REVIEW ARTICLE

DEEP LEARNING IN MULTISTAGE AND MULTISCALE REMOTE SENSING

Thomas U. Omali

National Biotechnology Development Agency (NABDA), Abuja, Nigeria.

*Corresponding author E-mail: t.omali@yahoo.com

DOI: https://doi.org/10.5281/zenodo.17262357

Abstract:

The use of remote sensing methods for accurate and continuous monitoring of the Earth is increasing. Thus, various approaches are adopted for analyzing its data (imagery) for precise results. For example, one may analyze remotely sensed data collected at different scales or resolutions giving rise to the concept of multiscale remote sensing. Remotely sensed data can also be analyzed in multiple stages or iterations, which form the foundation of multistage remote sensing. This study aims to present an extensive review that summarizes multistage and multiscale remote sensing methods with special emphasis on Deep Learning applications. Of course, Deep Learning algorithms are influencing remote sensing image analysis. The evidence is in many existing articles some of which were reviewed in this study. Summarily, Deep Learning for multiscale and multistage techniques in remote sensing enhances the accuracy and efficiency of image analysis.

Keywords: Deep Learning, Fusion, Image Classification, Remote Sensing, Small Target.

Introduction:

Remote sensing provides an effective and rapid method for monitoring natural resources. It is characterized by the ability to acquire data in remote and inaccessible areas. The development of Earth Observation (EO) technology has enhanced remote sensing applications (Zhang & Shen, 2022). Also, recent advancements in remote sensing sensors and information transmission technologies have considerably improved the spatial resolution of remote sensing systems (Marin *et al.,* 2015). However, many situations or analyses require the integration of multiple data from different or single remote sensing system(s) or the analysis of the imaeries at many stages. This has given rise to multi-image, which is a technique for generating images or information by combining multiple images. Of course, the images to be merged may be acquired from different sensors in the same area, at various points in time, or from different angles and viewpoints.

With integrative processing of more images or information, it is possible to obtain information that is not available in a single data. Coarse-resolution sensors can be used to acquire regular temporal measurements. However, they lack the required spatial detail for resolving individual fields in some regions of the world (Graesser & Ramankutty, 2017). Regardless of this limitation, they are highly effective in large-scale observations (Maus *et al.,* 2016). Furthermore, higher-resolution sensors do not regularly acquire data, and they frequently lack spectral bands in key wavelengths such as the shortwave infrared or the red edge. Thus, modeling and aligning multiple image frames will allow information sharing between sequential frames. The implication is that the image resolution will be enhanced. Apparently, multi-image remote sensing techniques have been successfully used in disaster monitoring (Liu *et al.,* 2022; Wang *et al.,* 2022), land cover assessment (Lepcha *et al.,* 2023; Malczewska & Wielgosz, 2024), and vegetation growth monitoring (Wang *et al.,* 2023).

The conventional method for exploiting remote sensing data depends on feature extraction and/or feature selection (Paheding *et al.,* 2024). Unfortunately, extraction and selection of features are suboptimal in the context of broad representation of original data for certain applications. To be specific, their appropriateness can be reduced in the framework of big data. This is because remote sensing data vary in time, geo-location, atmospheric conditions, and imaging platform (Zhang *et al.,* 2016; Zhu *et al.,* 2017). Besides, the complex nature of remote sensing data makes its analysis a challenging task. Thus, a need arises for an effective method that could be used for the automatic extraction of relevant features from various remote sensing data (Paheding *et al.,* 2024). Additionally, a dependable system for analysis is required (Adegun *et al.,* 2023). In this regard, Deep Learning (DL) has proven its effectiveness (Yann er at., 2015). Of course, DL as emerged as a robust tool for remote sensing application. It revolutionized ow remote sensing data are analyzed and interpreted. Also, it can automatically learn complex features in remote sensing data. Ferchichi *et al.* (2022) revealed that many studies on the application of DL in multidisciplinary approaches including its utilization in remote sensing are in tew literature. It is noteworthy that the application of Deep Learning in remote sensing is mainly used in three aspects, namely surface classification, object detection and change detection.

Fundamentals of Deep Learning

Deep Learning is a component of Machine Learning, which was developed for the purpose of solving a complex problem with a huge amount of data. Its concentration is mainly on the development of systems stimulated by the structural arrangement and functioning of the brain known as Artificial Neural Networks (ANN). Generally, the DL structure is characterised by multiple layers. With these numerous layers, it can automatically learn the abstract features from the raw data directly from which valuable representations can be determined (Rajagukguk *et al.,* 2020). Data abstraction and extraction from the lower to higher layers are done through simple nonlinear modules.

