

Deep Learning Techniques for Automated Classification of Satellite Imagery and Land-Use Patterns

Eva Novaković

Research Fellow, Central Adriatic Institute for Earth Observation Analytics, Slovenia

Abstract

Satellite imagery provides large-scale spatial information that can support land-use monitoring, environmental assessment, urban analysis, agricultural observation, and infrastructure planning. Traditional image-classification approaches generally depend on handcrafted spectral, textural, and spatial descriptors. Deep learning offers an alternative by learning hierarchical image representations directly from training data. Convolutional neural networks are particularly relevant because land-use patterns are represented not only by pixel values but also by spatial arrangements of roads, buildings, vegetation, water, and other landscape elements. This hypothetical methodological study examines deep learning techniques for automated classification of satellite imagery into eight land-use categories. A fictional dataset containing 6,000 optical satellite-image patches is divided into training, validation, and independent testing subsets. Three conceptual approaches are compared: support vector machine classification using handcrafted visual descriptors, a convolutional neural network trained specifically for the satellite dataset, and a transfer-learning convolutional network initialized from a large general-image recognition model.

The illustrative analysis evaluates overall accuracy, macro precision, macro recall, and macro F1 score. Under the assigned hypothetical values, the transfer-learning configuration produces the strongest overall classification performance. A second analysis examines class-specific precision, recall, and F1 score across urban built-up land, agricultural land, forest, water bodies, barren land, grassland, transportation corridors, and mixed residential-commercial patterns. Classification difficulty differs across categories according to hypothetical visual similarity and spatial heterogeneity. The study concludes that deep convolutional representations provide a useful methodological framework for automated interpretation of satellite imagery, particularly where spatial patterns are complex and handcrafted descriptors are insufficient. However, model reliability depends on representative training imagery, geographic independence of test data, seasonal variation, image resolution, class-definition quality, and domain shift. All numerical findings reported in this paper are illustrative and require verification using real remote-sensing imagery before practical conclusions can be established.

Keywords: Deep Learning, Satellite Imagery, Land-Use Classification, Remote Sensing, Convolutional Neural Networks, Transfer Learning, Image Classification, Earth Observation, Land-Cover Analysis, Geospatial Artificial Intelligence.

1. Introduction

Satellite remote sensing provides repeated observations of the Earth's surface and supports analysis of land-use and land-cover characteristics over large geographic areas. Optical imagery can represent urban development, agricultural fields, forests, water bodies, transportation infrastructure, and other spatial patterns that would be expensive to observe consistently through ground surveys alone. Automated classification can therefore assist large-scale mapping and monitoring by assigning semantic land-use categories to image regions according to their visual and spatial characteristics.

Traditional remote-sensing classification frequently relies on spectral information, texture descriptors, object-based representations, or manually engineered image features. Lu and Weng reviewed major image-classification techniques and emphasized the importance of feature selection, classifier design, and data quality. Object-based image analysis subsequently became influential for high-resolution remote sensing because pixels could be grouped into meaningful spatial objects before classification. Such methods remain valuable but can require substantial feature engineering and application-specific parameter selection.

Deep learning reduces dependence on manually specified feature representations by learning hierarchical visual patterns directly from imagery. Convolutional neural networks transformed general computer vision after large-scale image-recognition studies demonstrated their ability to learn progressively complex features from raw pixel arrays. Mnih and Hinton applied neural learning to road detection in high-resolution aerial imagery, demonstrating the potential of deep representations for remote-sensing interpretation. Later studies increasingly applied pretrained convolutional networks to land-use and scene-classification problems.

Castelluccio et al. examined convolutional neural networks for land-use classification and demonstrated the usefulness of pretrained architectures and fine-tuning. Penatti et al. investigated whether deep features learned from ordinary photographic imagery could transfer to aerial and remote-sensing scenes. Hu et al. similarly examined transfer of deep convolutional representations to high-resolution remote-sensing scene classification. Marmanis et al. extended this direction by applying ImageNet-pretrained networks to Earth-observation classification, showing how transfer learning can address limited labeled remote-sensing datasets.

The present hypothetical study compares handcrafted-feature classification, custom convolutional learning, and transfer learning for automated land-use classification from satellite images. A fictional eight-class dataset is used to illustrate a controlled methodological design with spatially independent data partitions. The study evaluates both overall and class-specific performance and discusses important challenges including spatial autocorrelation, seasonal variation, resolution differences, geographic transfer, and class ambiguity. No actual satellite image was classified for the numerical results presented here.

