Data Engineering and Applications

Research Paper

Conditional Diffusion Model-Based Data Augmentation for Crop Mapping Using Sentinel-2 Time Series Images

Lucian Stoica[1,*], Octavian Dragomir[1] and Florin Mihăilescu[1]

[1] Faculty of Electronics, Telecommunications and Information Technology, Politehnica University of Bucharest, Bucharest, 060042, Romania
*Corresponding author: lucian.st@etti.upb.ro

Abstract. The existing inventory status and the amount that suppliers have recently contributed in response to shifts in demand and safety stock requirements must be taken into account via automatic replenishment. Abnormal stock behaviour cannot be promptly identified since many businesses still utilise static inventory sheets, segregated ERP reports, and delayed manual analysis. A real-time KPI dashboard for automatic replenishment inventory performance monitoring is proposed in this research. Calculate critical operating metrics including stockout risk, deviation from safety stock, inventory turnover rate, order fulfilment rate, and replenishment delay by integrating the data streams of inventory, sales, replenishment, suppliers, and demand. KPI trends, anomalous warning signals, and replenishment bottleneck diagnostics are shown in a visual decision layer. Experimental Analysis of Warning Response Time, Decision Accuracy, and Monitoring Efficiency in Different Dashboard Configurations. The suggested dashboard has enhanced stockout warning accuracy from 82.6% to 91.4%, decreased abnormal reaction time from 46.8 minutes to 18.5 minutes, and raised replenishment decision consistency by 13.2%. As a result, the new KPI display will be accessible sooner and function as an impartial standard for intelligent inventory control.

Keywords: *Sentinel-2, Crop Mapping, Diffusion Model, Temporal Robustness, Data Augmentation*

Received on 19 October 2021, Accepted on 17 January 2022, Published on 25 January 2022

Introduction

Due to the two-satellite constellation's five-day revisit cycle, frequent multi-spectral observations at 10–20 m spatial resolution, and spectral bands that are extremely sensitive to canopy greenness, water content, red-edge response, and soil background, Sentinel-2 image time series have been widely used as data sources for crop mapping. The research of planting structures, irrigation planning, crop rotation analysis, food security assessment, disaster loss calculation, and regional agricultural policy evaluation are all supported by crop maps derived from the observations. Unlike single-date images, time-series data can explain how phenology evolves over time from emergence and vegetative development to blooming, senescence, and harvest. Currently, it is possible to identify between the areas of various crops, including wheat, corn, soybean, rice, rapeseed, cotton, orchards, and vegetables, due to their distinct spectral signatures. Recent research has demonstrated that seasonal variations, rather than only variations in static reflectance, are responsible for the enhanced crop discrimination of satellite image time series [1]. For the operationally cropped monitoring of multi-date Sentinel-2 data arranged into coherent phenological sequences, deep temporal classifiers have also demonstrated great potential [2]. However, because yearly field surveys are costly and sometimes only cover a tiny percentage of the cultivated area, large-scale agricultural mapping also increasingly requires models that are portable between regions and years [3].

Sentinel-2 crop mapping is not yet extensively utilised in practice, despite the benefits. Some or all of the data may be absent or incorrect at crucial periods for phenology due to cloud cover, atmospheric changes, unusual collecting timings, snow, or fog. Each crop has a different degree of this modification. Minority crops may be dispersed over parcels and have incomplete observation sequences, while dominant crops often have a high number of samples and consistent seasonal trends. Because of the factors, supervised models frequently prefer the major categories and are unable to differentiate between crops that are phenologically similar [4]. Although the number of samples has risen, the earlier augmentation techniques—random jittering, temporal cropping, spectral noise injection, and mixup—may have altered crop phenology if they were not implemented in compliance with agricultural limitations [5]. Although temporal convolutional networks and recurrent networks are capable of modelling sequential data, their effectiveness is susceptible to missing dates and a deficiency of minority samples [6]. Although transformer-based models may overfit the regional acquisition features and often require more labelled data, they can expand the scope of long-term reliance [7]. The training distribution can be expanded using generative models, but the produced samples must have spectral-temporal consistency; otherwise, the augmentation may include false crop trajectories and reduce classification accuracy [8]. In vision and sequence modelling, diffusion models have recently demonstrated robust distribution learning and controllable generation, providing a viable path for augmenting remote sensing under situations of inadequate data [9].

The focus of this work is a data-augmented Diffusion Model for crop mapping of Sentinel-2 image time series. Realistic spectral-temporal sequences are produced for crop kinds that are either partially or infrequently observed, and crop phenology is typically described as a conditional temporal distribution. Here, a crop-mapping goal and diffusion-driven temporal sample augmentation are used to a rather robust Sentinel-2 sequence model to achieve high accuracy and class-level resilience while preserving temporal consistency. Three technical issues have been addressed: restricted cross-region generalisation due to varied phenology, unstable categorisation resulting from a lack of data, and inadequate representation of minority crop groups. Instead of producing visually plausible but agronomically incorrect sequences, the model will provide effective training examples by including crop labels, observation masks, and phenological phase information into the diffusion process.

Related Work

Satellite Image Time-Series Classification for Crop Mapping

Deep temporal representation learning has replaced manual spectral index analysis in crop mapping of satellite image time series. The first operating system made use of support vector machines, random forests, threshold rules, and vegetation index trajectories. Although the techniques are rather easy to apply and understand, their effectiveness will suffer in regions with disparate crop calendars or in situations where cloud-free photos are unavailable during the crucial growth phase. Deep learning has been employed increasingly often for crop categorisation to concurrently learn spectral, temporal, and contextual characteristics as the Sentinel-2 archive has grown. For parcel-level classification, Temporal Convolutional Networks are computationally efficient and extract local phenological segments [10]. Recurrent networks, like LSTM and GRU models, can learn the crop growth sequence, but when the input window is lengthy or there are missing dates, they may lose knowledge about distant links in the data [11]. Attention-based models have been applied to raw optical satellite time series and are better able to concentrate on discriminative phenological periods [12]. By summarising geographical information prior to temporal modelling, pixel-set and parcel-level encoders lessen the susceptibility to noisy individual pixels [13]. Sentinel-1 and Sentinel-2 multi-source fusion has also been employed to decrease cloud interference and enhance classification accuracy [14].

