Remote Sensing of Grasslands: Performance Comparison of Radar and Optical Data in Machine Learning Classification

Konstantinos Christofi[1], Charalambos Chrysostomou[1*], Iason Tsardanidis[3], Michalis Mavrovouniotis[1], Giorgia Guerrisi[4], Charalampos Kontoes[3], Diofantos G. Hadjimitsis[1,2]

[1] ERATOSTHENES Centre of Excellence - konstantinos.christofi@eratosthenes.org.cy
[2] Dept. of Civil Engineering and Geomatics, Cyprus University of Technology - d.hadjimitsis@cut.ac.cy
[3] Institute for Astronomy, Astrophysics, Space Applications and Remote Sensing, National Observatory of Athens (IAASARS/NOA) - j.tsardanidis@noa.gr
[4] Dept. of Civil Engineering and Computer Science Engineering, Tor Vergata University of Rome - giorgia.guerrisi@uniroma2.it

Keywords: Grassland Classification, Remote Sensing, Machine Learning, Sentinel-1 SAR, Sentinel-2 Optical Imagery, Ecosystem Monitoring

Abstract

Classification of grasslands has an important role in environmental monitoring, and management. This study compares and evaluates the performance of various machine learning and deep learning algorithms in grassland classification using remote sensing data from Sentinel-1 and Sentinel-2 satellites. Sentinel-1 satellite provide Synthetic Aperture Radar data, which captures structural and moisture-related information. Sentinel-2 captures high-resolution optical images with rich spectral details. Both datasets from Sentinel-1 and Sentinel-2 satellites were used to train and evaluate a variety of machine learning models including Random Forest, Support Vector Machines, Logistic Regression, XGBoost and Deep Neural Networks. The results of this study show that Random Forest performs best on Sentinel-1 data and Neural Networks perform best when it comes to grassland classification using Sentinel-2 data. These results show how important it is to select a model based on the characteristics and the nature of the dataset.

1. Introduction

Grasslands are one of the major ecosystems in the world, supporting different species and also play an important role in carbon sequestration, helping mitigate change (d’Andrimont et al. (2018); Ojima et al. (1993); Bengtsson et al. (2019)). Accurate classification of grasslands is important for effective monitoring, management since it can allow informed decision-making for sustainable land use practices. In addition to their importance in ecological terms, grasslands also provide ecosystem services such as water management, control of soil erosion. and landscape stability (Boval and Dixon (2012); Carlier et al. (2009)). Farmers are encouraged to maintain permanent grasslands in order to enhance biodiversity and carbon capture by policies like the Common Agricultural Policy (CAP) (d’Andrimont et al. (2018)).

Remote sensing has become a key aspect when monitoring grasslands, since it can provide tools to evaluate cover, biomass, and degradation over large areas (Yong (2003)). Satellite data give the opportunity to researchers to monitor changes in grassland ecosystems using techniques such as vegetation indices and texture analysis (Guo et al. (2004); Weeks et al. (2013)). Advances in satellite technology now allow us to extract grassland characteristics with improved accuracy.

This study explores the use of Machine Learning (ML) and Deep Learning (DL) techniques for classifying grasslands based on remote sensing data retrieved from the Sentinel-1 (fig.1(a)) and Sentinel-2 (fig.1(b)) satellites. Sentinel-1’s Synthetic Aperture Radar (SAR) data, provide valuable insights into vegetation structure and moisture levels while being unaffected by cloud cover or lighting conditions. Sentinel-2 captures high-resolution optical images with a large scale of spectral information across multiple bands, giving the opportunity for a detailed analysis of vegetation health and land cover. This spectral richness improves the ability to find various differences in grassland composition and vitality.

This study evaluates the performance of these two data sources in grassland classification, highlighting their strengths and limitations. By understanding these differences, it can help validate farmer-provided data, support better informed grassland management practices, and also support more effective conservation efforts.