2. Objectives of the Study

The principal objective of this study is to examine deep learning approaches for automated classification of satellite imagery and land-use patterns. The investigation emphasizes representation learning, transfer learning, class-specific discrimination, geographic generalization, and methodological limitations affecting remote-sensing classification.

2.1 Examination of Deep Visual Representation Learning

The first objective is to examine how convolutional neural networks can learn spatial representations from satellite image patches without depending entirely on manually designed descriptors. Convolutional layers identify local visual structures such as edges, textures, boundaries, and repeated spatial arrangements, while deeper layers can represent combinations associated with roads, building clusters, vegetation patterns, field structures, and other land-use characteristics. Hierarchical representation learning is therefore especially relevant to high-resolution remote sensing.

The study also compares learned representations with handcrafted visual features. Traditional descriptors can be effective when classes have clear color, texture, or shape characteristics, but their representation capacity is constrained by predefined feature design. CNNs can potentially discover more complex spatial relationships automatically. Nevertheless, learned representations require adequate training diversity and may capture image-specific artifacts if data partitions are not designed carefully. The comparison therefore treats deep learning as an empirical alternative rather than assuming inherent superiority.

2.2 Evaluation of Transfer Learning for Satellite Classification

The second objective is to examine whether a convolutional network initialized from a large general-image dataset can provide useful features for satellite-image classification. Transfer learning is particularly relevant when labeled remote-sensing datasets are smaller than large natural-image collections. Early convolutional layers can encode general visual structures such as edges, corners, textures, and shapes that may remain useful even when the target imagery is acquired from an overhead perspective.

The study compares a network trained specifically from the hypothetical satellite dataset with a pretrained model that is subsequently fine-tuned. The comparison uses identical training, validation, and testing images. This methodological control is important because apparent transfer-learning benefits can otherwise arise from differences in data preparation or model capacity. The objective is to evaluate whether pretrained feature representations conceptually improve generalization when remote-sensing labels are comparatively limited.

2.3 Identification of Land-Use Classification Challenges

The third objective is to identify sources of error in satellite land-use classification. Different land-use categories can contain similar visual elements. Residential and commercial regions may both include buildings and roads, while grassland and agricultural areas may share vegetation-related spectral characteristics. Classification therefore depends on spatial arrangement, object density, season, resolution, and local landscape structure rather than isolated pixel values alone.

The study additionally considers geographic generalization. Neighboring satellite patches can be highly correlated because they share the same landscape, acquisition conditions, and local infrastructure. If nearby patches appear in both training and testing sets, model accuracy can become artificially high. The hypothetical design therefore assumes geographic separation among data partitions. Robust land-use classification should demonstrate performance on imagery from independent geographic areas rather than merely recognizing nearby examples.

3. Review of Literature

3.1 Conventional Remote-Sensing and Land-Use Classification

Lu and Weng reviewed remote-sensing image-classification methods and emphasized the importance of training data, feature selection, preprocessing, and classifier suitability. Traditional approaches have included maximum-likelihood classification, decision trees, neural networks, support vector machines, and knowledge-based procedures. High-resolution imagery increased the importance of spatial and textural information because individual scenes frequently contain several land-cover components arranged in characteristic land-use patterns.

Blaschke examined object-based image analysis as an important development in remote sensing, particularly for high-resolution imagery in which meaningful objects can extend across multiple pixels. Yang and Newsam subsequently applied bag-of-visual-words and spatial extensions to land-use classification in high-resolution overhead imagery. Their research demonstrated that spatial arrangements of local visual features can provide important discrimination among land-use classes and established a widely used methodological foundation for remote-sensing scene classification.

3.2 Deep Learning for Remote-Sensing Imagery

Mnih and Hinton applied a large neural network to road detection in high-resolution aerial imagery and demonstrated that deep representations could capture useful contextual information. Krizhevsky et al. later demonstrated the broader effectiveness of deep convolutional networks for large-scale image classification. The success of deep learning in general computer vision encouraged its transfer to remote-sensing applications, where spatial patterns can similarly be represented through hierarchical convolutional features.

