural Network (GNN) [80][82], and other DL methods including capsule network [72] and autoencoders [61] were the other DL algorithms used in 12, 5, and 4 papers, respectively.

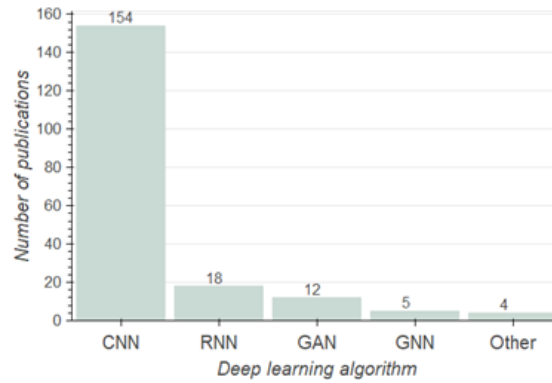


Figure 8. The improved DL algorithms with attention mechanism in the papers.

We investigated the performance of attention mechanism in DL methods for RS image processing in two manners; (i) by extracting the overall accuracies of the used At-DL methods for RS image processing tasks (**Figure 9**), and (ii) comparing the overall accuracies of the produced results with and without attention mechanism in the papers (**Figure 10**).

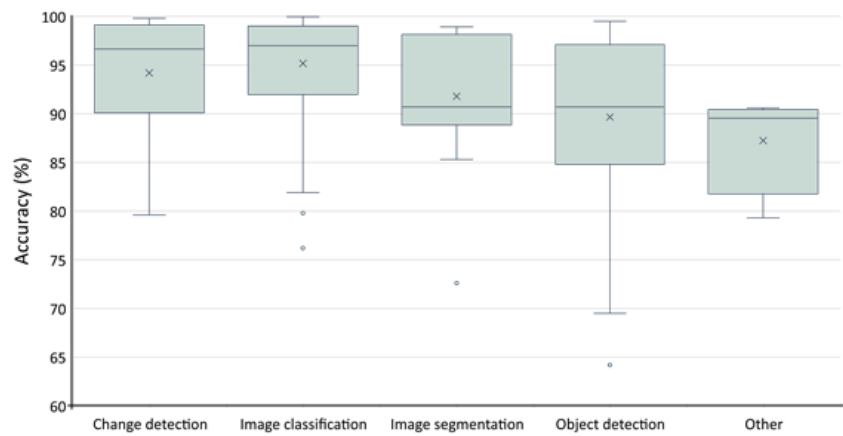


Figure 9. The produced accuracy of the developed At-DL methods for different tasks in the papers.

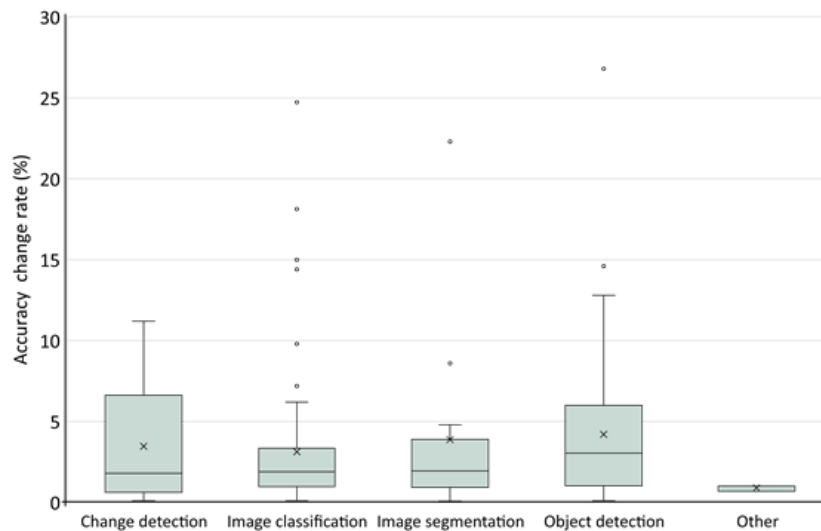


Figure 10. The effect of the used attention mechanism within the DL algorithms in terms of accuracy rate for different tasks in the papers.

External validity: This study reviewed the publications which employed At-DL methods for RS image processing applications. However, all of the existing DL methods have not been improved with attention mechanism or have not yet been used for RS image processing applications, and all the possible RS image processing applications were not addressed with At-DL and thus not included or discussed in this study. In addition, we only reviewed the publications that used At-DL for RS image processing applications, and thus, we cannot make judgments about the use and the effect of the At-DL in a broader scope or other applications.