2. Literature Review

Remote sensing has advanced significantly since Machine Learning (ML) and Deep Learning (DL) have also advanced, along with the application of remote sensing data for classifying grasslands. European Space Agency’s Sentinel-1 and Sentinel-2 satellites, that are part of the Copernicus program, provide important data that can be used for the aforementioned purpose. Sentinel-1’s Synthetic Aperture Radar (SAR) captures detailed information about earth’s surface structure and texture. On the other hand, Sentinel-2’s multi-spectral optical images, also provide detailed information about earth’s surface but with a focus on vegetation, land cover, and water bodies. Researchers have explored how well different ML and DL algorithms perform on

* Corresponding author

https://doi.org/10.5194/isprs-archives-XLVIII-G-2025-295-2025

(a) Sentinel-1: Synthetic Aperture Radar (SAR) image

(b) Sentinel-2: Multi-spectral optical image

Figure 1. Sentinel-1 and Sentinel-2 satellite images from the Netherlands region of interest (ROI), as used in our dataset.

these datasets, to find the best models for grassland classification.

Sentinel-1 SAR data is particularly useful for areas with regular cloud cover since it provides observations in all weather conditions, day and night. However, there are particular challenges when using SAR data, such as speckle noise and complex backscatter characteristics, which require specific processing techniques. Several studies have shown how well machine learning models and deep learning architectures handle these challenges for vegetation and land-cover classification.

A thorough analysis of Random Forest (RF) for remote sensing applications was carried out by Belgiu and Drăguţ (2016), which emphasized the model's ability to handle high-dimensional data and reduce overfitting. The authors of the study highlighted RF's effectiveness in classification using SAR data, due to its ability to incorporate multiple polarization features. These findings were further supported by Alharbi (2024), who analyzed the impact of decision tree-based models like RF and XGBoost (XGB) in handling radar-based backscatter variations. In their study they found that although XGB performed well in structured land types, RF outperformed XGB in classifying grassland areas due to its ensemble learning nature.

Beyond decision-tree models, Pelletier et al. (2019) explored the application of Temporal Convolutional Neural Networks (TempCNNs) on Sentinel-1 time-series data. They demonstrated that CNNs could effectively capture seasonal variations in radar backscatter, enhancing grassland classification accuracy. However, while CNNs excelled in learning spectral-temporal patterns, RF remained a computationally efficient alternative with comparable accuracy.

In addition to supervised learning techniques, Guo et al. (2004) explored SAR-derived texture measures for vegetation classification. Their results showed that integrating backscatter intensity with polarization information significantly improved model accuracy, reinforcing the idea that Sentinel-1 data contains structural and moisture-related features valuable for distinguishing grassland cover.

These studies indicate that Random Forest is a strong contender for Sentinel-1-based classification, consistently outperforming other traditional ML models in terms of stability and interpretability. While CNNs show promise in handling time-series SAR data, RF remains the preferred choice for standalone Sentinel-1 datasets due to its efficiency and adaptability to noisy radar data.

Sentinel-2's 13-band multispectral images rich information for vegetation monitoring, biomass estimation and land-cover classification. Sentinel-2 is well-suited for classifying grasslands since it has the ability to directly measure vegetation reflectance, in contrast to Sentinel-1, which records surface structure and moisture.

Nalepa et al. (2019) investigated the potential of Deep Neural Networks (DNNs) in their study for classification using Sentinel-2 data. Their research showed that when trained on the high-dimensional spectral data from Sentinel-2, DNNs performed better than traditional ML models. According to the results of the study, the neural networks can effectively model complex relationships within multispectral imagery, resulting to the improvement of classification accuracy.

Zhu et al. (2017) offered a broader perspective by reviewing deep learning applications in remote sensing. Their findings emphasized the superiority of Convolutional Neural Networks (CNNs) for spectral-spatial feature extraction, making them particularly effective for vegetation classification tasks. Their analysis showed that CNNs consistently outperformed RF and SVM when applied to high-resolution Sentinel-2 imagery.

Building on these insights, Helber et al. (2019) introduced EuroSAT, a benchmark dataset derived from Sentinel-2, which facilitated extensive land-use classification research. Their study revealed that CNN-based models trained on Sentinel-2 data achieved classification accuracies exceeding 98%, highlighting the benefits of deep learning in multispectral analysis.

Furthermore, Weeks et al. (2013) explored the impact of vegetation indices on Sentinel-2 classification accuracy. They demonstrated that incorporating indices such as $NDVI$ and Red Edge bands significantly enhanced model performance, particularly in discriminating between healthy and degraded grasslands.