By and large, Deep Learning has demonstrated potential for advances in various fields (Maduako *et al.,* 2022). Now, enhanced algorithms and multilayer networks such as Convolution Neural Networks (CNNs), Recurrent Neural Networks (RNNs), DBN (Deep Belief Network) and many more have shown to be more effective than standard methods.

Multiscale Remote Sensing

Remote sensing object detection is extremely important (Chen *et al.,* 2023) for natural disaster monitoring (Alkhatib *et al.,* 2023), large-scale scene detection, and resource surveying. Studies have revealed how it is of great significance to human life and societal development (e.g. Masita *et al.,* 2020). Also, remote sensing images are characterized by multiple scales (Cai *et al.,* 2023), and complex backgrounds (Cai *et al.,* 2022). Thus, it is often difficult to logically extract specific features from the high-resolution imageries that contain rich feature information. Furthermore, the image sensor information is weak such that it gives rise to many difficulties and challenges in small target detection. Many studies have been conducted on network structure, training, data processing, and others to improve the detection performance of small targets. Yet, there is a gap in the performance of small-target detection compared with that of large- and medium-target detection.

Considering the case of optical remote sensing data that involves long distance, images are normally presented at a smaller scale (small targets). Therefore, they contain fewer pixels and feature information. Optical remote sensing data are characterized by several kinds of targets with large-scale variations. Thus, remote sensing analysts should ensure that small targets and multiscale are considered in the designed framework for a better detection accuracy. Many recent studies have suggested enhancing the accuracy of small and multi-scale targets in optical remote sensing data (e. g. Deng *et al.,* 2018; Li *et al.,* 2019; Li *et al.,* 2020; Pham *et al.,* 2020; Amazirh *et al.,* 2024). These studies revealed the significance of multi-stage classifiers, multi-stage features quadratic fusion, and loss function weighting.

Multistage Remote Sensing

Multistage or multilevel remote sensing method is concerned with the use of a remote sensing system to capture data at various levels. It involves a hierarchical classification, which detects some important classes in the first level and progressively differentiates different sub-classes within each main class (Dibs *et al.,* 2017; Mao *et al.,* 2020). The method of multistage mapping is examined to a lesser extent compared to the use of different elements in a single classification process. However, separating diverse land cover types into a sequence of sub-classes produces better results than direct classifications of all classes. For example, Kwong *et al.* (2022) facilitated the effective mapping of diverse habitats by integrating VHR satellite images, GIS databases, and post-classification rules. In the first stage, initial classification was performed on a 2m WorldView-2/3 image to map several classes with a variety of variables and classification methods. In the second and third stages, modification procedures were adopted with GIS data and spatial relationships to identify habitats to produce the final results.

Landsat imagery has proven to be effective in Forest Inventory and Analysis (FIA). Unfortunately, Landsat doesn't greatly reduce the required amount of field data. However, it is possible to get more detailed information that observes rapid changes in forest cover and conditions using high-resolution data. Thus, a multistage statistical design can combine wall-to-wall Landsat data in the first stage, a sample of high-resolution imagery in the second stage, and conventional forest inventory and analysis in the third stage. It is also possible to improve the mapping results by using post-classification modification rules with thematic layers within a GIS (Thakkar *et al.,* 2017). Examples of such layers are terrain, land use, geologic and climatic data.

Progressive refinement is an important characteristic of multistage analysis in remote sensing (Liu *et al.,* 2025). Through multiple stages, coarse-grained features can be analyzed as the first step followed by the progressive refinement of those features. Another significance of multistage remote sensing is its ability to solve the small target detection problem. The features of small targets can be enriched using multistage feature maps. Evidence in many studies demonstrates the effectiveness of the multistage method in remote sensing. For instance, Zhang & Shen (2022) improved the accuracy of target detection in optical remote sensing using the Multi-stage Feature Enhancement Pyramid

Network. Esmaiel *et al.* (2022) proposed multistage feature mining to enhance Ship-Radiated Noise (SRN) feature extraction in the detected passive sonar signal. This was based on Enhanced Variational Mode Decomposition (EVMD), Weighted Permutation Entropy (WPE), Local Tangent Space Alignment (LTSA), and Particle Swarm Optimization-Based Support Vector Machine (PSO-SVM). In another study, Zhou *et al.* (2024) combined three levels (including image, feature, and decision) in a multistage interaction network designed for effective change detection. Similarly, Xiaogang *et al.* (2024) presented a Multi-Stage Progressive Change Detection Network (MSP-CD), particularly for urban growth change detection. The approach integrated invariant detection, knowledge distillation, and a coarse-to-fine change detection structure.

