

# Deep Learning with Explicit Geographic Coordinate Embedding for Improved Remote Sensing Image Classification

Xiaokui XIE, Riming WANG, and Zhijun DAI

**Abstract**—Remote sensing imagery inherently contains geographic coordinate information, and the spatial distribution of land-cover types often exhibits pronounced regional heterogeneity. However, most deep learning approaches in remote sensing still follow computer vision paradigms, treating images as ordinary pixel matrices while ignoring their intrinsic geographic attributes, which limits both representation capacity and spatial generalization. To address this issue, we propose a spatially enhanced deep learning framework that explicitly incorporates geographic reference information. Specifically, the geographic coordinates of image centers are introduced as additional input channels into convolutional neural networks through a simple embedding mechanism, enabling the model to adaptively capture regional variations while maintaining a unified and backbone-agnostic architecture. Based on the EuroSAT dataset, we construct a new benchmark, termed Geo-EuroSAT, by embedding coordinate information and conduct systematic comparative experiments under multiple training-validation split strategies. Both ResNet-34 and EfficientNet-B3 backbones are evaluated to verify the general applicability of the proposed method. Experimental results demonstrate that the coordinate-aware models consistently outperform conventional baselines across all settings ( $p < 0.005$ ), with overall accuracy improvements ranging from 0.4% to 1.8%. The gains are especially pronounced under limited training samples and for vegetation-related classes such as *Pasture* and *Herbaceous Vegetation*. These findings indicate that explicitly integrating geographic reference into deep networks via a lightweight, backbone-agnostic embedding provides a principled and generalizable way to enhance large-scale land-cover classification and geographic process modeling.

**Index Terms**—Deep learning, geographic coordinates, land-cover classification, remote sensing, spatial heterogeneity

## I. INTRODUCTION

With the rapid growth of Earth observation technologies, the volume of remote sensing imagery has increased exponentially[1]. For instance, the Sentinel-2 constellation generates more than 1.6 TB of

multispectral imagery daily[2], and global Earth observation archives exceeded 1000 PB by 2019, expanding at an annual rate of 20–30%[3], [4]. These large-scale datasets provide critical support for ecological monitoring, resource management, and climate change studies [5], yet efficiently extracting geospatial information remains a major challenge[6], [7].

Deep convolutional neural networks have become indispensable for intelligent remote sensing interpretation due to their hierarchical feature learning ability [8] [9]. Recent advances have introduced multisource data fusion, optical-SAR integration, spectral-spatial coupling, and time-series analysis [10], [11], [12], [13], as well as advanced architectures such as attention mechanisms, graph convolutional networks, and Transformers[14], [15]. For example, mountain land-cover mapping has been enhanced through systematic comparisons of machine learning algorithms in complex terrain environments, demonstrating the importance of contextual modeling in heterogeneous landscapes[16], while integrating Sentinel-1 SAR and Sentinel-2 optical data has significantly improved crop classification accuracy[17]. Ensemble-style learning without explicitly combining multiple classifiers has also shown strong robustness in scene-level classification tasks[18]. Correlation-aware and Transformer-style networks further demonstrate the value of explicitly modeling spatial dependencies in high-resolution imagery[19].

While deep learning has achieved remarkable success in remote sensing image classification, most existing approaches follow the conventional computer vision paradigm by treating remote sensing imagery as ordinary pixel arrays and focusing primarily on spectral and textural features[20], [21]. Spatial or geographic context is often modeled only implicitly through large receptive fields, multi-scale convolution, attention mechanisms, or neighborhood aggregation strategies. Such implicit approaches allow the network to capture local spatial dependencies, but cannot directly model large-scale spatial gradients or regional environmental constraints. According to

Manuscript received \*\* This work was supported by the National Natural Science Foundation of China (Key Program) under Grant No. 41930537; by the National Natural Science Foundation of China under Grant No. 42366009; and by the Open Fund of the Guangxi Key Laboratory of Marine Environmental Change and Disaster in the Beibu Gulf under Grant No. 2021KA02. (Corresponding author: Zhijun DAI)

Xiaokui XIE is with Guangxi Key Laboratory of Marine Environmental Disaster Processes and Ecological Protection Technology /College of Resources and Environment, Beibu Gulf University, Qinzhou 535011, China (e-mail: xiexiaokui@gmail.com).