These studies collectively suggest that deep learning models, especially CNNs, surpass traditional ML approaches for Sentinel-2-based classification by effectively utilizing high-dimensional spectral information. However, RF remains a viable alternative, particularly when computational efficiency and interpretability are essential considerations.

https://doi.org/10.5194/isprs-archives-XLVIII-G-2025-295-2025

3. Data

3.1 Space-level datasets

The dataset used in this study is a multi-level and multi-sensor resource provided by (Choumos et al. (2022)), and it aims at supporting grassland classification for agricultural monitoring. The dataset consists of Earth observation data extracted from the Sentinel-1 and Sentinel-2 satellites, part of the Copernicus program (Jutz and Milagro-Perez (2020)) managed by the European Space Agency (ESA). It is a pre-annotated dataset where Sentinel-1 and Sentinel-2 data were georeferenced and labeled with crop-type information from the Dutch Land Parcel Identification System (LPIS). Sentinel-1 and Sentinel-2 satellites provide complementary data, with Sentinel-1 capturing Synthetic Aperture Radar (SAR) data, and Sentinel-2 providing multi-spectral optical imagery data. Both of these data modalities contain data from 2017 in a span of 7 months.

3.1.1 Sentinel-1 The Synthetic Aperture Radar (SAR) data that Sentinel-1 data consist of, were gathered across several months in 2017. Sentinel-1 functions in the microwave spectrum, collecting data day and night in any weather situation compared to optical satellites. The current dataset, which was gathered for every month from April to October in 2017, has two different polarization modes: Vertical-Vertical (VV) and Vertical-Horizontal (VH), and two coherence values: Horizontal-Horizontal (HH) and Horizontal-Vertical (HV). The dataset contains different observation points, denoted by parcel_ids - the id of the polygon for the corresponding agricultural parcel - and the labels of these points, indicate whether the vegetation is dominated by grass or not.

The radar backscatter coefficients for each month and polarization are shown in the columns with names like 2017-04_VH_SAR, 2017-04_VV_SAR, etc. The strength of the radar signal reflected from the surface is indicated by the backscatter values, which can be influenced by vegetation structure, moisture content, and surface roughness.

3.1.2 Sentinel-2 Sentinel-2 provides 13 spectral bands of high-resolution optical imaging, including visible, near-infrared, and short-wave infrared bands (see Table 1). The table presents Sentinel-2's 13 spectral bands, which are used for a variety of Earth Observation applications. These bands cover various portions of the electromagnetic spectrum, including visible, near-infrared, and short-wave infrared regions. Every band is tailored to perform particular observational tasks, such as identifying water bodies, vegetation, or atmospheric properties. In this study, we utilize all bands except B1, B9, and B10, to extract meaningful spectral information for the classification of grassland areas.

The dataset contains of multi-spectral optical data that were gathered in 2017 for every month from March until the end of October. Like the Sentinel-1 dataset, each observation in the dataset is represented by a distinct id and a corresponding label that shows whether the area of interest (AOI) belongs to the $Grassland$ class or not. The dataset contains Sentinel-2 reflectance values from all spectral bands across multiple time points, without focusing on $B1$, $B9$ and $B10$ at all since they are primarily designed for coastal and atmospheric applications (Ali and Johnson (2022)).

The data collection records temporal sequences of spectral values that enable dynamic analysis, which is needed for the aforementioned task, land cover classification using the spectral properties of various land surfaces.

Band	Function
B1	Coastal Aerosol
B2	Blue
B3	Green
B4	Red
B5	Red-edge
B6	Red-edge
B7	Red-edge
B8	NIR
B8a	Red-edge
B9	Water vapour
B10	SWIR
B11	SWIR
B12	SWIR

Table 1. Sentinel-2's 13 spectral bands

3.2 Data Preparation

As previously mentioned, the dataset for each sensor contain over 35,000 records in total. However, only approximately 4,000 records from these datasets have a $parcel\_id$ that matches an entry in the parcel-level crop label dataset. This subset was derived by merging the Sentinel-1 and Sentinel-2 datasets with the parcel-level annotations, ensuring that only records with overlapping $parcel\_ids$ are included. This overlap enables a comparative analysis of models trained independently on each satellite dataset, offering valuable insights into the relative effectiveness of radar and optical data for classification tasks.