Deep Learning Application in Multiscale and Multistage Remote Sensing

The extraction of multi-scale features has been based on conventional techniques such as dense sampling of windows (Yan *et al.,* 2014), spatial pyramids, and their combination. With the improved DL models (Lei *et al.,* 2021), different change detection techniques have emerged (Pan *et al.,* 2022; Yan *et al.,* 2023). Of course, the current level of deep neural networks results in high-level network design solutions for object detection (Ren *et al.,* 2017; Law & Deng, 2020; Zhang & Shen, 2021). It is noteworthy tath *Remote sensing* data are *characterized by delicate and intricate features, which pose difficulties in using traditional techniques for accurate data extraction (*Almohsen, 2024; Liu *et al.,* 2025). However, DL algorithms differs from conventional machine learning algorithms and data mining because it can take very comprehensive representations of data from large datasets (Taye, 2023). It can also handle complex relationships and automate feature learning (Bengani, 2024). This has resulted in the recent shift of attention to Deep Learning among remote sensing analysts. The application of DL has recorded significant success in several remote sensing image analyses including land use and land cover classification, scene classification, and object detection (Chen *et al.,* 2015; Romero *et al.,* 2016; Marmanis *et al.,* 2016; Kussul *et al.,* 2017; Vetrivel *et al.,* 2018). Some of the DL algorithms in multistage analysis are presented in Table 1.

Table 1: Deep Learning in Multistage Analysis

Method	**Application**	**Authors**
Convolutional Neural Networks (CNNs)	Feature extraction from remote sensing imagery	Alshehhi *et al.,* 2017; Bai *et al.,* 2018; Li *et al.,* 2018; Yi *et al.,* 2019; Aryal and Neupane, 2023
Recurrent Neural Networks (RNNs)	Capturing temporal dependencies in multi-temporal data	Ienco *et al.,* 2017; Sharma, Liu, and Xiaojun, 2017; Ndikumana *et al.,* 2018; Mou *et al.,* 2019; Grünig, *et al.,* 2021
Transformers	Capturing long-range dependencies and contextual information in remote sensing imagery	He *et al.,* 2020; Chen, Shi, Qi, 2021; Bazi *et al.,* 2021; Yang *et al.,* 2022; Ma *et al.,* 2022; Boulila *et al.,* 2024; Bolcek *et al.,* 2025
Generative Adversarial Networks (GANs)	Image super-resolution and data augmentation	Crivellari *et al.,* 2023; Huang, and Cao, 2023

Over the last few years, a strong interest has been raised in CNNs as a powerful machine-learning tool. CNNs have been employed on multiple tasks such as object detection (Lin *et al.,* 2017) and classification (He *et al.,* 2016). This is because they can autonomously extract highly representative features, overshadowing traditional learning approaches which are based on hand-crafted features. Wang *et al.* (2022) conducted a comprehensive investigation with a special emphasis on deep learning in multiscale agricultural remote and proximal sensing. They specifically focussed on the applications of CNN supervised learning, transfer learning, and few-shot learning in crop sensing at land, field, canopy, and leaf scales. It should be noted that Feature pyramids are usually constructed with CNN. This is due to the pooling and multi-convolution kernel operations that make them suitable for building deep multiscale representations (Jiao *et al.,* 2021). Furthermore, there are two common methods for CNN-based multi-scale learning. The first technique uses external or preset architectures such as the multi-scale kernel (Szegedy *et al.,* 2015; Zou & Cheng, 2021) and multi-scale inputs (Dollár & Zitnick, 2014). In the second method, internal layers of the network including skip and dense connections are designed (Chen *et al.,* 2017).

Another model commonly utilized for supervised learning is the RNN, which was conventionally employed for discrete sequence analysis. However, advances in their architecture and training methods make them applicable to more complex tasks involving remote sensing data analysis (e.g. Lyu *et al.,* 2016; Ho Tong Minh *et al.,* 2018).