Riming WANG is with Guangxi Key Laboratory of Marine Environmental Disaster Processes and Ecological Protection Technology /College of Resources and Environment, Beibu Gulf University, Qinzhou 535011, China (e-mail: wangriming@bbgu.edu.cn).

Zhijun DAI is with State Key Laboratory of Estuarine and Coastal Research, East China Normal University, Shanghai 200062, China (e-mail: zjdai@sklec.ecnu.edu.cn).

The Geo-EuroSAT dataset supporting the findings of this study is openly available at Zenodo: <https://doi.org/10.5281/zenodo.16920405>.

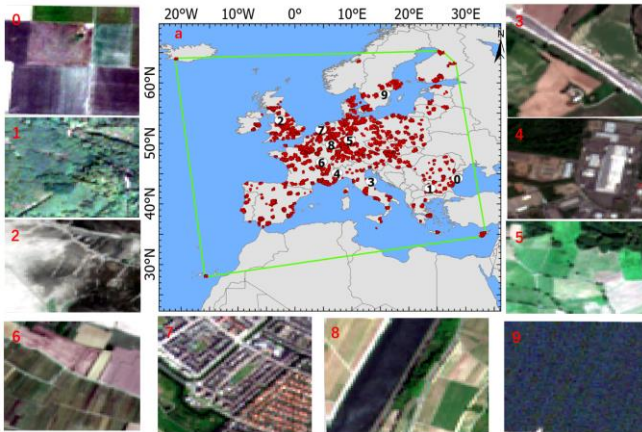
the first and second laws of geography [21], [22], geographic phenomena such as climate, soil, and vegetation typically vary across space, making spatial awareness essential for robust classification.

To address this limitation, we propose a deep learning approach that explicitly incorporates geographic coordinates. By explicitly embedding latitude and longitude as additional input channels to a convolutional neural network, the model learns region-aware feature sensitivities under a unified architecture, effectively realizing adaptive submodels without manual spatial partitioning. Based on the EuroSAT dataset[23], we construct an extended dataset, Geo-EuroSAT[24], and systematically evaluate the impact of coordinate embedding on classification accuracy and generalization using ResNet-34 and EfficientNet-B3.

## II. DATA AND METHODS

### A. Data Sources

This study utilizes the EuroSAT dataset[23], a widely adopted benchmark for land use/land cover (LU/LC) classification derived from Sentinel-2 imagery provided by the European Space Agency (ESA). The dataset spans both urban and rural areas across 34 European countries, ensuring strong geographic representativeness and class diversity. As illustrated in Fig. 1, the samples are distributed throughout Europe, covering a broad spectrum of landscapes ranging from high-latitude forests to Mediterranean coastal regions.



**Fig. 1.** The EuroSAT dataset. (a) Spatial distribution of samples across Europe. (b) Representative examples of the ten land cover classes (0–9), with class codes: 0-Annual Crop, 1-Forest, 2-Herbaceous Vegetation, 3-Highway, 4-Industrial, 5-Pasture, 6-Permanent Crop, 7-Residential, 8-Sea/Lake, and 9-River

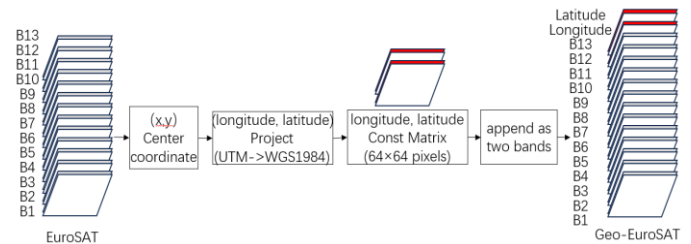
The EuroSAT dataset contains 27,000 multispectral  $64 \times 64$  image patches across ten balanced LU/LC classes (2,500–3,000 samples each), such as crops, forests, pastures, residential areas, and water. This balance reduces class imbalance and supports fair evaluation of spatial reference enhancement. All bands were resampled to 10 m for deep learning input.