3.2.1 Handling Class Imbalance The dataset used in this study shows a significant class imbalance, with 3,513 records labeled as $Grassland$ and only 533 records labeled as $Non-Grassland$. As a result of this class imbalance, ML models can become biased towards the majority class and perform poorly on the minority class which can make the training challenging (Chakraborty et al. (2021); Wang et al. (2023)). In order to solve this problem and make sure that the minority class is not underrepresented during training, we computed the inverse frequency of each class in the training data to establish determine the proper class weights. The class weights that are determined after calculating the class counts, are then allocated to individual samples, by using the PyTorch (Paszke et al. (2019)) module, $WeightedRandomSampler$. This sampler guarantees that each class contributes proportionately during the training process. This method helps to reduce the impact of class imbalance and encourages more balanced model learning.

4. Methodology

4.1 Data Preprocessing

The datasets used in this study, Sentinel-1 and Sentinel-2, were pre-aligned geographically and linked to specific parcels to ensure consistency in spatial representation. No missing values were observed in either dataset, eliminating the need for imputation or other handling techniques. Input features were normalized using the $StandardScaler$ method from $scickit-learn$ (Pedregosa et al. (2011)) to standardize their scale and improve model convergence during training. The datasets were split into training, validation and test sets with an $80:10:10$ ratio. Additionally, a 5-fold cross-validation procedure was applied to evaluate model robustness across varying subsets of data.

4.2 Feature Selection and Input Details

For the Sentinel-1 dataset, the input features comprised 28 values, representing VV and VH polarizations collected monthly between April and October 2017. These values were treated independently without temporal aggregation. Similarly, Sentinel-2 dataset included 290 input features derived from 10 spectral bands, also collected independently over the same period. No additional feature engineering or transformations, such as the computation of vegetation indices, were applied.

4.3 Model Architectures and Configurations

To evaluate the effectiveness of various machine learning algorithms for grassland classification, we assessed Neural Networks (NN), Support Vector Machines (SVM), Random Forest (RF), Logistic Regression (LogReg), and XGBoost (XGB), fine-tuning each model's hyperparameters for optimal performance. The **Neural Network** models were designed with distinct architectures based on the dataset: for Sentinel-1 the network had 28 input features with three layers comprising 32, 16, and 8 neurons, whereas for Sentinel-2 it utilized 290 input features with layers of 512, 128, and 16 neurons. Both architectures were trained using a learning rate of 0.00001, a dropout rate of 0.2, and the $CrossEntropyLoss$ function, running for 500 epochs while selecting the best-performing parameters based on MCC. The **SVM** model employed an RBF kernel with a regularization parameter $C = 1.0$ and a balanced class weight to address class imbalance. The **Random Forest** classifier as configured with 100 trees, a maximum depth of 10, and a balanced class weight. For **Logistic Regression**, we used L2 regularization, the LBFGS solver, and a maximum of 1000 iterations, with class weights adjusted o handle imbalance. Lastly, the XGBoost model was trained with 100 boosting rounds, a maximum tree depth of 6, and a learning rate of 0.1, incorporating the $scale\_pos\_weight$ parameter to mitigate class imbalance. These configurations ensured that each model was optimized for the classification task, maximizing performance across both Sentinel-1 and Sentinel-2 datasets.

4.4 Evaluation Metrics

Neural network models were trained using the RAdam optimizer, chosen for its ability to handle adaptive learning rates while ensuring stable convergence. In preliminary experiments, RAdam outperformed Adam in terms of both convergence speed and final classification performance. Cross-entropy loss was used as the objective function for the classification task. The training was conducted for 500 epochs with a batch size of 64, and the best-performing model parameters were saved based on the highest MCC achieved during the evaluation of the model.

To compare the performance of models trained on Sentinel-1 and Sentinel-2 data, we analyzed accuracy, loss, and MCC for each algorithm. Metrics were computed independently for both datasets to assess the effectiveness of radar versus optical data in grassland classification tasks. Additionally, we compared results statistically by observing trends in MCC values during training and testing across folds.

All experiments were conducted on a workstation with the following specifications: a 12th Gen Intel® Core™ i9-12900 CPU, 64 GB of RAM, and an NVIDIA RTX A2000 12 GB GPU. The training pipeline was implemented using PyTorch for neural networks and Scikit-learn for traditional machine learning algorithms.