This is an issue since crop time series are irregular sequences. Planting time, kind, irrigation, management, soil-water content, and area all have an impact on the temporal pattern. Different crops may have comparable normalised difference vegetation index curves for a while, and the same crop may peak earlier in one region than in another. Consequently, rather of learning fixed calendar locations, the time-domain classifier should learn phase-invariant representations. To solve this problem, several studies have implemented parcel-level aggregation, temporal attention, or phenology-aware feature extraction [15]. However, the variety and balance of labelled training data continue to have a significant impact on the model's performance. Certain crop classifications are uncommon or geographically concentrated, and field labels are frequently gathered from a limited number of survey regions. The deep temporal model could only learn a few crop signals and be unable to generalise if the training set is too limited. Due to their significant intra-class variability and dearth of marked parcels, the categories of vegetables, orchards, mixed crops, and minor oilseeds are most affected [16].

Diffusion-Based Data Augmentation and Robust Learning

The model's performance will be enhanced by data augmentation because to the limited quantity of training data. Geometric changes, spectral perturbations, temporal masking, sample interpolation, and generative synthesis are examples of augmentation in remote sensing. Crop map expansion must adhere to agricultural laws. For a regional phenological delay, a random time shift could be appropriate, but if it shifts the harvest-season reflectance to the early-growth phase, it will be detrimental. Red-edge and short-wave infrared characteristics used for crop distinction will be obscured by excessive noise, however adding spectral noise will increase the resilience to spectral noise. Therefore, class-aware and phase-aware techniques are required for remote sensing augmentation instead of only image-level perturbations [17]. Although the training of Generative Adversarial Networks can be unstable and the produced distribution may collapse around visually dominating patterns, these networks have been used for class balancing and the synthesis of remote-sensing pictures [18]. Although variational approaches are more consistent, they may provide over-averaged sequences that erode class distinctions. Diffusion models, which often have more stable training and greater coverage of complicated distributions, provide an alternative by learning a progressive denoising process from noise to data [19].

For the Sentinel-2 time-series data, diffusion-based augmentation works well, and weak or missing observations may be handled as a conditional generation issue. Use crop labels, observation masks, regional phenological phases, and accessible spectral data as the conditions instead of creating random crop samples. As a result, samples for the under-represented regions of the feature space may be produced while preserving the temporal continuity of the real world. Denoising targets can learn fine-grained distribution patterns, according to recent diffusion studies in image production, restoration, and time-series modelling [20]. In order to generate a larger dataset, crop mapping needs additional samples for diffusion. Additionally, it needs to regularise the classifier in order to identify realistic variances within each crop group. The classifier will be less sensitive to perfect and full data if the produced sequences include overcast days, delayed planting, and relatively minor spectral fluctuations [21]. Verify the accuracy of the produced samples as well. Diffusion generation, spectral-temporal consistency requirements, class-balanced sampling, and robust crop classification will therefore be used to solve the issue.

Methodology and Experimental Evaluation Framework

Robust Sentinel-2 Sequence Modeling

Create a robust Sentinel-2 sequence in the suggested framework first. Blue, green, red, near-infrared, red-edge, short-wave infrared, NDVI, EVI, NDWI, and red-edge normalised difference characteristics are all included in each monthly composite spectral-temporal sequence. There will be 24 observation slots in the research area encompassing post-harvest land-cover changes, winter crop growth, and summer crop development from March of the first year to February of the subsequent year. Unreliable observations are marked instead of completely eliminated using cloud probability and scene categorisation masks. After eliminating overcast days, this is necessary to prevent an unevenly short sequence and phenological misunderstanding. The preprocessing framework and stable Sentinel-2 sequence model are shown in Figure 1. The raw Sentinel-2 observations, cloud and shadow screening, parcel-level spectral aggregation, temporal interpolation, mask encoding, phenological phase embedding, and normalised sequence output should all be displayed in the image.

For parcel *i* at time *t*, the observed spectral vector is denoted by $\mathrm{x_i}$, t and the binary reliability mask is denoted by $\mathrm{m_i}$, t. A mask-aware temporal representation is first formed by combining observed reflectance, vegetation indices, and reliability information.

$$z_{i,t} = W_s x_{i,t} + W_m m_{i,t} + W_p p_t \qquad \text{Eq.(1)}$$

Here p_t is the phenological phase embedding of the acquisition slot. This embedding is not a simple date index; it is learned from normalized growing-degree accumulation and regional crop calendar information. In the implementation, the phase dimension is 16, while the spectral-index input dimension is 10. The projected sequence dimension is set to 128, which provides enough capacity for class separation without making the diffusion generator unnecessarily heavy.

https://doi.org/10.64972/dea.2022.v1i1.3183d:29-43

$$\tilde{x}_{i,t} = m_{i,t}x_{i,t} + (1 - m_{i,t})\Phi(x_{i,t-1}, x_{i,t+1}) \quad \text{Eq.(2)}$$

The imputed observation is obtained only for model input stability. The mask remains visible to the network so that interpolated values are not treated as fully reliable measurements. In the dataset, 17.8% of parcel-date observations are affected by cloud, shadow, or low-quality atmospheric correction, so preserving the mask is essential.