The original EuroSAT images include georeferenced metadata in WGS 1984 UTM. To exploit this, we built **Geo-EuroSAT** by embedding explicit spatial coordinates as additional features.

### B. Spatial Reference Feature Encoding

To integrate spatial information into model learning, we embed the image center’s latitude and longitude as additional feature bands into the CNN. The procedure is as follows:

1. **Projection and Coordinate Extraction:** Extract UTM-projected coordinates from Sentinel-2 metadata and convert central pixel positions to WGS 1984 latitude–longitude using GDAL, ensuring a unified spatial reference.
2. **Normalization and Multi-channel Construction:** Normalize latitude and longitude to (0, 1), generate two  $64 \times 64$  constant matrices with the center’s values, and stack them with the 13 spectral bands to form a 15-channel dataset, **Geo-EuroSAT**(Fig. 2).



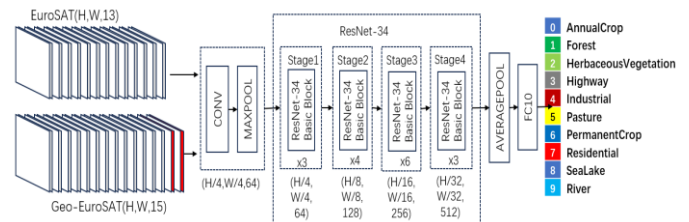
**Fig. 2.** Workflow of spatial reference feature extraction

Latitude/longitude channels have zero spatial gradient inside a patch. However, they do not act merely as a bias term. Instead, they provide a region-level conditioning signal that modulates convolutional feature responses. In other words, the coordinates allow the network to adaptively adjust channel sensitivities and learn region-specific feature preferences under a shared architecture.

### C. Deep Learning Model Training

#### 1. Model Selection and Architecture

To assess the effect of explicit geographic coordinates, we trained a unified deep CNN on both EuroSAT and Geo-EuroSAT. ResNet-34[25] was adopted as the backbone, consisting of residual convolutional blocks, pooling layers, and a fully connected classification head. Residual connections ease gradient vanishing, convolutions extract multi-scale features, pooling downsamples, and the final layer outputs class probabilities for ten land-use categories (Fig. 3).



**Fig. 3.** ResNet-34 model for 10-class classification using EuroSAT and Geo-EuroSAT datasets

#### 2. Training Procedure

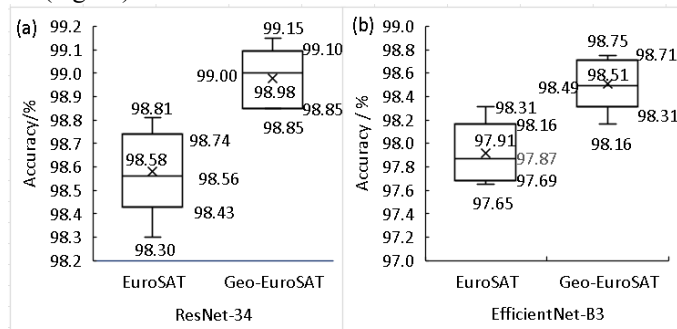
The ResNet-34 input layer was extended from 3 channels to 13 (EuroSAT) or 15 (Geo-EuroSAT), with ImageNet-pretrained weights transferred to RGB channels and the others randomly initialized, following multi-channel transfer learning practices[26], [27]. Both datasets were split 9:1 into training

and test sets using a fixed random seed for reproducibility. Training used a batch size of 64 for up to 60 epochs with early stopping (patience = 5). Data augmentation included random rotations ( $\pm 30^\circ$ ), brightness/contrast adjustments (0.8–1.2), scaling (0.9–1.1), and random cropping/translation. Cross-entropy loss was adopted, and all experiments were implemented in ArcGIS Pro. We also adopt EfficientNet-B3[28] as a second backbone to evaluate whether the effectiveness of coordinate embedding is consistent across architectures.