Furthermore, the Generative Adversarial Networks (GANS) as presented in Goodfellow *et al.* (2014) are pervasive unsupervised Deep Learning algorithms. GANs have been effective in computer vision and image processing (Zhan *et al.,* 2018) including remote sensing image processing applications. It comprises generative and discriminative networks. The generative network learns from a hidden space to a specific data distribution of interest. On the other hand, a discriminative network differentiates between the real and generated data. Both the generative and discriminative networks enhance different and disparate loss functions in a zero-zum game (Oliehoek *et al.,* 2017).

In recent times, there has been more interest in transformer-based architectures. Notably, the Vision Transformer (ViT) as shown in Dosovitskiy *et al.* (2020) has been particularly successful. When compared to CNN, the ViT balances the global and local features much better. Also, other efforts have been made for multi-scale feature learning including the multi-scale Deformable Attention and Multilevel Features Aggregation (MSDAM and MLFAM) network (Dong *et al.,* 2022), and the Shunted Self-Attention (SSA) network (Ren *et al.,* 2022).

Conclusion:

The applications of remote sensing in monitoring the Earth-related occurrences are increasing. Thus, there is need to improve the interpretation of remote sensing data. In that regard, many approaches have emerged including multistage and multiscale remote sensing methods. The multistage feature maps can enhance the features of small-scale targets. Models can also rely on more feature difference information for target classification. Furthermore, the methods to solve the multiscale target detection problem use multistage classifiers, multistage feature quadratic fusion, and loss function weighting. Although these methods can deal with the specificities of targets in sensing images, they can reduce detection efficiency. Of course, the detection accuracy will not be sufficiently high as feature usage is insufficient.

The existing and potential applications of Deep Learning was highlighted in this study. The main intention was to encourage increased interest and engagement among vision and remote sensing experts. This will inspire the use of Deep Learning models not only for multiscale and multistage remote sensing image analysis but also for broader applications in the remote sensing field. DL possesses significant advantages for processing time-series data such as the RNN, which was traditionally applied to a discrete sequence analysis. Also, CNNs have been successfully used in remote sensing image fusion.

References:

1. Alkhatib, R., Sahwan, W., Alkhatieb, A., & Schütt, B. (2023). A brief review of machine learning algorithms in forest fires science. *Applied Sciences, 13*, 8275.
2. Almohsen, A. S. (2024). Challenges facing the use of remote sensing technologies in the construction industry: a review. *Buildings*, *14*(9), 2861. https://doi.org/10.3390/buildings14092861
3. Alshehhi, R., Marpu, P. R., Woon, W. L., & Mura, M. D. (2017). Simultaneous extraction of roads and buildings in remote sensing imagery with convolutional neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, *130*, 139–149.
4. Aryal, J., & Neupane, B. (2023). Multi-scale feature map aggregation and supervised domain adaptation of fully convolutional networks for urban building footprint extraction. *Remote Sensing, 15*(2), 488.
5. Amazirh, A., Ouassanouan, Y., Bouimouass, H., Baba, M. W., Bouras, E. H., Rafik, A., Benkirane, M., Hajhouji, Y., Ablila, Y., & Chehbouni, A. (2024). Remote Sensing-Based Multiscale Analysis of Total and Groundwater Storage Dynamics over Semi-Arid North African Basins. *Remote Sensing*, *16*(19), 3698. https://doi.org/10.3390/rs16193698
6. Bai, Y., Jiang, D. M., Pei, J. J., Zhang, N., & Bai, Y. (2018). Application of an improved ELU convolution neural network in the SAR image ship detection. *Bulletin of Surveying and Mapping, 1*, 125–128.
7. Bazi, Y., Bashmal, L., Rahhal, M. M. A., Dayil, R. A., & Ajlan, N. A. (2021). Vision transformers for remote sensing image classification. *Remote Sensing, 13*, 516. https://doi.org/10.3390/rs13030516
8. Bengani, V. (2024). Hybrid learning systems: Integrating traditional machine learning with deep learning techniques. 10.13140/RG.2.2.10461.22244/1.
9. Bolcek, J., Gibril, M., Al-Ruzouq, R., Shanableh, A., Jena, R., Hammouri, N., Sachit, M. S., Ghorbanzadeh, O. (2025). A Comprehensive Evaluation of Deep Vision Transformers for Road Extraction from Very-high-resolution Satellite Data. *Science of Remote Sensing, 11*, 100190. https://doi.org/10.1016/j.srs.2024.100190.
10. Boulila, W., Ghandorh, H., Masood, S., Alzahem, A., Koubaa, A., Ahmed, F., Khan, Z., Ahmad, J. A. (2024). Transformer-based approach empowered by a self-attention technique for semantic

segmentation in remote sensing. *Heliyon*, *10*(8), e29396. https://doi.org/10.1016/j.heliyon.2024.e29396.