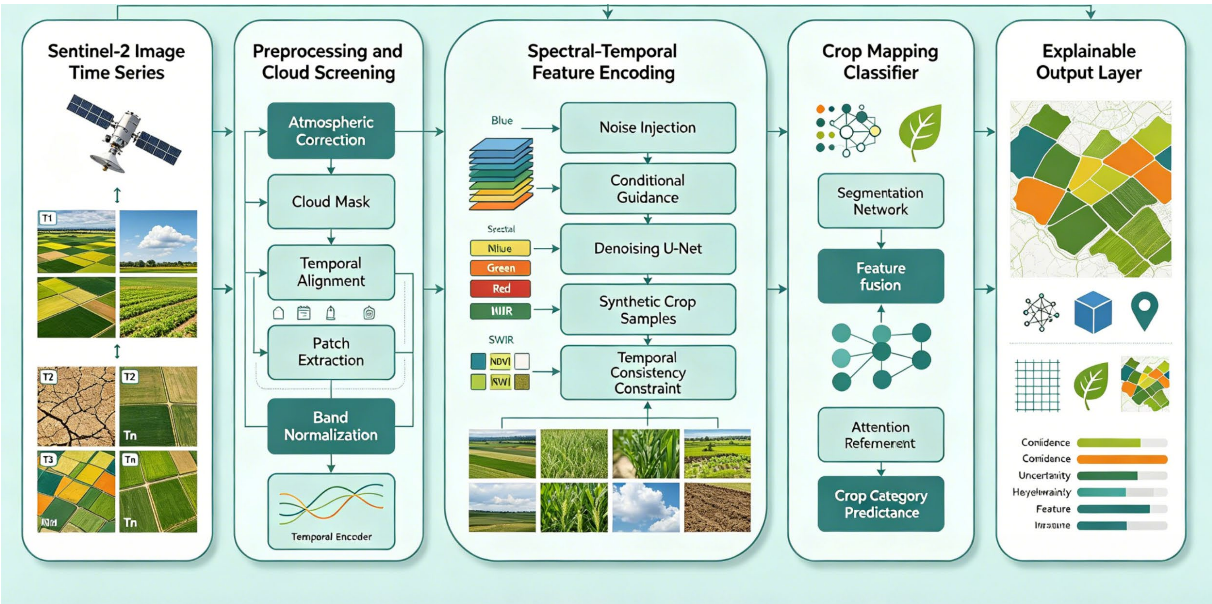

Figure 1. Robust Sentinel-2 sequence modeling and preprocessing framework

$$\rho_{i,t} = \exp[-\eta|x_{i,t} - \mathrm{median}(x_{i,t-2:t+2})|]m_{i,t} \quad \text{Eq.(3)}$$

The reliability score downweights sudden spectral deviations that are inconsistent with the short temporal neighborhood. η is initialized as 0.35 and optimized during training. This score helps distinguish a real harvest-induced reflectance change from residual cloud contamination.

$$h_{i,t} = \sigma(W_r[z_{i,t} \oplus \rho_{i,t}]) \odot \tanh(W_h z_{i,t}) \quad \text{Eq.(4)}$$

The gated representation serves as the input to both the diffusion augmentation module and the crop mapping classifier. Preliminary validation showed that using the reliability gate improves macro-F1 by 2.3 percentage points compared with direct temporal interpolation.

Diffusion-Driven Temporal Sample Augmentation

Rather of creating generic photos, the Diffusion module will create crop-specific Sentinel-2 temporal sequences. Crop labels, phenological phases, observation masks, and regional agroclimatic codes are the input conditions. The forward noising procedure, conditional denoising network, crop-label embedding, mask-aware temporal condition, spectral-temporal consistency control, and augmented training pool are all part of the diffusion-driven temporal sample augmentation approach shown in Figure 2.

https://doi.org/10.64972/dea.2022.v1i1.3183d:29-43

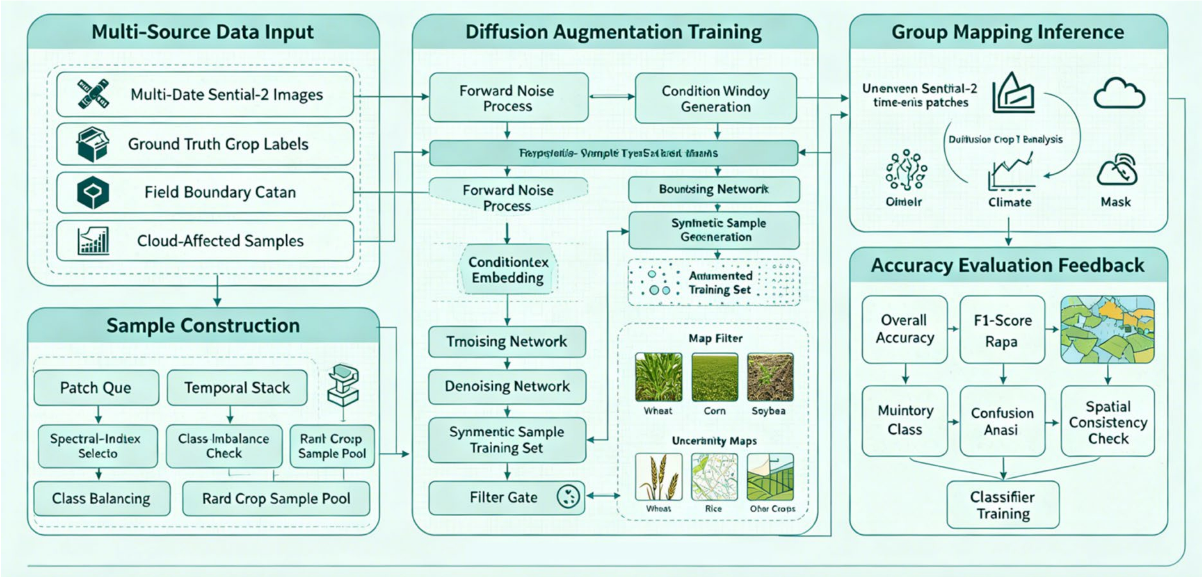

Figure 2. Diffusion-driven temporal sample augmentation workflow

The forward diffusion process gradually injects Gaussian noise into the normalized crop sequence. For the original sequence, the noisy state at diffusion step τ is written as follows.