### III. RESULTS AND ANALYSIS

#### A. Results and Statistical Analysis

Given the inherent randomness in deep learning training, model outputs may vary slightly even under identical hyperparameter settings. To mitigate this stochastic effect, five independent training runs were conducted on both EuroSAT and Geo-EuroSAT, and classification accuracies were recorded on the test set (Fig. 4a).



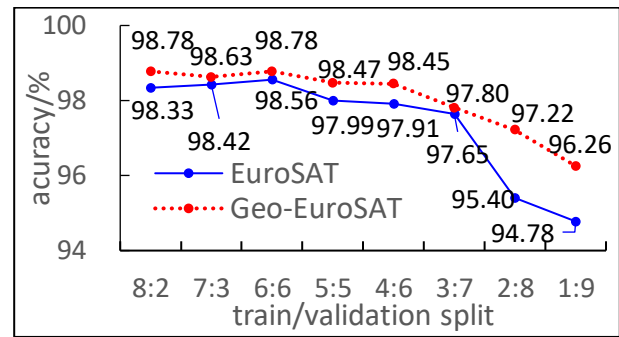
**Fig. 4.** Classification accuracies (%) of EuroSAT and Geo-EuroSAT across five independent experiments by ResNet-34 (a) and EfficientNet-B3 (b)

For EuroSAT by ResNet-34, accuracies ranged from 98.30%–98.81% (mean = 98.58%, SD = 0.19%), while Geo-EuroSAT achieved 98.85%–99.15% (mean = 98.98%, SD = 0.13%). Compared to EuroSAT, Geo-EuroSAT improved mean accuracy by 0.40%, reduced variance by 31.6%, and narrowed the accuracy range from 0.51% to 0.30%, indicating enhanced stability. Both models outperformed the official EuroSAT benchmark (98.57%), confirming the robustness of the adopted framework. An F-test confirmed homogeneity of variances ( $F = 2.13 < F_{0.05}(4,4) = 6.39$ ,  $p > 0.05$ ), and a paired t-test showed a significant performance gain for Geo-EuroSAT ( $t = 5.92$ ,  $p = 0.0042$ ).

Using EfficientNet-B3, Geo-EuroSAT consistently outperformed EuroSAT across five runs, with a higher mean accuracy (98.51% vs. 97.91%) (Fig. 4b). An independent-samples *t*-test shows the improvement is statistically significant (mean gain = 0.59%,  $p < 0.01$ ).

#### B. Impact of Training–Validation Split Ratios on Accuracy

To further evaluate the robustness of the proposed approach under varying training sample sizes, the dataset was partitioned into training and validation sets with eight different ratios ranging from 8:2 to 1:9. As illustrated in Fig. 5, the Geo-EuroSAT models consistently outperformed the EuroSAT models across all partitioning schemes.

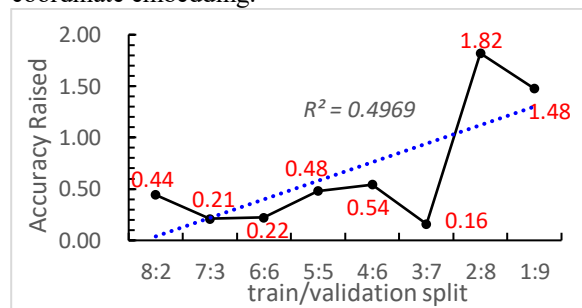


**Fig. 5.** Classification accuracy (%) of different training-test splits on the EuroSAT and Geo-EuroSAT dataset by ResNet-34

A paired-sample *t*-test revealed an average improvement of 0.67% (95% CI: 0.45%–0.88%), with a standard deviation of 0.30%,  $t = 6.29$ , and  $p = 0.00034$  ( $p < 0.001$ ). These results indicate that the explicit incorporation of geographic coordinate information significantly enhances classification accuracy across different training–validation split conditions, particularly under limited training data scenarios.