References

1. Weiss, M.; Jacob, F.; Duveiller, G. Remote sensing for agricultural applications: A meta-review. *Remote Sens. Environ.* 2020, 236, 111402.
2. Ghaffarian, S.; Turker, M. An improved cluster-based snake model for automatic agricultural field boundary extraction from high spatial resolution imagery. *Int. J. Remote Sens.* 2019, 40, 1217–1247.
3. Valente, J.; Sari, B.; Kooistra, L.; Kramer, H.; Mùcher, S. Automated crop plant counting from very high-resolution aerial imagery. *Precis. Agric.* 2020, 21, 1366–1384.
4. Zhang, C.; Valente, J.; Kooistra, L.; Guo, L.; Wang, W. Orchard management with small unmanned aerial vehicles: A survey of sensing and analysis approaches. *Precis. Agric.* 2021.
5. Nielsen, M.M. Remote sensing for urban planning and management: The use of window-independent context segmentation to extract urban features in Stockholm. *Comput. Environ. Urban Syst.* 2015, 52, 1–9.
6. Kadhim, N.; Mourshed, M.; Bray, M. Advances in remote sensing applications for urban sustainability. *Euro-Mediterr. J. Environ. Integr.* 2016, 1, 7.
7. Ghaffarian, S.; Ghaffarian, S. Automatic building detection based on Purposive FastICA (PFICA) algorithm using monocular high resolution Google Earth images. *ISPRS J. Photogramm. Remote Sens.* 2014, 97, 152–159.
8. Ghaffarian, S.; Kerle, N.; Filatova, T. Remote Sensing-Based Proxies for Urban Disaster Risk Management and Resilience: A Review. *Remote Sens.* 2018, 10, 1760.
9. Ghaffarian, S.; Rezaie Farhadabad, A.; Kerle, N. Post-Disaster Recovery Monitoring with Google Earth Engine. *Appl. Sci.* 2020, 10, 4574.
10. Ghaffarian, S.; Emtehani, S. Monitoring Urban Deprived Areas with Remote Sensing and Machine Learning in Case of Disaster Recovery. *Climate* 2021, 9, 58.
11. Ma, L.; Liu, Y.; Zhang, X.; Ye, Y.; Yin, G.; Johnson, B.A. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* 2019, 152, 166–177.
12. Belgiu, M.; Drăguț, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* 2016, 114, 24–31.
13. Sheykhou, M.; Mahdianpari, M.; Ghanbari, H.; Mohammadimanesh, F.; Ghamisi, P.; Homayouni, S. Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2020, 13, 6308–6325.
14. Zhang, L.; Zhang, L.; Du, B. Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art. *IEEE Geosci. Remote Sens. Mag.* 2016, 4, 22–40.
15. Li, Y.; Zhang, H.; Xue, X.; Jiang, Y.; Shen, Q. Deep learning for remote sensing image classification: A survey. *WIREs Data Min. Knowl. Discov.* 2018, 8, e1264.
16. Kattenborn, T.; Leitloff, J.; Schiefer, F.; Hinz, S. Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. *ISPRS J. Photogramm. Remote Sens.* 2021, 173, 24–49.
17. Ghanbari, H.; Mahdianpari, M.; Homayouni, S.; Mohammadimanesh, F. A Meta-Analysis of Convolutional Neural Networks for Remote Sensing Applications. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2021, 14, 3602–3613.
18. Liu, X.; Wang, Y.; Liu, Q. Psgan: A Generative Adversarial Network for Remote Sensing Image Pan-Sharpening. In *Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP)*, Athens, Greece, 7–10 October 2018; pp. 873–877.
19. Yan, X.; Ai, T.; Yang, M.; Yin, H. A graph convolutional neural network for classification of building patterns using spatial vector data. *ISPRS J. Photogramm. Remote Sens.* 2019, 150, 259–273.