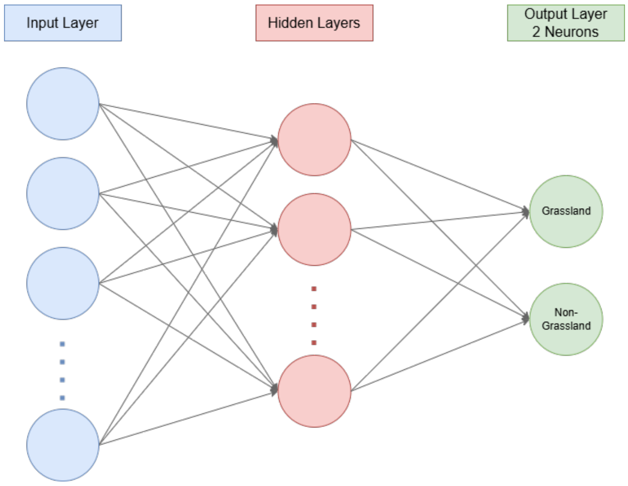

Figure 2. Neural Network architecture. NN architecture used for Sentinel-1 data: $[28, 32, 16, 8, 2]$. NN architecture used for Sentinel-2 data: $[290, 512, 256, 16, 2]$

5. Results

The table of the results (see table 2) shows the performance of various machine learning models when applied to Sentinel-1 and Sentinel-2 datasets, as determined by two key metrics: accuracy and the Matthews Correlation Coefficient (MCC).

For the **Sentinel-1** dataset, the Random Forest model demonstrated the highest accuracy (0.9172 ± 0.0086) and MCC (0.5888 $\pm$ 0.0500), indicating that it was the most reliable and accurate model when classifying grasslands using radar data. While the neural network achieved a high accuracy of 0.9040 ± 0.0192, its MCC score (0.4704 ± 0.2388) was notable lower compared to the Random Forest and SVM models, indicating less consistent performance in handling imbalanced data. Logistic Regression had the lowest accuracy and MCC, suggesting its limitations for more complex classification tasks in this dataset.

For the **Sentinel-2** dataset, which consists of optical imagery, the neural network outperformed other models with the highest accuracy highest accuracy (0.9649 ± 0.0082) and MCC (0.8381 $\pm$ 0.0396). Given that the Sentinel-2 dataset we used has 290 input features, this result demonstrates how well neural network can handle high-dimensional data. With accuracies of 0.9632 $\pm$ 0.0046 and 0.9622 ± 0.0030, and MCC scores slightly lower than the neural network's, other models like Random Forest and SVM, respectively, also showed strong performance.

Overall, these results show how dataset characteristics have a big impact on the performance of the model. While the Random Forest model performed well with Sentinel-1 radar data, offering a balanced performance in terms of accuracy and MCC, the neural network demonstrated outstanding performance with the high-dimensional Sentinel-2 data. These results highlight how crucial it is to select machine learning models that are relevant to the data's characteristics.

6. Discussion

6.1 Summary of Key Findings

The results of the study show notable differences in the performance of the machine learning models when they are used on

Dataset	Metric	Model				
		Neural Network	Logistic Regression	SVM	Random Forest	XGB
Sentinel-1	**Accuracy**	0.904 ± 0.019	0.861 ± 0.006	0.887 ± 0.007	$\mathbf{0.917 \pm 0.009}$	0.908 ± 0.011
	MCC	0.470 ± 0.239	0.533 ± 0.019	0.569 ± 0.027	$\mathbf{0.589 \pm 0.050}$	0.589 ± 0.032
Sentinel-2	**Accuracy**	$\mathbf{0.965 \pm 0.008}$	0.921 ± 0.009	0.962 ± 0.003	0.963 ± 0.005	0.961 ± 0.006
	MCC	$\mathbf{0.838 \pm 0.040}$	0.690 ± 0.020	0.826 ± 0.015	0.831 ± 0.023	0.834 ± 0.025

Table 2. Performance metrics (Accuracy, MCC, and Loss) of various machine learning models applied on Sentinel-1 and Sentinel-2 datasets. Bold values indicate the best performance in terms of Accuracy and MCC within each dataset.

Sentinel-1 and Sentinel-2 datasets for grassland classification. For the Sentinel-1 dataset, the Random Forest (RF) model outperformed all the other models, obtaining the highest accuracy and MCC scores. The neural network on the other hand, outperformed the other models in terms of accuracy and MCC score when tested on the Sentinel-2 data. Regarding the aforementioned results, we can see how important it is to select models based on the dataset's characteristics in order to get the best performance out of it.