#### C. Adaptability of Spatial Reference Enhancement Under Varying Training Data Scales

To explore the potential of spatial reference enhancement under data-scarce conditions, we further analyzed the relative accuracy improvements of Geo-EuroSAT over EuroSAT across different training–validation split ratios. The results (Fig. 6) reveal a negative correlation ( $R^2 = 0.499$ ) between the accuracy gain and the proportion of training samples, indicating that the fewer the training samples, the greater the relative benefit of coordinate embedding.



**Fig. 6.** Accuracy gains of Geo-EuroSAT across varying training/testing partition ratios by ResNet-34

Specifically, when the training/validation ratios were 2:8 and 1:9, the classification accuracy improvements reached 1.82% and 1.48%, respectively, which are substantially higher than the typical gains (<0.5%) observed under large-sample conditions. This demonstrates that geographic coordinate information can effectively compensate for insufficient training data, thereby enhancing the model’s discriminative power and generalization performance.

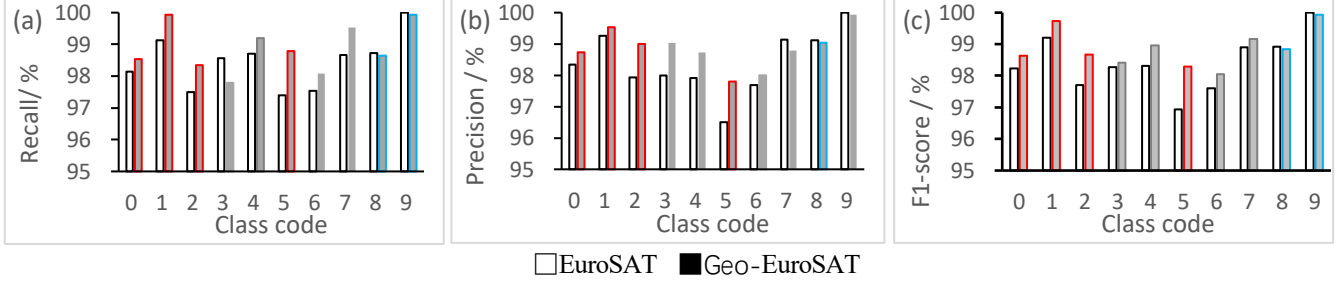
## IV. DISCUSSION

#### A. Response of Spatial Enhancement Across Land-Cover Classes

Confusion matrices were generated by ResNet-34 from five independent runs on EuroSAT and Geo-EuroSAT, Precision /

Recall/ F1-scores were computed for each class (Fig. 7). Geo-EuroSAT achieved consistent F1-score improvements across most classes, with the largest gains in Annual Crop(+0.40%), Forest (+0.53%), *Herbaceous Vegetation* (+0.96%), and *Pasture* (+1.36%) (Fig.7c). These vegetation-related categories are shaped by climatic, edaphic, and topographic gradients, and thus benefit from coordinate embedding. The slight performance decrease

observed for the Sea/Lake and River classes is mainly attributed to the scene-level annotation scheme of EuroSAT. An analysis of the confusion matrices shows frequent confusion between water-related classes and Highway, Industrial, and Pasture. By inspecting the misclassified samples, we found that these patches often contain mixed land-cover elements, including water bodies, roads, and vegetation.