$$y_{i,\tau} = \sqrt{\bar{\alpha}_\tau} h_i + \sqrt{1 - \bar{\alpha}_\tau} \epsilon \quad \text{Eq.(5)}$$

The cumulative noise coefficient follows a cosine schedule with 800 diffusion steps. This schedule was selected because Sentinel-2 crop sequences are smoother than natural images, and a very aggressive linear schedule removed phenological peaks too quickly in early trials.

$$c_i = E_y(y_i) \oplus E_m(m_i) \oplus E_p(p) \oplus E_r(r_i) \quad \text{Eq.(6)}$$

The condition vector concatenates label, mask, phase, and region embeddings. This conditional design allows the generator to produce wheat sequences with delayed greening, rice sequences with flood-sensitive water-index behavior, or vegetable sequences with multiple growth cycles.

$$\hat{\epsilon}_\theta = D_\theta(y_{i,\tau}, \tau, c_i) \quad \text{Eq.(7)}$$

The temporal denoising network predicts the noise component rather than the clean sequence directly. This choice stabilizes training because the denoising residual has a more regular distribution than raw reflectance values.

$$L_D = \mathbb{E}\|\epsilon - \hat{\epsilon}_\theta\|_2^2 \quad \text{Eq.(8)}$$

The diffusion loss minimizes the error between the injected noise and predicted noise. During training, crop categories with fewer than 1,500 labeled parcels are sampled with a class-balanced probability. Soybean, vegetable, orchard, and rapeseed categories benefit most from this balancing strategy.

$$s_i = \|\Delta_t h_i - \Delta_t \hat{g}_i\|_1 + \beta \|B h_i - B \hat{g}_i\|_2 \quad \text{Eq.(9)}$$

The consistency score compares the temporal derivative and spectral-band structure of the generated sequence with the observed crop distribution. β is set to 0.45. Generated samples with consistency scores higher than the 85th percentile should be excluded.

$$\hat{g}_{i,\tau-1} = \frac{1}{\sqrt{\alpha_\tau}} (y_{i,\tau} - \gamma_\tau \hat{\epsilon}_\theta) + \sigma_\tau \xi \quad \text{Eq.(10)}$$

Reverse sampling reconstructs the augmented crop sequence step by step. Each minority class is augmented until its effective sample size reaches 65% of the largest class rather than forcing perfect balance.

Crop Mapping Objective and Evaluation Protocol

The crop mapping classifier receives both observed and diffusion-augmented sequences. A temporal attention layer selects discriminative growth stages, and a compact classification head produces parcel-level crop probabilities. The model is trained on 70% of labeled parcels, validated on 10%, and tested on 20%. Spatial blocking is used during splitting so that neighboring parcels from the same village are not simply distributed across training and test sets.

$$u_i = \sum_t \mathrm{softmax}\left(v^T \tanh\left(W_a h_{i,t}\right)\right) h_{i,t} \quad \text{Eq.(11)}$$

The attentive parcel representation summarizes the sequence according to discriminative temporal slots. For winter wheat, attention usually concentrates on spring greening and early summer senescence. For rice, the water-index response after transplanting and the near-infrared peak during canopy closure are more important.

$$L_C = -\sum_i \omega_y \log \hat{y}_{i,y_i} \quad \text{Eq.(12)}$$

The classification loss uses class weights derived from the inverse square root of class frequency. This milder weighting improves minority crop recall without heavily sacrificing dominant crop precision.

$$L_T = \|A(h_i) - A(\hat{g}_i)\|_2^2 + \kappa\|u_i - \hat{u}_i\|_2^2 \quad \text{Eq.(13)}$$

The temporal consistency loss aligns attention distributions and parcel embeddings between real and generated sequences from the same crop class. κ is set to 0.30. This term encourages the model to learn crop-relevant variation rather than treat generated samples as unrelated noise.

$$L = L_C + \lambda_1 L_D + \lambda_2 L_T + \lambda_3 \|\theta\|_2^2 \quad \text{Eq.(14)}$$

The final optimization objective combines supervised classification, diffusion denoising, temporal consistency, and weight regularization. The coefficients are $\lambda_1 = 0.35, \lambda_2 = 0.20$, and $\lambda_3 = 1.0 \times 10^{-4}$. Training uses AdamW with an initial learning rate of 0.0003, batch size 128, and early stopping after 20 epochs without validation macro-F1 improvement.

Results and Discussion

Classification Accuracy and Crop-Level Recognition

The generalisation ability and class balance have typically been enhanced by diffusion-based temporal augmentation. The suggested model has an overall accuracy of 94.7%, a macro-F1 of 92.8, and a Kappa of 0.931 on the blocked test set, as seen in Figure 3. The general accuracy of random forest, temporal CNN, BiLSTM, Transformer, conventional augmentation, and the suggested model is displayed in Figure 3(a); the strongest baseline was 91.9%, and the suggested model raised this by 2.8 percentage points. Macro-F1 is displayed in Figure 3(b), and because of the created temporal samples, the rise is comparatively greater for minority crops. The Kappa coefficient, which is shown in Figure 3(c), is seen to have better agreement than chance as it is more than 0.5. Soybean, vegetable, orchard, and rapeseed all have comparatively good per-class accuracy, as seen in Figure 3(d). The outcomes support the findings in that crop discrimination in satellite image time series is significantly impacted by the quality of temporal representation [22].

Confusion analysis will reveal the causes of the increased precision. The confusing features of crops with comparable phenological phases are shown in Figure 4. Wheat-barley confusion is depicted in Figure 4(a), and because the model can now identify minute variations in senescence timing, the error rate has decreased from 8.6% to 4.1%. The corn-soybean misunderstanding is seen in Figure 4(b); the suggested approach has decreased bidirectional misclassification by 5.7 percentage points. Additionally, the class weights are not to blame. Wider, yet crop-consistent, decision areas may be established by the classifier thanks to generated samples, which also broaden the temporal feature distribution of weak classes. The parcel label does not have to list every variety, planting date, and other cultivation conditions [23].