11. Cai, D., Lu, Z., Fan, X., Ding, W., & Li, B. (2023). Improved YOLOv4-tiny target detection method based on adaptive self-order piecewise enhancement and multiscale feature optimization. *Applied Sciences, 13*, 8177.
12. Cai, Y., Zhou, Y., Zhang, H., Xia, Y., Qiao, P., & Zhao, J. (2022). Review of target geo-location algorithms for aerial remote sensing cameras without control points. *Applied Sciences, 12*, 12689.
13. Chen, H., Shi, Zhenwei, & Qi, Z. (2021). Remote sensing image change detection with transformers. *IEEE Transactions on Geoscience and Remote Sensing, 60*. 10.1109/TGRS.2021.3095166.
14. Chen, Y., Ma, L., Liu, T., Huang, X., & Sun, G. (2023). The synergistic effect between precipitation and temperature for the NDVI in Northern China from 2000 to 2018. *Applied Sciences, 13*, 8425.
15. Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2017). Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 40*(4), 834–848.
16. Chen, Y., Zhao, X., Jia, X. (2015). Spectral-spatial classification of hyperspectral data based on deep belief network. *Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8*(6SI), 2381–2392.
17. Cheng G, Han J, Guo L, Qian X, Zhou P, Yao X, Xintao H. (2013). Object detection in remote sensing imagery using a discriminatively trained mixture model. *ISPRS Journal of Photogrammetry and Remote Sensing*, *85*, 32–43.
18. Crivellari, A., Wei, H., Wei, C., & Shi, Y. (2023). Super-resolution GANs for upscaling unplanned urban settlements from remote sensing satellite imagery– the case of Chinese Urban village detection. *International Journal of Digital Earth. 16*, 2623–2643. 10.1080/17538947.2023.2230956.
19. Deng, Z., Sun, H., Zhou, S., Zhao, J., Lei, L., & Zou, H. (2018). Multi-scale object detection in remote sensing imagery with convolutional neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, *145*, 3–22.
20. Dibs, H., Idrees, M. O., & Alsalhin, G. B. A. (2017). Hierarchical classification approach for mapping rubber tree growth using per-pixel and object-oriented classifiers with SPOT-5 imagery. *Egyptian Journal of Remote Sensing and Space Science, 20*, 21–30.
21. Dollár, P., & Zitnick, C. L. (2014). Fast edge detection using structured forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 37*(8), 1558–1570.
22. Dong, S., Hangel, G., Bogner, W., Widhalm, G., Rössler, K., Trattnig, S., You, C., de Graaf, R., Onofrey, J., & Duncan, J. S., (2022). Multi-scale Super-resolution Magnetic Resonance Spectroscopic Imaging with Adjustable Sharpness. *In International Conference on Medical Image Computing and Computer-Assisted Intervention*, 410–420. Springer.

23. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, Uszkoreit, S. J., Houlsby, N., & Gelly, S. (2020). An Image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929.
24. Esmaiel, H., Xie, D., Qasem, Z. A. H., Sun, H., Qi, J., & Wang, J. (2022). Multi-stage feature extraction and classification for ship-radiated noise. *Sensors*, *22*(1), 112. https://doi.org/10.3390/s22010112.
25. Ferchichi, A., Abbes, A. B., Barra, V., & Farah, I. R. (2022). Forecasting vegetation indices from spatio temporal remotely sensed data using deep learning-based approaches: A systematic literature review. Ecological Informatics, 68, 101552.
26. Graesser, J., & Ramankutty, N. (2017). Detection of cropland field parcels from Landsat imagery. *Remote Sensing of Environment, 201*, 165–180.
27. Grünig, M., Razavi, E., Calanca, P., Mazzi, D., Wegner, J. D., Pellissier, L. (2021). Applying deep neural networks to predict incidence and phenology of plant pests and diseases. *Ecosphere*, *12*(10), e03791.
28. Hao, S., Wu, B., Zhao, K., Ye, Y., & Wang, W. (2022). Two-stream swin transformer with differentiable sobel operator for remote sensing image classification. *Remote Sensing*, *14*(6), 1507.
29. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.
30. He, J., Zhao, L., Yang, H., Zhang, M., & Li, W. (2020). HSI-BERT: Hyperspectral image classification using the bidirectional encoder representation from transformers. *IEEE Transactions on Geoscience and Remote Sensing, 58*, 165–178.
31. Ho Tong Minh, D., Ienco, D., Gaetano, R., Lalande, N., Ndikumana, E., Osman, F., & Maurel, P. (2018). Deep recurrent neural networks for winter vegetation quality mapping via multitemporal SAR Sentinel-1. *IEEE Geoscience and Remote Sensing Letter, 15*(3), 464–468.
32. Huang, Y., & Cao, L. (2023). Privacy-preserving remote sensing image generation and classification with differentially private GAN. *IEEE Sensors Journal,* 1-1. doi: 10.1109/JSEN.2023.3267001.
33. Ienco, D., Gaetano, R., Dupaquier, C., & Maurel, P. (2017). Land cover classification via multitemporal spatial data by deep recurrent neural networks. *IEEE Geoscience and Remote Sensing Letter, 14*(10), 1685-1689.
34. Maduako, I., Igwe, C. F., Abah, J. E., Onwuasaanya, O. E., Chukwu, A., Ezeji, F., & Okeke, F. I. (2022). Deep learning for component fault detection in electricity transmission lines. *Journal of Big Data,* 9(81), 34 pages. https://doi.org/10.1186/s40537-022-00630-2
35. Kussul, N., Lavreniuk, M., Skakun, S., Shelestov, A., 2017. Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letter*, *14*(5), 778–782.