**Fig. 7.** Precision, recall and harmonic mean F1-scores across 10 classes (averaged over five models) by ResNet-34. Class codes: 0-Annual Crop, 1-Forest, 2-Herbaceous Vegetation, 3-Highway, 4-Industrial, 5-Pasture, 6-Permanent Crop, 7-Residential, 8-Sea/Lake, and 9-River

### B. Essence of Spatial Reference-Enhanced Deep Learning

Spatial heterogeneity, regarded as the *second law of geography*[21], emphasizes that geographic phenomena exhibit pronounced non-uniformity with respect to location. Processes such as climate, soil, and vegetation often follow region-specific patterns, necessitating localized modeling. For example, the FLAASH atmospheric correction algorithm requires distinct atmospheric parameters for tropical or temperate, urban or rural zones. Yet, explicit partitioning increases workload and risks overfitting in subregions with limited samples, whereas a single global function  $z=f(p)$  ignores spatial context and weakens local predictive performance.

To address this dilemma, this study introduces geographic coordinates  $(x, y)$  as explicit input features, thereby constructing a spatial reference-enhanced deep learning framework. The implicit mathematical equivalence can be expressed as:

$$z(x, y, p) = \begin{cases} f_A(p), & \text{if } (x, y) \in A \\ f_B(p), & \text{if } (x, y) \in B \\ f_C(p), & \text{if } (x, y) \in C \\ f_D(p), & \text{if } (x, y) \in D \end{cases} \quad (1)$$

where  $f_i$  denotes local submodels corresponding to subregions  $A, B, C, D$ . Since geographic phenomena are generally governed by similar physical mechanisms, all submodels share the same structure, differing only in parameter weights. For regions lacking relevant samples, these weights effectively approach zero.

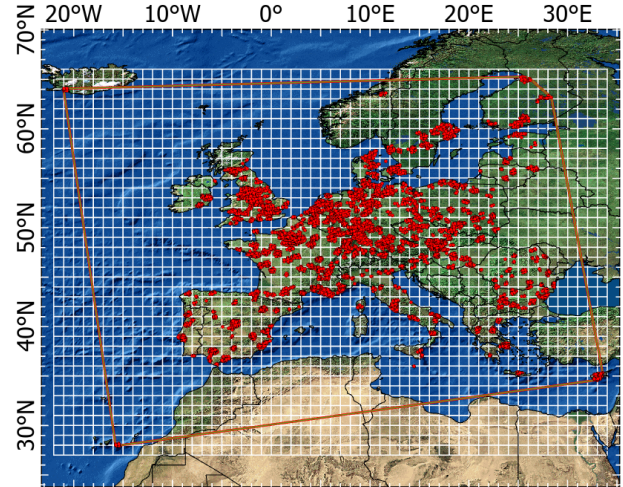
Within a deep neural network, this mechanism manifests as the integration of  $(x, y)$  coordinates alongside spectral bands into the input tensor, thereby participating in each convolutional operation:

$$\hat{F}^{(l)} = \sigma(W^{(l)} * [B, p, x, y] + b^{(l)}) \quad (2)$$

where  $[B, p, x, y]$  represents the multi-channel tensor containing spectral, textural, and geographic coordinate features,  $\sigma$  is the activation function, and  $*$  denotes convolution.

This formulation enables the network to learn an implicit mapping between coordinates and class labels by adaptively adjusting channel sensitivities. Conceptually, it is equivalent to

creating localized grid-based submodels within a unified architecture (Fig. 8), which share global feature extraction while dynamically reweighting regional parameters. Experiments confirm that this mechanism improves classification of spatially heterogeneous classes—such as *Pasture* and *Herbaceous Vegetation*—demonstrating the effectiveness of coordinate-driven adaptive learning for large-scale remote sensing interpretation.



**Fig. 8.** Implicit grid-based submodels in spatial enhanced deep learning

### V. CONCLUSIONS

(1) Embedding explicit geographic coordinates enhances both classification accuracy and stability, delivering consistent performance gains across different backbones (ResNet-34 and EfficientNet-B3). This demonstrates the robustness and model-agnostic effectiveness of the proposed approach.