Hu et al. applied deep convolutional networks to hyperspectral image classification, while Castelluccio et al. investigated CNN-based land-use classification using remote-sensing scenes. Penatti et al. evaluated the generalization of deep features from everyday imagery to aerial and remote-sensing domains. These studies provided evidence that representations learned by deep networks can outperform or complement handcrafted image descriptors, particularly when scene-level spatial structure contains important semantic information.

3.3 Transfer Learning and Semantic Interpretation

Hu et al. investigated transferred convolutional representations for high-resolution remote-sensing scene classification and showed that features extracted from pretrained deep networks could provide strong scene representations. Marmanis et al. similarly used ImageNet-pretrained networks for Earth-observation classification, illustrating how transfer learning can reduce problems associated with limited training data. Pretrained networks therefore provide an important alternative to training complex deep architectures entirely from remote-sensing datasets.

Deep learning has also extended from scene classification toward pixel-level semantic interpretation. Long et al. developed fully convolutional networks for semantic segmentation, while Marmanis et al. applied ensembles of CNNs to semantic segmentation of aerial imagery. These approaches demonstrate that deep remote-sensing analysis can operate at several spatial levels, ranging from complete scene categories to dense per-pixel land-cover labeling. The present paper focuses specifically on image-patch land-use classification.

4. Research Methodology

4.1 Research Design and Classification Models

The present study adopts a hypothetical comparative image-classification design involving three models. The first is a support vector machine operating on handcrafted color-histogram, local-texture, and edge-orientation features. This model provides a conventional machine-learning baseline. The second is a custom convolutional neural network trained from randomly initialized parameters using only the hypothetical satellite training dataset. The third uses a pretrained convolutional architecture followed by target-domain fine-tuning.

All models are evaluated using the same geographic training, validation, and testing partitions. Model-selection decisions are conceptually based only on training and validation results. The independent test images remain unavailable until final evaluation. Data augmentation for the CNN models includes horizontal and vertical reflection, moderate rotation, limited scaling, and brightness variation. Transformations are applied only when they preserve the conceptual land-use category of the image.

4.2 Hypothetical Satellite-Image Dataset

The fictional dataset contains 6,000 optical satellite-image patches at a conceptual spatial resolution of 1.5 meters per pixel. Each patch is represented as a 256×256 RGB image. Four thousand two hundred patches are allocated to training, nine hundred to validation, and nine hundred to independent testing. Images are assumed to originate from geographically separated fictional regions so that neighboring patches do not occur across different data partitions.

Eight land-use categories are represented: urban built-up land, agricultural land, forest, water bodies, barren land, grassland, transportation corridors, and mixed residential-commercial land. Class proportions are approximately balanced. Labels are assumed to have been assigned according to a predefined land-use interpretation guide.

4.3 Evaluation Metrics and Analytical Procedure

Overall classification performance is evaluated through accuracy, macro precision, macro recall, and macro F1 score. Macro averaging gives each land-use class equal contribution to the final metric and therefore prevents large classes from dominating the evaluation. The second analysis examines precision, recall, and F1 score separately for each category using the strongest hypothetical model. Class-specific analysis is important because overall accuracy can conceal systematic confusion among visually similar land-use patterns.

A real remote-sensing experiment would additionally report a full confusion matrix, class-support values, confidence intervals, repeated model training, and geographic cross-validation. Cohen's kappa or other agreement measures could also be included where methodologically appropriate. Because the present paper is hypothetical, numerical values are deliberately constructed examples designed to demonstrate reporting structure rather than to claim actual classification performance.

5. Analysis

5.1 Comparative Classification Performance

The first hypothetical analysis compares handcrafted-feature classification with two deep-learning configurations. The support vector machine achieves an assigned accuracy of 82.4%, demonstrating that

conventional visual descriptors can provide substantial discrimination among land-use scenes. The custom CNN improves illustrative accuracy to 89.8%, while the transfer-learning model records the highest assigned accuracy at 93.1%. Macro precision, recall, and F1 score demonstrate the same general pattern.

The hypothetical advantage of transfer learning is methodologically consistent with the idea that pretrained convolutional features provide useful low- and intermediate-level representations before domain-specific fine-tuning. However, these values do not establish that pretrained models always outperform custom networks. Performance depends on image resolution, domain similarity, available labels, network capacity, training strategy, and augmentation. A genuine comparison would require repeated training and statistical assessment of the resulting performance distributions.