20. Bahdanau, D.; Cho, K.; Bengio, Y. Neural machine translation by jointly learning to align and translate. *arXiv* 2014, arXiv:1409.0473.
21. Niu, Z.; Zhong, G.; Yu, H. A Review on the Attention Mechanism of Deep Learning. *Neurocomputing* 2021.
22. Zhang, J.; Zhou, Q.; Wu, J.; Wang, Y.C.; Wang, H.; Li, Y.S.; Chai, Y.Z.; Liu, Y. A Cloud Detection Method Using Convolutional Neural Network Based on Gabor Transform and Attention Mechanism with Dark Channel Subnet for Remote Sensing Image. *Remote Sens.* 2020, 12, 3261.
23. Zeng, Y.L.; Ritz, C.; Zhao, J.H.; Lan, J.H. Attention-Based Residual Network with Scattering Transform Features for Hyperspectral Unmixing with Limited Training Samples. *Remote Sens.* 2020, 12, 400.
24. Yu, Y.; Li, X.; Liu, F. Attention GANs: Unsupervised Deep Feature Learning for Aerial Scene Classification. *IEEE Trans. Geosci. Remote Sens.* 2020, 58, 519–531.
25. Gao, F.; He, Y.S.; Wang, J.; Hussain, A.; Zhou, H.Y. Anchor-free Convolutional Network with Dense Attention Feature Aggregation for Ship Detection in SAR Images. *Remote Sens.* 2020, 12, 2619.
26. Li, F.; Feng, R.; Han, W.; Wang, L. High-Resolution Remote Sensing Image Scene Classification via Key Filter Bank Based on Convolutional Neural Network. *IEEE Trans. Geosci. Remote Sens.* 2020, 58, 8077–8092.
27. Yang, H.; Wu, P.H.; Yao, X.D.; Wu, Y.L.; Wang, B.; Xu, Y.Y. Building Extraction in Very High Resolution Imagery by Dense-Attention Networks. *Remote Sens.* 2018, 10, 1768.
28. Zhu, X.X.; Tuia, D.; Mou, L.; Xia, G.; Zhang, L.; Xu, F.; Fraundorfer, F. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geosci. Remote Sens. Mag.* 2017, 5, 8–36.
29. Li, S.; Kang, X.; Fang, L.; Hu, J.; Yin, H. Pixel-level image fusion: A survey of the state of the art. *Inf. Fusion* 2017, 33, 100–112.
30. Galassi, A.; Lippi, M.; Torroni, P. Attention in Natural Language Processing. *IEEE Trans. Neural Netw. Learn. Syst.* 2020, 1–18.
31. Koščević, K.; Subašić, M.; Lončarić, S. Attention-based Convolutional Neural Network for Computer Vision Color Constancy. In *Proceedings of the 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA)*, Dubrovnik, Croatia, 23–25 September 2019; pp. 372–377.
32. Li, W.; Liu, K.; Zhang, L.; Cheng, F. Object detection based on an adaptive attention mechanism. *Sci. Rep.* 2020, 10, 11307.
33. Cui, W.; Wang, F.; He, X.; Zhang, D.Y.; Xu, X.X.; Yao, M.; Wang, Z.W.; Huang, J.J. Multi-Scale Semantic Segmentation and Spatial Relationship Recognition of Remote Sensing Images Based on an Attention Model. *Remote Sens.* 2019, 11, 1044.
34. Alshehri, A.; Bazi, Y.; Ammour, N.; Almubarak, H.; Alajlan, N. Deep Attention Neural Network for Multi-Label Classification in Unmanned Aerial Vehicle Imagery. *IEEE Access* 2019, 7, 119873–119880.
35. Sun, H.; Zheng, X.; Lu, X.; Wu, S. Spectral-Spatial Attention Network for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* 2020, 58, 3232–3245.
36. Feng, J.; Wu, X.; Shang, R.; Sui, C.; Li, J.; Jiao, L.; Zhang, X. Attention Multibranch Convolutional Neural Network for Hyperspectral Image Classification Based on Adaptive Region Search. *IEEE Trans. Geosci. Remote Sens.* 2020.
37. Zhao, Z.; Wang, H.; Yu, X. Spectral-Spatial Graph Attention Network for Semisupervised Hyperspectral Image Classification. *IEEE Geosci. Remote Sens. Lett.* 2021.
38. Censi, A.M.; Ienco, D.; Gbodio, Y.J.E.; Pensa, R.G.; Interdonato, R.; Gaetano, R. Attentive Spatial Temporal Graph CNN for Land Cover Mapping from Multi Temporal Remote Sensing Data. *IEEE Access* 2021, 9, 23070–23082.
39. Hu, J.; Shen, L.; Sun, G. Squeeze-and-Excitation Networks. In *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
40. Xu, K.; Ba, J.; Kiros, R.; Cho, K.; Courville, A.; Salakhudinov, R.; Zemel, R.; Bengio, Y. Show, Attend and Tell: Neural Image Caption Generation with Visual Attention. In *Proceedings of the 32nd International Conference on Machine Learning*, *Proceedings of Machine Learning Research*, Lille, France, 7–9 July 2015; pp. 2048–2057.
41. Guo, Y.; Ji, J.; Lu, X.; Huo, H.; Fang, T.; Li, D. Global-Local Attention Network for Aerial Scene Classification. *IEEE Access* 2019, 7, 67200–67212.
42. Ma, J.; Ma, Q.; Tang, X.; Zhang, X.; Zhu, C.; Peng, Q.; Jiao, L. Remote Sensing Scene Classification Based on Global and Local Consistent Network. In *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS)*, Brussels, Belgium, 11–16 July 2020; pp. 537–540.