(2) The method proves particularly effective for vegetation-related classes (e.g., *Pasture*, *Herbaceous Vegetation*), where spatial heterogeneity is pronounced.

(3) Under small-sample conditions, coordinate embedding offers notable advantages, compensating for limited training data and enhancing model generalization.

In this study, training and validation samples are randomly split from the full dataset, which primarily evaluates average classification performance under independent and identically distributed assumptions. A more rigorous assessment of geographic generalization would require spatial cross-validation, such as block-based or region-holdout schemes, where entire geographic areas are excluded from training and used only for testing. Implementing such spatially disjoint validation will be an important direction of future work to further examine the ability of coordinate embedding to support spatial extrapolation to unseen regions.

In addition, future work will explore extending this framework to multimodal remote sensing data (e.g., integrating optical, SAR, and LiDAR), which introduces challenges in coordinate alignment and sensor-specific normalization. For large-scale monitoring tasks, issues such as computational overhead and global coordinate normalization also need to be addressed to ensure efficiency and robustness at continental or global scales.

#### ACKNOWLEDGMENT

The authors thank Prof. Qimin Qin from Peking University and Prof. Hui Lin from Beibu Gulf University for their insightful discussions, valuable guidance, and encouragement.

#### REFERENCES

- [1] X. Deng *et al.*, "Geospatial Big Data: New Paradigm of Remote Sensing Applications," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 10, pp. 3841–3851, 2019, doi: 10.1109/JSTARS.2019.2944952.
- [2] M. Drusch *et al.*, "Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services," *Remote Sens. Environ.*, vol. 120, pp. 25–36, 2012, doi: 10.1016/j.rse.2011.11.026.
- [3] A. S. Belward and J. O. Skoien, "Who launched what, when and why: trends in global land-cover observation capacity from civilian earth observation satellites," *ISPRS J. Photogramm. Remote Sens.*, vol. 103, pp. 115–128, 2015, doi: 10.1016/j.isprsjprs.2014.03.009.
- [4] Z. Zhu *et al.*, "Benefits of the free and open Landsat data policy," *Remote Sens. Environ.*, vol. 224, pp. 382–385, 2019, doi: 10.1016/j.rse.2019.02.016.
- [5] Y. Ma *et al.*, "Remote sensing big data computing: Challenges and opportunities," *Future Gener. Comput. Syst.*, vol. 51, pp. 47–60, 2015, doi: 10.1016/j.future.2014.10.029.
- [6] R. Dian, S. Li, B. Sun, and A. Guo, "Recent advances and new guidelines on hyperspectral and multispectral image fusion," *Information Fusion*, vol. 69, pp. 40–51, 2021, doi: 10.1016/j.inffus.2020.11.001.
- [7] S. L. Ustin and E. M. Middleton, "Current and near-term advances in Earth observation for ecological applications," *Ecological Processes*, vol. 10, no. 1, p. 1, 2021, doi: 10.1186/s13717-020-00255-4.
- [8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.
- [9] X. Zhu *et al.*, "Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8–36, 2017, doi: 10.1109/MGRS.2017.2762307.
- [10] J. Zhang, "Multi-source remote sensing data fusion: Status and trends," *International Journal of Image and Data Fusion*, vol. 1, pp. 5–24, 2010, doi: 10.1080/19479830903561035.
- [11] A. Gettelman *et al.*, "The future of Earth system prediction: Advances in model-data fusion," *Sci. Adv.*, vol. 8, no. 14, p. eabn3488, 2022, doi: 10.1126/sciadv.abn3488.
- [12] G. Vivone, "Multispectral and hyperspectral image fusion in remote sensing: A survey," *Information Fusion*, vol. 89, pp. 405–417, 2023, doi: 10.1016/j.inffus.2022.08.032.