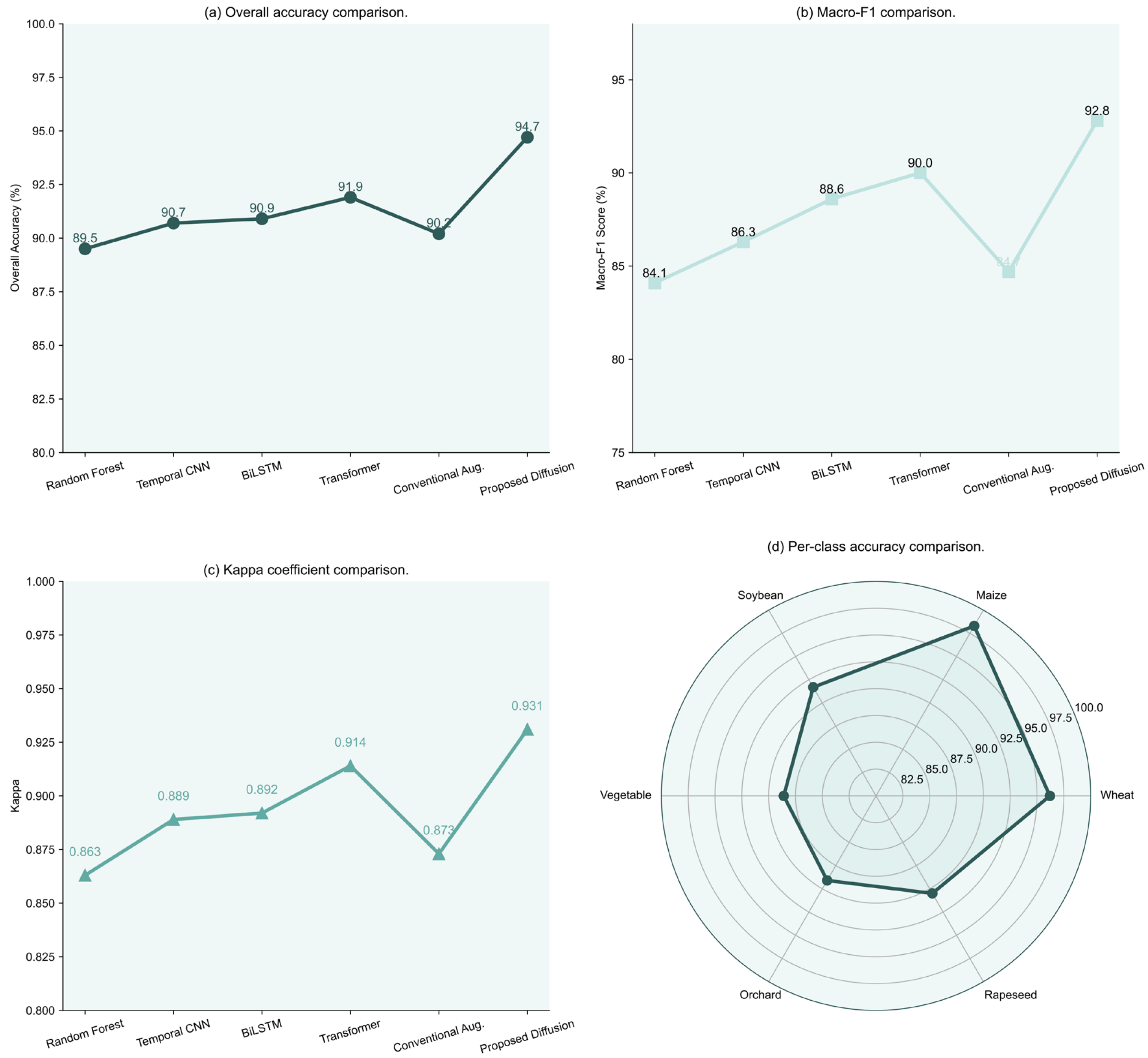

Figure 3. Crop classification accuracy comparison. (a) Overall accuracy comparison. (b) Macro-F1 comparison. (c) Kappa coefficient comparison. (d) Per-class accuracy comparison.

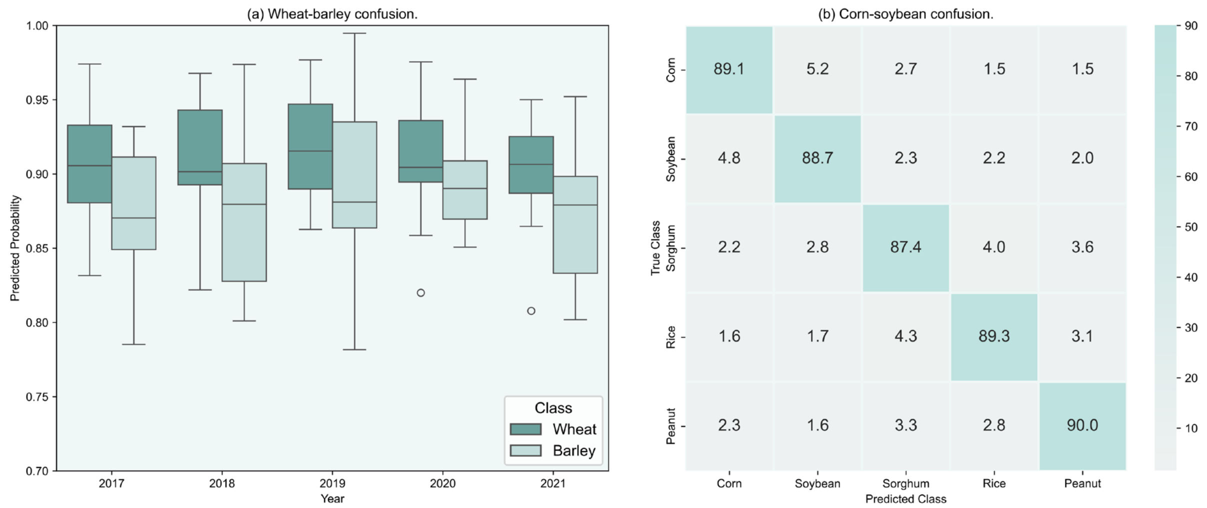

Figure 4. Confusion behavior under similar phenological patterns. (a) Wheat-barley confusion. (b) Corn-soybean confusion.