43. Hu, J.; Shen, L.; Albanie, S.; Sun, G.; Wu, E. Squeeze-and-Excitation Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 2020, 42, 2011–2023.
44. Wang, L.; Peng, J.T.; Sun, W.W. Spatial-Spectral Squeeze-and-Excitation Residual Network for Hyperspectral Image Classification. *Remote Sens.* 2019, 11, 884.
45. Alswayed, A.S.; Alhichri, H.S.; Bazi, Y. SqueezeNet with Attention for Remote Sensing Scene Classification. In *Proceedings of the ICCAIS 2020—3rd International Conference on Computer Applications and Information Security*, Riyadh, Saudi Arabia, 19–21 March 2020.
46. Li, C.Y.; Luo, B.; Hong, H.L.; Su, X.; Wang, Y.J.; Liu, J.; Wang, C.J.; Zhang, J.; Wei, L.H. Object Detection Based on Global-Local Saliency Constraint in Aerial Images. *Remote Sens.* 2020, 12, 1435.
47. Zhou, M.; Zou, Z.; Shi, Z.; Zeng, W.J.; Gui, J. Local Attention Networks for Occluded Airplane Detection in Remote Sensing Images. *IEEE Geosci. Remote Sens. Lett.* 2020, 17, 381–385.
48. Ding, L.; Tang, H.; Bruzzone, L. LANet: Local Attention Embedding to Improve the Semantic Segmentation of Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* 2021, 59, 426–435.
49. Chaudhari, S.; Mithal, V.; Polatkan, G.; Ramanath, R. An attentive survey of attention models. *arXiv* 2019, arXiv:1904.02874.
50. Lu, J.; Yang, J.; Batra, D.; Parikh, D. Hierarchical question-image co-attention for visual question answering. In *Proceedings of the NIPS*, Barcelona, Spain, 5–10 December 2016; pp. 289–297.
51. Jiang, H.W.; Hu, X.Y.; Li, K.; Zhang, J.M.; Gong, J.Q.; Zhang, M. PGA-SiamNet: Pyramid Feature-Based Attention-Guided Siamese Network for Remote Sensing Orthoimagery Building Change Detection. *Remote Sens.* 2020, 12, 484.
52. He, N.; Fang, L.; Li, Y.; Plaza, A. High-Order Self-Attention Network for Remote Sensing Scene Classification. In *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS)*, Yokohama, Japan, 28 July–2 August 2019; pp. 3013–3016.
53. Wu, Z.C.; Li, J.; Wang, Y.S.; Hu, Z.W.; Molinier, M. Self-Attentive Generative Adversarial Network for Cloud Detection in High Resolution Remote Sensing Images. *IEEE Geosci. Remote Sens. Lett.* 2020, 17, 1792–1796.
54. Cao, R.; Fang, L.; Lu, T.; He, N. Self-Attention-Based Deep Feature Fusion for Remote Sensing Scene Classification. *IEEE Geosci. Remote Sens. Lett.* 2021, 18, 43–47.
55. Wu, H.L.; Zhao, S.Z.; Li, L.; Lu, C.Q.; Chen, W. Self-Attention Network With Joint Loss for Remote Sensing Image Scene Classification. *IEEE Access* 2020, 8, 210347–210359.
56. Xiao, T.; Xu, Y.; Yang, K.; Zhang, J.; Peng, Y.; Zhang, Z. The application of two-level attention models in deep convolutional neural network for fine-grained image classification. In *Proceedings of the CVPR*, IEEE Computer Society, Boston, MA, USA, 7–12 June 2015; pp. 842–850.
57. Sumbul, G.; Cinbis, R.G.; Aksoy, S. Multisource Region Attention Network for Fine-Grained Object Recognition in Remote Sensing Imagery. *IEEE Trans. Geosci. Remote Sens.* 2019, 57, 4929–4937.
58. Li, J.; Tu, Z.; Yang, B.; Lyu, M.R.; Zhang, T. Multi-Head Attention with Disagreement Regularization. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, Brussels, Belgium, 31 October–4 November 2018; pp. 2897–2903.
59. Zhang, S.Y.; Li, C.R.; Qiu, S.; Gao, C.X.; Zhang, F.; Du, Z.H.; Liu, R.Y. EMMCNN: An ETPS-Based Multi-Scale and Multi-Feature Method Using CNN for High Spatial Resolution Image Land-Cover Classification. *Remote Sens.* 2020, 12, 66.
60. Cheng, B.; Li, Z.Z.; Xu, B.T.; Yao, X.; Ding, Z.Q.; Qin, T.Q. Structured Object-Level Relational Reasoning CNN-Based Target Detection Algorithm in a Remote Sensing Image. *Remote Sens.* 2021, 13, 281.
61. Wu, Z.; Hou, B.; Jiao, L. Multiscale CNN with Autoencoder Regularization Joint Contextual Attention Network for SAR Image Classification. *IEEE Trans. Geosci. Remote Sens.* 2021, 59, 1200–1213.
62. Shen, T.; Zhou, T.; Long, G.; Jiang, J.; Pan, S.; Zhang, C. DiSAN: Directional Self-Attention Network for RNN/CNN-free Language Understanding. In *Proceedings of the AAAI*, New Orleans, LA, USA, 2–7 February 2018.
63. Du, J.; Han, J.; Way, A.; Wan, D. Multi-Level Structured Self-Attentions for Distantly Supervised Relation Extraction. In *Proceedings of the EMNLP*, Brussels, Belgium, 31 October–4 November 2018.
64. Carrasco, M. Visual attention: The past 25 years. *Vis. Res.* 2011, 51, 1484–1525.
65. Beuth, F.; Hamker, F.H. A mechanistic cortical microcircuit of attention for amplification, normalization and suppression. *Vis. Res.* 2015, 116, 241–257.