Table 1: Hypothetical Comparative Performance for Satellite Land-Use Classification

Classification Model	Accuracy (%)	Macro Precision (%)	Macro Recall (%)	Macro F1 Score (%)
Handcrafted Features + SVM	82.4	82.1	81.7	81.9
Custom Convolutional Neural Network	89.8	89.6	89.2	89.4
Transfer-Learning Convolutional Network	93.1	92.9	92.7	92.8

Note: All model outputs and performance values are fictional and were not generated through actual satellite-image classification.

5.2 Class-Specific Land-Use Recognition

The second hypothetical analysis examines the transfer-learning model across the eight land-use categories. Water bodies and forest receive the strongest assigned F1 values because their visual characteristics are assumed to be comparatively distinct within the fictional image collection. Agricultural and urban built-up categories also achieve strong results. Mixed residential-commercial land receives the lowest F1 score because its images are assumed to contain combinations of structures that overlap with several other classes.

Barren land and grassland also produce lower hypothetical performance than the most distinctive classes. Their appearance can vary according to season, soil condition, vegetation density, and image illumination. The results illustrate why class-specific reporting is important in remote sensing.

Table 2: Hypothetical Class-Specific Performance of the Transfer-Learning Model

Land-Use Category	Precision (%)	Recall (%)	F1 Score (%)
Urban Built-Up	94.0	93.0	93.5
Agricultural Land	92.0	94.0	93.0
Forest	95.0	96.0	95.5

Water Bodies	97.0	96.0	96.5
Barren Land	90.0	88.0	89.0
Grassland	89.0	90.0	89.5
Transportation Corridors	92.0	91.0	91.5
Mixed Residential-Commercial	87.0	85.0	86.0

Note: All land-use classes, model predictions, and numerical values in this table are illustrative.

6. Discussion

6.1 Deep Representations and Satellite Scene Understanding

The hypothetical analysis illustrates the potential advantage of convolutional representation learning over fixed handcrafted descriptors. Satellite land-use patterns often depend on combinations of local and regional spatial structure. Urban scenes contain characteristic relationships among buildings, roads, vegetation, and open areas, while agricultural scenes may contain repeated field boundaries and regular textures. CNNs can learn hierarchical combinations of such patterns without requiring every relevant descriptor to be designed manually.

This interpretation is consistent with remote-sensing studies that transferred or trained convolutional models for aerial and satellite imagery. Castelluccio et al. demonstrated the effectiveness of CNN architectures for land-use classification, while Penatti et al. showed that deep representations learned from ordinary images could generalize to aerial-scene analysis. The benefit of learned features is therefore not limited to raw spectral discrimination but extends to increasingly complex spatial representation.

6.2 Transfer Learning and Limited Remote-Sensing Labels

The strongest hypothetical performance is assigned to transfer learning because pretrained convolutional networks begin with feature representations learned from much larger image collections. Hu et al. and Marmanis et al. showed that pretrained deep networks can provide effective features for high-resolution remote-sensing and Earth-observation classification. This is methodologically relevant because collecting satellite imagery may be relatively straightforward while generating reliable land-use labels often requires substantial expert effort.

The usefulness of transfer learning nevertheless depends on domain similarity. Natural photographs generally contain perspective views of individual objects, while satellite images provide overhead representations with different scales, orientations, and spatial structures. Fine-tuning is therefore necessary to adapt higher-level representations. In multispectral or hyperspectral applications, transfer is even more complex because standard pretrained networks generally expect three-channel imagery.

6.3 Class Confusion and Spatial Pattern Complexity

The hypothetical class-specific results demonstrate that land-use classification difficulty depends on internal class heterogeneity and interclass similarity. Water bodies may exhibit relatively consistent

color and texture characteristics, whereas mixed urban classes contain several shared objects. Residential, commercial, and transportation regions can all contain roads, vehicles, roofs, trees, and paved areas. Classification therefore requires understanding differences in spatial organization rather than merely detecting the presence of individual objects.

This distinction between land cover and land use is methodologically important. Land cover describes physical surface characteristics such as vegetation or water, whereas land use often reflects functional organization. A commercial district and residential district may contain similar physical materials but differ in density, layout, and surrounding infrastructure. Deep scene representations can potentially capture these broader relationships, although ambiguous labels and mixed scenes will continue to limit achievable classification accuracy.