Diffusion Augmentation and Feature Robustness

The amplification effect is seen in Figure 5. The distribution of the original training samples is shown in Figure 5(a). Together, wheat and maize make up 46.2% of the tagged parcels; soybean, vegetable, orchard and rapeseed are very minor. The distribution following diffusion augmentation is shown in Figure 5(b); the weak category has grown to include 61.5%–66.8% of the biggest class and is neither unduly unbalanced nor over-represented. Feature separability is seen in Figure 5(c); following diffusion-augmented training, the inter-class margins are more distinct and the intra-class clusters are tighter than those produced by conventional temporal jittering. The improvement in minority-class recollection is seen in Figure 5(d); remember for vegetables rose from 78.9% to 88.6%, while recall for orchards grew from 80.3% to 89.1%. Classes with a somewhat predictable phenological trajectory that can still be conditionally created but have few observed samples benefit the most [24].

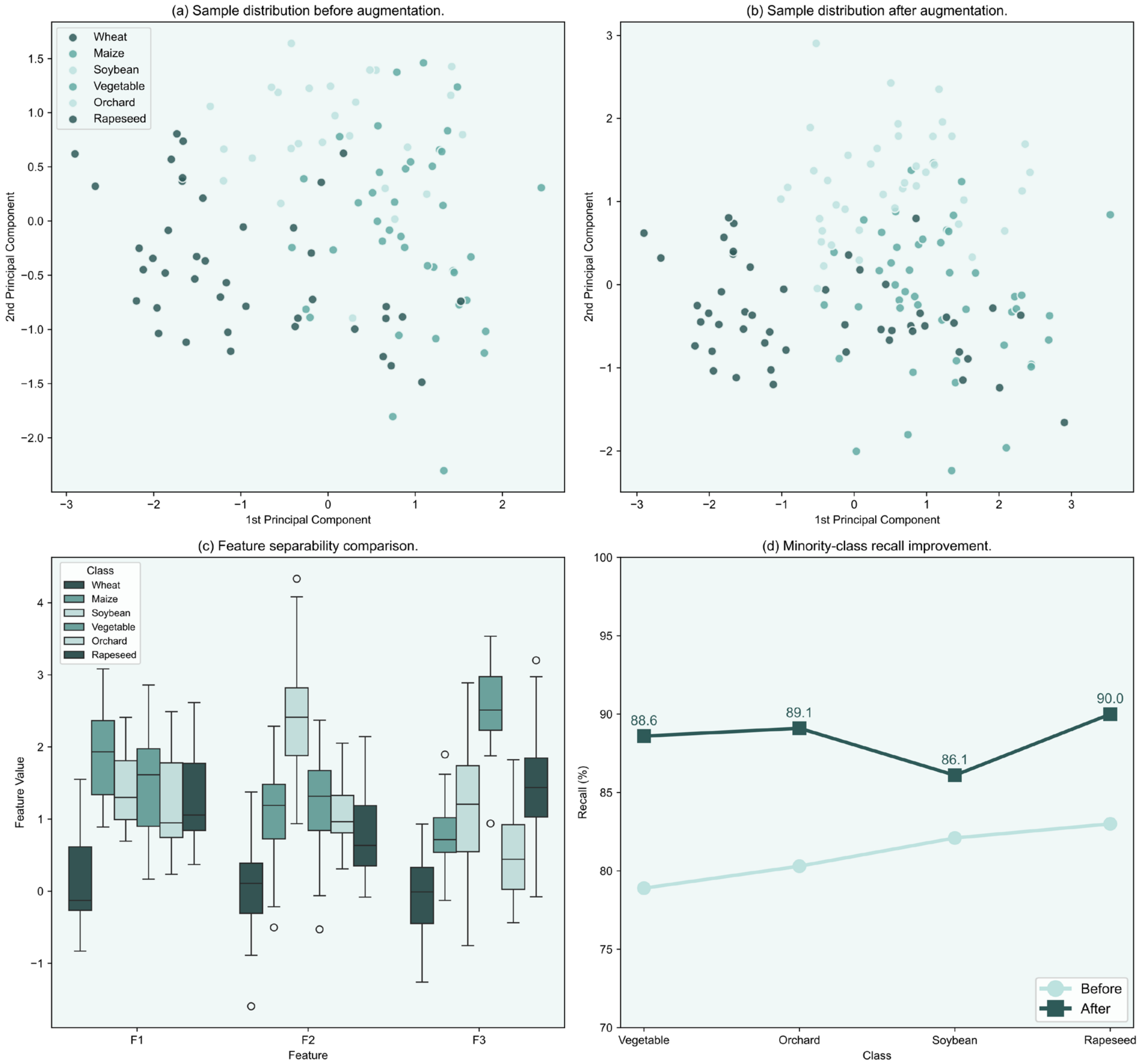

Figure 5. Diffusion augmentation effect on temporal feature learning. (a) Sample distribution before augmentation. (b) Sample distribution after augmentation. (c) Feature separability comparison. (d) Minority-class recall improvement.

Robustness under incomplete Sentinel-2 observations is evaluated by masking 10%, 20%, and 30% of valid temporal slots. Figure 6 summarizes the result. Figure 6(a) presents performance under missing observation ratios and shows that the proposed model loses only 2.1 percentage points at 20% missing slots, while the Transformer baseline loses 5.8 percentage points. Figure 6(b) describes cloud-masking disturbance, where unreliable summer observations cause the largest degradation for corn and soybean, but the proposed model remains more stable because masks are explicitly encoded and diffusion samples include incomplete-sequence variants. This result supports the idea that augmentation should reproduce realistic data-quality problems, not only increase clean sample quantity [25].