66. Itti, L.; Koch, C.; Niebur, E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* 1998, 20, 1254–1259.
67. Ma, W.P.; Zhao, J.L.; Zhu, H.; Shen, J.C.; Jiao, L.C.; Wu, Y.; Hou, B.A. A Spatial-Channel Collaborative Attention Network for Enhancement of Multiresolution Classification. *Remote Sens.* 2021, 13, 106.
68. Tong, W.; Chen, W.; Han, W.; Li, X.; Wang, L. Channel-Attention-Based DenseNet Network for Remote Sensing Image Scene Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2020, 13, 4121–4132.
69. Guo, D.; Xia, Y.; Luo, X. Scene Classification of Remote Sensing Images Based on Saliency Dual Attention Residual Network. *IEEE Access* 2020, 8, 6344–6357.
70. Hang, R.L.; Li, Z.; Liu, Q.S.; Ghamisi, P.; Bhattacharyya, S.S. Hyperspectral Image Classification With Attention-Aided CNNs. *IEEE Trans. Geosci. Remote Sens.* 2021, 59, 2281–2293.
71. Zhu, M.; Jiao, L.; Liu, F.; Yang, S.; Wang, J. Residual Spectral-Spatial Attention Network for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* 2021, 59, 449–462.
72. Ren, Y.F.; Yu, Y.T.; Guan, H.Y. DA-CapsUNet: A Dual-Attention Capsule U-Net for Road Extraction from Remote Sensing Imagery. *Remote Sens.* 2020, 12, 2866.
73. Ren, Y.; Li, X.; Yang, X.; Xu, H. Development of a Dual-Attention U-Net Model for Sea Ice and Open Water Classification on SAR Images. *IEEE Geosci. Remote Sens. Lett.* 2021.
74. He, N.; Fang, L.; Plaza, A. Hybrid first and second order attention Unet for building segmentation in remote sensing images. *Sci. China Inf. Sci.* 2020, 63.
75. Pan, X.; Yang, F.; Gao, L.; Chen, Z.; Zhang, B.; Fan, H.; Ren, J. Building extraction from high-resolution aerial imagery using a generative adversarial network with spatial and channel attention mechanisms. *Remote Sens.* 2019, 11, 966.
76. Liu, R.C.; Cheng, Z.H.; Zhang, L.L.; Li, J.X. Remote Sensing Image Change Detection Based on Information Transmission and Attention Mechanism. *IEEE Access* 2019, 7, 156349–156359.
77. Wang, Q.; Liu, S.T.; Chanussot, J.; Li, X.L. Scene Classification With Recurrent Attention of VHR Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* 2019, 57, 1155–1167.
78. Li, Z.T.; Chen, G.K.; Zhang, T.X. Temporal Attention Networks for Multitemporal Multisensor Crop Classification. *IEEE Access* 2019, 7, 134677–134690.
79. Mei, X.G.; Pan, E.T.; Ma, Y.; Dai, X.B.; Huang, J.; Fan, F.; Du, Q.L.; Zheng, H.; Ma, J.Y. Spectral-Spatial Attention Networks for Hyperspectral Image Classification. *Remote Sens.* 2019, 11, 963.
80. Ma, F.; Gao, F.; Sun, J.P.; Zhou, H.Y.; Hussain, A. Attention Graph Convolution Network for Image Segmentation in Big SAR Imagery Data. *Remote Sens.* 2019, 11, 2586.