7. Challenges

7.1 Geographic Dependence and Spatial Data Leakage

Satellite image datasets contain strong spatial autocorrelation. Adjacent image patches often share roads, buildings, vegetation, illumination conditions, and acquisition characteristics. Random image-level splitting can therefore place nearly identical neighboring scenes into both training and testing sets. A model may appear to generalize when it is partly recognizing geographic patterns already represented during training. Such evaluation can substantially overestimate operational performance.

Geographic separation should therefore be considered during dataset design. Training, validation, and testing regions can be divided by spatial blocks, administrative areas, or independent acquisition zones. Cross-region testing provides stronger evidence that a model learns transferable land-use representations. The present hypothetical methodology assumes this form of separation. Real research should report geographic partition procedures explicitly so that performance can be interpreted appropriately.

7.2 Seasonal, Atmospheric, and Resolution Variation

Satellite images can differ because of season, sun angle, atmospheric conditions, sensor characteristics, and spatial resolution. Agricultural land may change dramatically between planting, growth, harvest, and fallow periods. Forest color and canopy appearance can vary seasonally, while shadows in built-up areas depend on illumination geometry. A classifier trained on images from one acquisition period may therefore experience reduced accuracy on imagery captured under different conditions.

Spatial resolution also changes the visual representation of land-use patterns. At coarse resolution, individual buildings and roads may disappear into mixed pixels, while very-high-resolution imagery reveals detailed objects and texture. Deep architectures developed for one resolution should not be assumed to transfer directly to another. Multi-season and multi-resolution validation can therefore provide stronger evidence of robust remote-sensing classification.

7.3 Label Quality and Geographic Generalization

Land-use labeling can contain ambiguity because real landscapes rarely follow perfectly separated semantic categories. Urban boundaries may contain mixed residential, commercial, industrial, and transportation uses. Agricultural regions can contain buildings, trees, roads, and water channels. Assigning one label to an entire image patch may therefore simplify genuinely heterogeneous spatial information and introduce reference-label uncertainty.

Geographic generalization compounds this problem because land-use appearance differs among regions. Building materials, street layouts, vegetation, field geometry, and settlement density can vary across countries and climates. A model trained in one geographic environment may therefore perform less effectively elsewhere even when category names remain unchanged. Reliable automated land-use classification should consequently be validated on independently collected regions with meaningful geographic diversity.

8. Conclusion

Deep learning provides a powerful methodological framework for automated classification of satellite imagery and land-use patterns because convolutional networks can learn hierarchical spatial representations directly from image data. The present hypothetical study compared handcrafted-feature classification, a custom convolutional neural network, and a transfer-learning architecture across eight conceptual land-use categories.

The illustrative analysis assigned the strongest overall performance to the transfer-learning model and showed substantial variation among individual land-use categories. Water and forest scenes produced the strongest hypothetical recognition, while mixed residential-commercial areas were more difficult because of internal heterogeneity and visual overlap with related urban classes. Every numerical value is fictional and does not constitute empirical evidence from actual satellite imagery.

The study concludes that model architecture is only one component of reliable remote-sensing classification. Geographic independence, seasonal variability, spatial resolution, label quality, class definition, and domain shift must also be addressed. Future empirical research should evaluate deep-learning systems across independent regions, multiple acquisition periods, different spatial resolutions, and sufficiently diverse land-use conditions before automated classifications are used for operational mapping or decision support.

References

1. Adegun, A. A., Viriri, S., & Tapamo, J.-R. (2023). Review of deep learning methods for remote sensing satellite images classification: Experimental survey and comparative analysis. *Journal of Big Data*, 10, Article 93. <https://doi.org/10.1186/s40537-023-00772-x>
2. Cheng, G., Han, J., & Lu, X. (2017). Remote sensing image scene classification: Benchmark and state of the art. *Proceedings of the IEEE*, 105(10), 1865–1883. <https://doi.org/10.1109/JPROC.2017.2675998>
3. Cheng, G., Yang, C., Yao, X., Guo, L., & Han, J. (2018). When deep learning meets metric learning: Remote sensing image scene classification via learning discriminative CNNs. *IEEE Transactions on Geoscience and Remote Sensing*, 56(5), 2811–2821. <https://doi.org/10.1109/TGRS.2017.2783902>
4. Helber, P., Bischke, B., Dengel, A., & Borth, D. (2019). EuroSAT: A novel dataset and deep learning benchmark for land use and land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(7), 2217–2226. <https://doi.org/10.1109/JSTARS.2019.2918242>
5. 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 Letters*, 14(5), 778–782. <https://doi.org/10.1109/LGRS.2017.2681128>