36. Kwong, I. H. Y., Wong, F. K. K., Fung, T., Liu, E. K. Y., Lee, R. H., & Ng, T. P. T. (2022). A multi-stage approach combining very high-resolution satellite images, GIS database, and post-classification modification rules for habitat mapping in Hong Kong. *Remote Sensing*, *14*(1), 67. https://doi.org/10.3390/rs14010067
37. Jiao, L., Gao, J., Liu, X., Liu, F., Yang, S., Hou, B. (2021). Multi-scale representation learning for image classification: a survey. *IEEE Transactions on Artificial Intell*igence, *99*, 1-1. 10.1109/TAI.2021.3135248.
38. Law, H., & Deng, J. (2020). CornerNet: detecting objects as paired keypoints. *International Journal of Computer Vision, 128*, 642–656.
39. Lei, Y., Peng, D., Zhang, P., Ke, Q., Li, H. (2021). Hierarchical paired channel fusion network for street scene change detection. *IEEE Transactions on Image Processing, 30*, 55–67. http://dx.doi.org/10.1109/TIP.2020.3031173.
40. Lepcha, D. C., Goyal, B., Dogra, A. & Goyal, V. (2023). Image super-resolution: a comprehensive review, recent trends, challenges and applications. *Information Fusion*, *91*, 230–260.
41. Li, Y., Huang, Q., Pei, X., Chen, Y., Jiao, L., & Shang, R. (2020). Cross-layer attention network for small object detection in remote sensing imagery. *Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14*, 2048–2146.
42. Li, J. W., Qu, C. W., & Peng, S. J. (2018). Ship classification for unbalanced sar dataset based on convolutional neural network. *Journal of Applied Remote Sensing, 12*(3), 035010.
43. Li, C., Xu, C., Cui, Z., Wang, D., Zhang, T., & Yang, J. (2019, September 22–25). *Feature-attentioned object detection in remote sensing imagery*. In: *proceedings of the 2019 IEEE International Conference on Image Processing (pp. 3886–3890).* Manhattan, NY, USA.
44. Lin, T. -Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In: *2017 IEEE International Conference on Computer Vision (pp. 2999-3007).* Venice, Italy, doi: 10.1109/ICCV.2017.324.
45. Liu, Feiqiang & Yu, Qiang & Chen, Lihui & Jeon, Gwanggil & Keese Albertini, Marcelo & Yang, Xiaomin. (2022). Aerial image super-resolution based on deep recursive dense network for disaster area surveillance. *Personal and Ubiquitous Computing*, *26*. 1-10. 10.1007/s00779-020-01516-x.
46. Liu, B., Zhan, C., Guo, C., & Ruan, S. (2025). Efficient remote sensing image classification using the novel STConvNeXt convolutional network. *Scientific Reports, 15*, 8406. https://doi.org/10.1038/s41598-025-92629-x
47. Liu, D., Zhong, L., Wu, H., Songyang Li, S., & Li, Y. (2025). Remote sensing image super-resolution reconstruction by fusing multi-scale receptive fields and hybrid transformer. *Scientific Reports, 15*, 2140. https://doi.org/10.1038/s41598-025-86446-5
48. Lyu, H., Lu, H., Mou, L. (2016). Learning a transferable change rule from a recurrent neural network for land cover change detection. *Remote Sensing*, *8*(6), 506.