81. Luo, X.; Li, X.; Wu, Y.; Hou, W.; Wang, M.; Jin, Y.; Xu, W. Research on Change Detection Method of High-Resolution Remote Sensing Images Based on Subpixel Convolution. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2021, 14, 1447–1457.
82. Li, Y.; Chen, R.; Zhang, Y.; Zhang, M.; Chen, L. Multi-label remote sensing image scene classification by combining a convolutional neural network and a graph neural network. *Remote Sens.* 2020, 12, 4003.
83. Li, J.; Lin, D.Y.; Wang, Y.; Xu, G.L.; Zhang, Y.Y.; Ding, C.B.; Zhou, Y.H. Deep Discriminative Representation Learning with Attention Map for Scene Classification. *Remote Sens.* 2020, 12, 1366.
84. Bahri, A.; Majelan, S.G.; Mohammadi, S.; Noori, M.; Mohammadi, K. Remote Sensing Image Classification via Improved Cross-Entropy Loss and Transfer Learning Strategy Based on Deep Convolutional Neural Networks. *IEEE Geosci. Remote Sens. Lett.* 2020, 17, 1087–1091.
85. Zhang, C.; Yue, J.; Qin, Q. Global prototypical network for few-shot hyperspectral image classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2020, 13, 4748–4759.
86. Lei, P.C.; Liu, C. Inception residual attention network for remote sensing image super-resolution. *Int. J. Remote Sens.* 2020, 41, 9565–9587.
87. Cheng, W.S.; Yang, W.; Wang, M.; Wang, G.; Chen, J.Y. Context Aggregation Network for Semantic Labeling in Aerial Images. *Remote Sens.* 2019, 11, 1158.
88. Gbodio, Y.J.E.; Ienco, D.; Leroux, L.; Interdonato, R.; Gaetano, R.; Ndao, B. Object-based multi-temporal and multi-source land cover mapping leveraging hierarchical class relationships. *Remote Sens.* 2020, 12, 2814.
89. Liang, L.; Wang, G. Efficient recurrent attention network for remote sensing scene classification. *IET Image Process.* 2021.

90. Wang, Z.S.; Zou, C.; Cai, W.W. Small Sample Classification of Hyperspectral Remote Sensing Images Based on Sequential Joint Deeping Learning Model. *IEEE Access* 2020, 8, 71353–71363.
91. You, H.; Tian, S.; Yu, L.; Lv, Y. Pixel-Level Remote Sensing Image Recognition Based on Bidirectional Word Vectors. *IEEE Trans. Geosci. Remote Sens.* 2020, 58, 1281–1293.
92. Zhang, X.K.; Pun, M.O.; Liu, M. Semi-Supervised Multi-Temporal Deep Representation Fusion Network for Landslide Mapping from Aerial Orthophotos. *Remote Sens.* 2021, 13, 548.
93. Wong, R.; Zhang, Z.J.; Wang, Y.M.; Chen, F.S.; Zeng, D. HSI-IPNet: Hyperspectral Imagery Inpainting by Deep Learning With Adaptive Spectral Extraction. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2020, 13, 4369–4380.

Retrieved from <https://encyclopedia.pub/entry/history/show/33365>