- [13] B. Chen, B. Huang, and B. Xu, "Multi-source remotely sensed data fusion for improving land cover classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 124, pp. 27–39, 2017, doi: 10.1016/j.isprsjprs.2016.12.008.
- [14] S. Jozdani, D. Chen, D. Pouliot, and B. Alan Johnson, "A review and meta-analysis of Generative Adversarial Networks and their applications in remote sensing," *International Journal of Applied Earth Observation and Geoinformation*, vol. 108, p. 102734, 2022, doi: 10.1016/j.jag.2022.102734.
- [15] H. Shen, T. Zhong, Y. Jia, and C. Wu, "Remote sensing image dehazing using generative adversarial network with texture and color space enhancement," *Scientific Reports*, vol. 14, no. 1, p. 12382, 2024, doi: 10.1038/s41598-024-63259-6.
- [16] G. Amin, I. Intiaz, E. Haroon, N. us Saqib, M. I. Shahzad, and M. Nazeer, "Assessment of Machine Learning Algorithms for Land Cover Classification in a Complex Mountainous Landscape," *J. Geovis. Spat. Anal.*, vol. 8, no. 2, p. 34, Aug. 2024, doi: 10.1007/s41651-024-00195-z.
- [17] A. Chakhar, D. Hernández-López, R. Ballesteros, and M. A. Moreno, "Improving the Accuracy of Multiple Algorithms for Crop Classification by Integrating Sentinel-1 Observations with Sentinel-2 Data," *Remote Sensing*, vol. 13, no. 2, Jan. 2021, doi: 10.3390/rs13020243.
- [18] P. Dou, C. Huang, W. Han, J. Hou, Y. Zhang, and J. Gu, "Remote sensing image classification using an ensemble framework without multiple classifiers," *ISPRS J. Photogramm. Remote Sens.*, vol. 208, pp. 190–209, Feb. 2024, doi: 10.1016/j.isprsjprs.2023.12.012.
- [19] D. He, Q. Shi, J. Xue, P. M. Atkinson, and X. Liu, "Very fine spatial resolution urban land cover mapping using an explicable sub-pixel mapping network based on learnable spatial correlation," *Remote Sens. Environ.*, vol. 299, p. 113884, Dec. 2023, doi: 10.1016/j.rse.2023.113884.
- [20] J. Deng, W. Dong, R. Socher, L. Li, K. Li, and F. Li, "ImageNet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255. doi: 10.1109/CVPR.2009.5206848.
- [21] M. F. Goodchild, "The Validity and Usefulness of Laws in Geographic Information Science and Geography," *Annals of the Association of American Geographers*, vol. 94, no. 2, pp. 300–303, 2004, [Online]. Available: <http://www.jstor.org/stable/3693988>
- [22] W. R. Tobler, "A Computer Movie Simulating Urban Growth in the Detroit Region," *Econ. Geogr.*, vol. 46, pp. 234–240, 1970, doi: 10.2307/143141.
- [23] P. Helber, B. Bischke, A. Dengel, and D. Borth, "EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 12, no. 7, pp. 2217–2226, Jul. 2019, doi: 10.1109/JSTARS.2019.2918242.
- [24] X. XIE, "Geo-EuroSAT: An Extended EuroSAT Dataset with Explicit Geographic Coordinate Embedding for Land Cover Classification," Aug. 2025, Accessed: Aug. 22, 2025. [Online]. Available: <https://zenodo.org/records/16920405>
- [25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.
- [26] D. Ienco, R. Gaetano, C. Dupaquier, and P. Maurel, "Land Cover Classification via Multitemporal Spatial Data by Deep Recurrent Neural Networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, 2017, doi: 10.1109/lgrs.2017.2728698.
- [27] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, "Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 778–782, 2017, doi: 10.1109/LGRS.2017.2681128.
- [28] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *Proc. 36th Int. Conf. Mach. Learn. (ICML)*, K. Chaudhuri and R. Salakhutdinov, Eds., PMLR, May 2019, pp. 6105–6114. Accessed: Feb. 01, 2026. [Online]. Available: <http://proceedings.mlr.press/v97/tan19a.html>