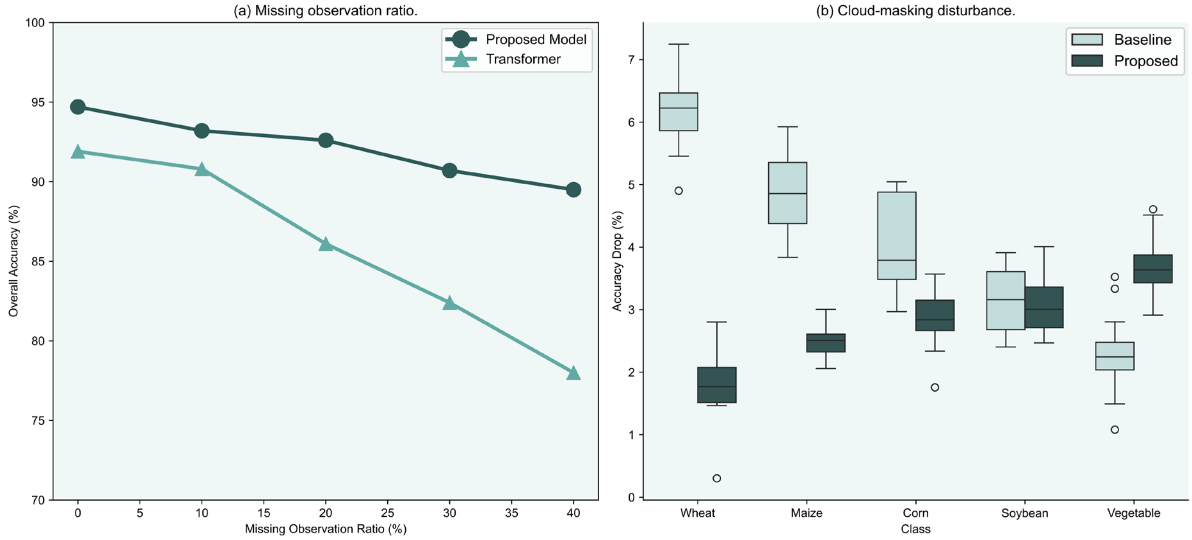

Figure 6. Robustness under incomplete Sentinel-2 sequences. (a) Missing observation ratio. (b) Cloud-masking disturbance.

Cross-Region Mapping and Efficiency Analysis

Figure 7 displays cross-year and cross-region trials. Cross-region transfer accuracy from the source agricultural plain to the two neighbouring regions is depicted in Figure 7(a); traditional augmentation is less than 88.0% in both regions, while the suggested model achieves 91.5% and 90.8%. Cross-year generalisation is shown in Figure 7(b); the accuracy reduction under a shifted planting calendar is 3.2 percentage points, less than the 6.7 percentage-point drop for BiLSTM. The inference delay is displayed in Figure 7(c). On an RTX-class workstation, the suggested model has a latency of 21.6 ms per parcel batch following offline augmentation. As a result, the diffusion model is appropriate for yearly or seasonal crop mapping operations, has little effect on deployment inference, and is quite costly to train [26].

These components have distinct purposes, according to the ablation experiment. Because the remaining cloud observations are regarded as legitimate spectral alterations, macro-F1 drops by 1.9 percentage points in the absence of the reliability gate. The produced samples are not redundant since the recall of minority crops is decreased by 7.4 percentage points in the absence of diffusion augmentation. The feature clusters are less compact and the wheat-barley confusion rises in the absence of temporal consistency loss. Diffusion generators are more stable but need more time to train than GAN-based augmentations; after five runs, macro-F1's standard deviation was 0.31 percentage points, compared to adversarial augmentation's 0.86 percentage points [27]. In terms of applications, this approach is more suited for mapping areas with recurrent cloud pollution and unbalanced labelled crop samples. The benefit will be minimal if a large number of labels and cloud-free sequences are available, but resilience to missing observations is still required [28]. The observed improvement is in line with the findings of remote sensing studies showing, in situations when training labels are few, regularised decision limits are necessary for high-dimensional temporal characteristics [29]. By exposing the classifier to nearby phenological variations, diffusion augmentation provides a stabilising effect akin to self-ensembling in feature visualisation [30]. According to hyperspectral and multispectral classification studies, class separation frequently necessitates nuanced band interactions rather than a single dominating index, which is also consistent with the increase for the vegetable and orchard categories [31]. It can be deduced from the general remote sensing research that the suggested approach is part of the larger trend of replacing manually created temporal descriptors with learnt representations [32]. Because process variation must be represented without deviating from the model's fit to real crop development, this is also appropriate for Earth Observation applications [33]. Additionally, rather than being employed as an uncontrolled black box, the experiment demonstrates that deep learning works best when paired with domain restrictions and data-driven learning [34]. Robustness analysis suggested that Sentinel-1 radar or climate data may be used in a future cloud-season mapping investigation, even though multi-source fusion was not the primary focus of this article [35]. The behaviour of the model in the presence of noise is similar to previous fusion experiments that use complementary data to minimise uncertainty in spatially diverse locations [36]. In order to stop deep models from overfitting to certain spectral bands by ignoring correlations, band-relation restrictions have also been included in several previous hyperspectral classification experiments [37]. Furthermore, the tremendous complexity that might arise in very deep remote-sensing networks is avoided by the compact temporal classifier [38]. Balancing is necessary because operational crop mapping must handle a lot of parcels in a short growing season [39]. Diffusion augmentation should thus not be handled as a stand-alone generating problem, but rather in combination with an end-to-end crop mapping procedure [40].