6.2 Interpretation of Results

The good performance of Random Forest on the Sentinel-1 dataset may be attributed to its ensemble-based architecture, which allows the model to effectively handle moderate dimensional radar data. Its architecture also provides robustness against noise and variability which it can be seen in Synthetic Aperture Radar (SAR) images. As mentioned by Alharbi (2024), RF can efficiently handle the speckle noise and variability due to environmental factors that SAR data contain by combining multiple decision trees to improve generalization and reduce overfitting. In addition, dual-polarization (VV and VH) features in Sentinel-1 data provide valuable information about surface properties. The performance advantage that RF had, was probably influenced by its ability to handle non-linear relationships and a wide range of input features. As also mentioned in the literature, Belgiu and Drăguţ (2016) show in their study Random Forest's ensemble learning nature as benefit for radar datasets, while Pelletier et al. (2019) have demonstrated the algorithm's ability to use both polarization and texture for land cover classification.

On the other hand, the neural network's outperforming performance on the Sentinel-2 dataset is probably because of its capacity to handle multi-spectral, high dimensional input. Sentinel-2 data provide rich information across several spectral bands, and the neural network's architecture allows it to learn complex relationships between these data points. The neural network can recognize complex patterns in the data that other models would overlook thanks to its capacity to process large-scale input features and advanced optimization techniques like $RAdam$. This capability is particularly significant when working with high-dimensional optical imagery, where subtle spectral variations can be crucial for accurate classification. This finding is also supported by the literature, which shows that neural networks are effective at modeling non-linear interactions in hyper-spectral and multi-spectral remote sensing data (Nalepa et al. (2019)), while Zhu et al. (2017) discuss how neural networks can be used to exploit spectral information for vegetation classification.

6.3 Limitations and Future Work

The limitations of this study should be addressed in future research. First, because the analysis was limited to a specific region, the results may not be as applicable to other places with different meteorological or environmental characteristics. Second, although we handled the class imbalance using a $WeightedSampler$, a larger dataset could provide more balanced class representation and improve model performance further.

Future research can investigate the integration of Sentinel-1 and Sentinel-2 data in order to take advantage of the complementary information that each dataset provides. This approach could enhance the classification accuracy by providing a more complete representation of grassland features. Performance could also be further improved by exploring more advanced machine learning architectures, such as transformer-based models or ensemble methods that combine machine learning algorithms with deep learning neural networks. Finally, testing these algorithms on many datasets from different regions, can also make the results more reliable, opening the door for grassland monitoring systems that are more broadly applicable.

7. Conclusion

Using the sentinel-1 and Sentinel-2 datasets, this study assessed how well different Machine Learning models performed in grassland classification. The study's findings showed that the type of remote sensing data had a big impact on model performance. Random Forest (RF) excelled on Sentinel-1 data, likely due to its ability to handle noise and incorporate polarization and texture features that are inherent in Synthetic Aperture Radar (SAR) images. The neural network's ability to model complex relationships across multiple spectral bands on the other hand, achieved the best results on Sentinel-2 data.

These results have significant implications for grassland monitoring in remote sensing applications. The robust performance of RF on Sentinel-1 data shows its reliability for radar-based classification tasks, especially in areas with persistent cloud cover where optical imagery is less effective. Conversely, the neural network's success with Sentinel-2 data highlights how well suited is is for high-dimensional spectral analysis, making it an effective tool for in-depth vegetation evaluation.

Despite these positive outcomes of the study, certain limitations should be acknowledged. While class imbalance in the datasets was mitigated using a $WeightSampler$, increasing the volume of training data could further improve model performance. Moreover, this study concentrated on the separate analysis of Sentinel-1 and Sentinel-2 data, while further research could investigate the integration of both datasets to capitalize on their complementary strengths. Exploring more advanced machine learning architectures could also enhance classification accuracy.

Overall, this study provides insightful information about the applicability of different machine learning models for grassland classification using remote sensing data. By identifying the strengths of RF for radar-based data and neural networks for multi-spectral imagery, these findings aid in the development of more effective, data-driven approaches for environmental monitoring and land cover classification.