6. Li, Y., Zhang, H., Xue, X., Jiang, Y., & Shen, Q. (2018). Deep learning for remote sensing image classification: A survey. *WIREs Data Mining and Knowledge Discovery*, 8(6), Article e1264. <https://doi.org/10.1002/widm.1264>
7. Paheding, S., Saleem, A., Siddiqui, M. F. 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. <https://doi.org/10.1007/s00521-024-10165-7>
8. Xia, G.-S., Hu, J., Hu, F., Shi, B., Bai, X., Zhong, Y., Zhang, L., & Lu, X. (2017). AID: A benchmark data set for performance evaluation of aerial scene classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(7), 3965–3981. <https://doi.org/10.1109/TGRS.2017.2685945>
9. Zhang, L., Zhang, L., & Du, B. (2016). Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and Remote Sensing Magazine*, 4(2), 22–40. <https://doi.org/10.1109/MGRS.2016.2540798>
10. Zhang, X., Han, L., Han, L., & Zhu, L. (2021). How well do deep learning-based methods for land cover classification and object detection perform on high resolution remote sensing imagery? *Remote Sensing*, 13(2), Article 210. <https://doi.org/10.3390/rs13020210>
11. Zhao, W., Yan, L., & Zhang, Z. (2023). Land use and land cover classification meets deep learning: A review. *Sensors*, 23(21), Article 8966. <https://doi.org/10.3390/s23218966>
12. Li, Z., Chen, B., Wu, S., Su, M., Chen, J. M., & Xu, B. (2024). Deep learning for urban land use category classification: A review and experimental assessment. *Remote Sensing of Environment*, 311, Article 114290. <https://doi.org/10.1016/j.rse.2024.114290>
13. Mei, S., Lian, J., Wang, X., Su, Y., Ma, M., & others. (2024). A comprehensive study on the robustness of deep learning-based image classification and object detection in remote sensing: Surveying and benchmarking. *Journal of Remote Sensing*, 4, Article 0219. <https://doi.org/10.34133/remotesensing.0219>
14. Rehman, M. Z. U., Islam, S. M. S., Blake, D., Ulhaq, A., & Janjua, N. (2025). Deep learning for land use classification: A systematic review of HS-LiDAR imagery. *Artificial Intelligence Review*, 58, Article 272. <https://doi.org/10.1007/s10462-025-11265-z>
15. AlAli, R. (2025). Artificial intelligence for land cover and land use classification in remote sensing: Review study. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-G-2025, 115–122. <https://doi.org/10.5194/isprs-archives-XLVIII-G-2025-115-2025>
16. Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., Li, C., Li, X., Fu, W., Liu, C., Xu, Y., Wang, X., Cheng, Q., Hu, L., Yao, W., ... Song, L. (2013). Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data. *International Journal of Remote Sensing*, 34(7), 2607–2654. <https://doi.org/10.1080/01431161.2012.748992>
17. Maggiori, E., Tarabalka, Y., Charpiat, G., & Alliez, P. (2017). Can semantic labeling methods generalize to any city? The Inria aerial image labeling benchmark. In 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 3226–3229). IEEE. <https://doi.org/10.1109/IGARSS.2017.8127684>
18. Sumbul, G., Charfuelan, M., Demir, B., & Markl, V. (2019). BigEarthNet: A large-scale benchmark archive for remote sensing image understanding. In 2019 IEEE International Geoscience and Remote

Sensing Symposium (IGARSS) (pp. 5901–5904). IEEE.
<https://doi.org/10.1109/IGARSS.2019.8898549>

19. Yang, Y., & Newsam, S. (2010). Bag-of-visual-words and spatial extensions for land-use classification. In Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems (pp. 270–279). Association for Computing Machinery.
<https://doi.org/10.1145/1869790.1869829>
20. Cheng, G., Han, J., & Lu, X. (2017). Remote sensing image scene classification: Benchmark and state of the art. Proceedings of the IEEE, 105(10), 1865–1883.
<https://doi.org/10.1109/JPROC.2017.2675998>