49. Ma, J., Li, M., Tang, X., Zhang, X., Liu, F., & Jiao, L. (2022). Homo–heterogenous transformer learning framework for remote sensing scene classification. *Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15*, 2223–2239.
50. Malczewska, A. & Wielgosz, M. (2025). How does super-resolution for satellite imagery affect different types of land cover? Sentinel-2 Case. *Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17*, 340–363.
51. Marmanis, D., Datcu, M., Esch, T., Stilla, U. (2016). Deep learning earth observation classification using imagenet pre-trained networks. *IEEE Geoscience and Remote Sensing Letter, 13*(1), 105–109.
52. Marin, C., Bovolo, F., Bruzzone, L., (2015). *Building change detection in multitemporal very high-resolution SAR images*. IEEE Transactions on Geoscience and Remote Sensing*, 53*(5), 2664–2682. http://dx.doi.org/10.1109/TGRS.2014.2363548.
53. Mao, D., Wang, Z., Du, B., Li, L., Tian, Y., Jia, M., Zeng, Y., Song, K., Jiang, M., & Wang, Y. (2020). National wetland mapping in china: a new product resulting from object-based and hierarchical classification of Landsat 8 OLI Images. *ISPRS Journal of Photogrammetry and Remote Sensing, 164*, 11–25.
54. Masita, K. L., Hasan, A. N., & Shongwe, T. (2020, August 6–7). *Deep Learning in Object Detection: A Review.* In: Proceedings of the 2020 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (pp. 1–11). Durban, South Africa.
55. Maus, V., Camara, G., Cartaxo, R., Sanchez, A., Ramos, F. M., de Queiroz, & G. R. (2016). A time-weighted dynamic time-warping method for land-use and land-cover mapping. *Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9*, 3729–3739.
56. Mou, L., Bruzzone, L., & Zhu, X. X. (2019). Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(2), 924–935.
57. Ndikumana, E., Ho Tong Minh, D., Baghdadi, N., Courault, D., & Hossard, L. (2018). Deep recurrent neural network for agricultural classification using multitemporal SAR Sentinel-1 for Camargue, France. *Remote Sensing, 10*(8), 1217. https://doi.org/10.3390/rs10081217
58. Pan, J., Cui, W., An, X., Huang, X., Zhang, H., Zhang, S., Zhang, R., Li, X., Cheng, W., Hu, Y., 2022. MapsNet: multi-level feature constraint and fusion network for change detection. *International Journal of Applied Earth Observation and Geoinformation, 108*, 102676. http://dx.doi.org/10.1016/j.jag.2022.102676.
59. Pham, M. T., Courtrai, L., Friguet, C., Lefèvre, S., & Baussard, A. (2020). YOLO-fine: one-stage detector of small objects under various backgrounds in remote sensing images. *Remote Sensing*, *12*, 2501.
60. Rajagukguk, R. A., Ramadhan, R. A. A., & Lee, H.-J. (2020). A Review on deep learning models for forecasting time series data of solar irradiance and photovoltaic power. *Energies*, 13, 6623; doi:10.3390/en13246623

61. Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*, 1137–1149.
62. Ren, S., Zhou, D., He, S., Feng, J., & Wang, X. (2022). *Shunted Self-attention via Multi-scale Token Aggregation.* In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10853–10862).
63. Romero, A., Gatta, C., Camps-Valls, G. (2016). Unsupervised deep feature extraction for remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing, 54*(3), 1349–1362.
64. Sharma, A., Liu, X., & Yang, X. (2017). Land cover classification from multi-temporal, multi-spectral remotely sensed imagery using patch-based recurrent neural networks. *Neural Networks.* 105. 10.1016/j.neunet.2018.05.019.
65. Paheding, S., Saleem, A., Faridul, M., Siddiqui, H., Rawashdeh, N., Essa, A., & Reyes, A. A. (2024). Advancing horizons in remote sensing: a comprehensive survey of deep learning models and applications in image classification and beyond. *Neural Computing and Applications*, *36*, 16727–16767.
66. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A. (2015). *Going deeper with convolutions*. In: Proceedings of the IEEE conference on computer vision and pattern recognition. Boston, 1-9. https://doi.org/10.1109/CVPR.2015.7298594
67. Taye, M. M. (2023). Understanding of machine learning with deep learning: architectures, workflow, applications and future directions. *Computers, 12*(5), 91. https://doi.org/10.3390/computers12050091
68. Thakkar, A. K., Desai, V. R., Patel, A., & Potdar, M. B. (2017). Post-classification corrections in improving the classification of land use/land cover of arid region using RS and GIS: The Case of Arjuni Watershed, Gujarat, India. *Egyptian Journal of Remote Sensing and Space Science, 20*, 79–89.
69. Vetrivel, A., Gerke, M., Kerle, N., Nex, F., Vosselman, G. (2018). Disaster damage detection through synergistic use of deep learning and 3d point cloud features derived from very high-resolution oblique aerial images, and multiple-kernel-learning. *ISPRS Journal of Photogrammetry Remote Sensing, 140*, 45–59.
70. Wang, D., Cao, W., Zhang, F., Li, Z., Xu, S., & Wu, X. (2022). A review of deep learning in multiscale agricultural sensing. *Remote Sensing, 14*, 559. https://doi.org/10.3390/rs14030559
71. Wang, H., Cao, H., Kai, Y., Bai, H., Chen, X., Yang, Y., Xing, L., & Zhou, C. (2022). Multi-source remote sensing intelligent characterization technique-based disaster regions detection in high-altitude mountain forest areas. *IEEE Geoscience and Remote Sensing Letters.* 19. 1-1. 10.1109/LGRS.2022.3185420.
72. Wang, X., Hu, Q., Cheng, Y. & Ma, J. (2023). Hyperspectral image super-resolution meets deep learning: a survey and perspective. *IEEE/CAA Journal of Automatica Sinica,* 10, 1668–1691.

73. Xiaogang, N, Hanchao, Z., Ruiqian, Z., & Xiao, H. (2024). Multi-stage progressive change detection on high-resolution remote sensing imagery. *ISPRS Journal of Photogrammetry and Remote Sensing, 207*, 231–244.
74. Yan, S., Xu, X., Xu, D., Lin, S., & Li, X. (2014). Image Classification with Densely Sampled Image Windows and Generalized Adaptive Multiple Kernel Learning. *IEEE Transactions on Cybernetics*, *45*(3), 381–390.
75. Yan, L., Yang, J., Wang, J. (2023). Domain knowledge-guided self-supervised change detection for remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 16*, 4167–4179. http://dx.doi.org/10.1109/JSTARS.2023.3270498.
76. Yann, L. C., Yoshua, B., & Geoffrey, H. (2015). Deep learning. *Nature*, *521*(7553), 436–44.
77. Yang, L., Yang, Y., Yang, J., Zhao, N., Wu, L., Wang, L., & Wang, T. (2022). FusionNet: a convolution-transformer fusion network for hyperspectral image classification. *Remote Sensing*, *14*, 4066.
78. Yi, Y., Zhang, Z., Zhang, W., Zhang, C., Li, W., & Zhao, T. (2019). Semantic segmentation of urban buildings from VHR remote sensing imagery using a deep convolutional neural network. *Remote Sensing, 11*, 1774.
79. Zhan, Y., Hu, D., Wang, Y., & Yu, X. (2018). Semi-supervised hyperspectral image classification based on generative adversarial networks. *IEEE Geoscience and Remote Sensing Letter*, *15*(2), 212–216.
80. Zhang, K. H., & Shen, H. K. (2021). Solder joint defect detection in the connectors using improved faster-rcnn algorithm. *Applied Sciences, 11*, 576.
81. Zhang, K., & Shen, H. (2022). Multi-stage feature enhancement pyramid network for detecting objects in optical remote sensing images. *Remote Sensing, 14*, 579. https://doi.org/10.3390/rs14030579
82. Zhang, L., Zhang, L., & Du, B. (2016). Deep learning for remote sensing data: a technical tutorial on the state of the art. *IEEE Transactions on Geoscience and Remote Sensing Magazine, 4*(2), 22–40.
83. Zhou, M., Qian, W., & Ren, K. (2024). Multistage interaction network for remote sensing change detection. *Remote Sensing, 16*, 1077. https://doi.org/10.3390/rs16061077
84. Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: a comprehensive review and list of resources. *IEEE Transactions on Geoscience and Remote Sensing Magazine, 5*(4), 8–36.
85. Zou, Y., & Cheng, L. (2021). A transfer learning model for gesture recognition based on the deep features extracted by CNN. *IEEE Transactions on artificial intelligence, 2*(5), 447–458.