8. Acknowledgments

This work was partially supported by the European Union's HORIZON Research and Innovation Programme by the AI-OBSERVER project funded from the European Union's Horizon Europe Framework Programme HORIZON WIDERA-2021-ACCESS-03 (Twinning) under the Grant Agreement No 101079468, and the 'EXCELSIOR': ERATOSTHENES: Excellence Research Centre for Earth Surveillance and Space-Based Monitoring of the Environment H2020 Widespread Teaming project (www.excelsior2020.eu). The 'EXCELSIOR' project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510, from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development and the Cyprus University of Technology.

References

Alharbi, A. A., 2024. Classification Performance Analysis of Decision Tree-Based Algorithms with Noisy Class Variable. *Discrete Dynamics in Nature and Society*, 2024(1), 6671395.

Ali, K., Johnson, B. A., 2022. Land-use and land-cover classification in semi-arid areas from medium-resolution remote-sensing imagery: A deep learning approach. *Sensors*, 22(22), 8750.

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: A review of applications and future directions. *ISPRS journal of photogrammetry and remote sensing*, 114, 24–31.

Bengtsson, J., Bullock, J., Egoh, B., Everson, C., Everson, T., O'connor, T., O'farrell, P., Smith, H., Lindborg, R., 2019. Grasslands—more important for ecosystem services than you might think. *Ecosphere*, 10(2), e02582.

Boval, M., Dixon, R., 2012. The importance of grasslands for animal production and other functions: a review on management and methodological progress in the tropics. *Animal*, 6(5), 748–762.

Carlier, L., Rotar, I., Vlahova, M., Vidican, R., 2009. Importance and functions of grasslands. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, 37(1), 25–30.

Chakraborty, J., Majumder, S., Menzies, T., 2021. Bias in machine learning software: Why? how? what to do? *Proceedings of the 29th ACM joint meeting on european software engineering conference and symposium on the foundations of software engineering*, 429–440.

Choumos, G., Koukos, A., Sitokonstantinou, V., Kontoes, C., 2022. Towards space-to-ground data availability for agriculture monitoring. *2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*, 1–5.

d'Andrimont, R., Lemoine, G., Van der Velde, M., 2018. Targeted grassland monitoring at parcel level using sentinels, street-level images and field observations. *Remote Sensing*, 10(8), 1300.

Guo, X., Wilmshurst, J. F., McCanny, S. J., Fargey, P. J., Richard, P. J. H., 2004. Measuring Spatial and Vertical Heterogeneity of Grasslands Using Remote Sensing Techniques. *Journal of Environmental Informatics*, 3, 24-32. https://api.semanticscholar.org/CorpusID:111223162.

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.

Jutz, S., Milagro-Perez, M., 2020. Copernicus: the european earth observation programme. *Revista de Teledetección*, V–XI.

Nalepa, J., Tulczyjew, L., Myller, M., Kawulok, M., 2019. Segmenting hyperspectral images using spectral-spatial convolutional neural networks with training-time data augmentation. *arXiv preprint arXiv:1907.11935*.

Ojima, D. S., Dirks, B. O., Glenn, E. P., Owensby, C. E., Scurlock, J. O., 1993. Assessment of C budget for grasslands and drylands of the world. *Water, Air, and Soil Pollution*, 70, 95–109.

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

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. et al., 2011. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825–2830.

Pelletier, C., Webb, G. I., Petitjean, F., 2019. Temporal convolutional neural network for the classification of satellite image time series. *Remote Sensing*, 11(5), 523.

Wang, R., Chaudhari, P., Davatzikos, C., 2023. Bias in machine learning models can be significantly mitigated by careful training: Evidence from neuroimaging studies. *Proceedings of the National Academy of Sciences*, 120(6), e2211613120.

Weeks, E. S., Ausseil, A.-G. E., Shepherd, J. D., Dymond, J. R., 2013. Remote sensing methods to detect land-use/cover changes in New Zealand's 'indigenous' grasslands. *New Zealand Geographer*, 69, 1-13. https://api.semanticscholar.org/CorpusID:128710565.

Yong, Z., 2003. Most Recent Progress of International Research on Remote Sensing of Grassland Resources. *Progress in geography*. https://api.semanticscholar.org/CorpusID:133119279.

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 geoscience and remote sensing magazine*, 5(4), 8–36.