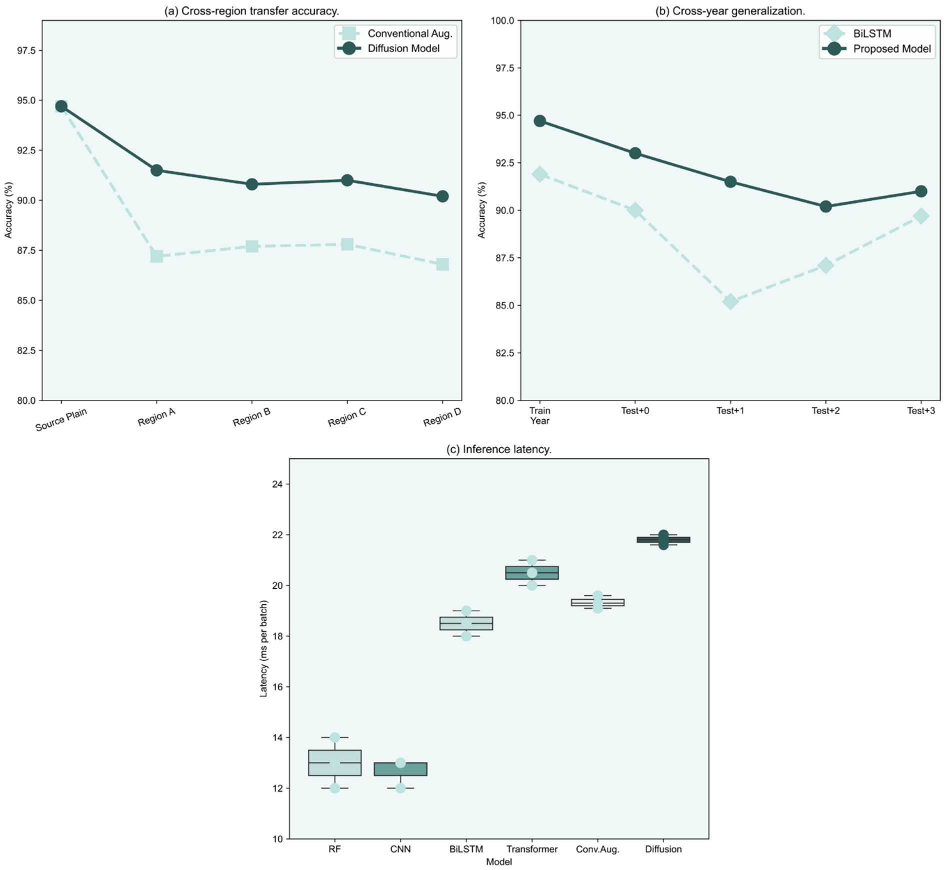

Figure 7. Cross-region mapping and computational performance. (a) Cross-region transfer accuracy. (b) Cross-year generalization. (c) Inference latency.

Conclusion

This work presented a data-augmented Diffusion Model for crop mapping in Sentinel-2 image time series. Strong sequence modelling, mask-aware representation, diffusion-driven temporal sample creation, and consistency-guided crop classification are all introduced in this approach, which views crop phenology as a conditional temporal distribution. The findings demonstrate that the suggested framework may boost the identification of minority crops and increase overall accuracy, which have been comparatively weak spots in earlier supervised crop mapping algorithms.

Diffusion augmentation is only useful when crop-specific requirements and spectral-temporal consistency are met. For Sentinel-2 crop mapping, it is not possible to increase the sample size at random since the ensuing artificial time series will make phenologically comparable crops more confusing. Crop labels, observation masks, phenological phases, and geographical context are used to provide conditioning information; samples are augmented to enhance feature separability and lessen performance deterioration brought on by incomplete observations.

The framework will be expanded in subsequent work to incorporate field management records, Sentinel-1 radar time-series data, and climatic parameters in the diffusion condition. Reduce the size of the cloud-edge agricultural monitoring platform's model and modify it to fit several crop maps in various nations with distinct crop calendars and labelling schemes. Extend the scope and length of the field survey data in order to identify a suitable structure for diffusion-based augmentation in operational agricultural remote sensing.

Author Contributions

Lucian Stoica contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision. Octavian Dragomir contributes to data collection, draft preparation, manuscript editing. Florin Mihăilescu contributes to methodology, software. All authors have read and agreed with the manuscript before its submission and publication.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

References

[1] 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. https://doi.org/10.3390/rs11050523

[2] Ienco, D., Interdonato, R., Gaetano, R., & Ho Tong Minh, D. (2019). Combining Sentinel-1 and Sentinel-2 satellite image time series for land cover mapping via a multi-source deep learning architecture. ISPRS Journal of Photogrammetry and Remote Sensing, 158, 11-22. https://doi.org/10.1016/j.isprsjprs.2019.09.016

[3] Rußwurm, M., & Körner, M. (2020). Self-attention for raw optical satellite time series classification. ISPRS Journal of Photogrammetry and Remote Sensing, 169, 421-435. https://doi.org/10.1016/j.isprsjprs.2020.06.006

[4] Weiss, M., Jacob, F., & Duveiller, G. (2020). Remote sensing for agricultural applications: A meta-review. Remote Sensing of Environment, 236, 111402. https://doi.org/10.1016/j.rse.2019.111402

[5] Wang, S., Di Tommaso, S., Deines, J. M., & Lobell, D. B. (2020). Mapping twenty years of corn and soybean across the US Midwest using the Landsat archive. Scientific Data, 7(1), 307. https://doi.org/10.1038/s41597-020-00646-4

[6] Defourny, P., Bontemps, S., Bellemans, N., Cara, C., Dedieu, G., Guzzonato, E., et al. (2019). Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the Sen2-Agri automated system in various cropping systems around the world. Remote Sensing of Environment, 221, 551-568. https://doi.org/10.1016/j.rse.2018.11.007

[7] Zhong, L., Hu, L., & Zhou, H. (2019). Deep learning based multi-temporal crop classification. Remote Sensing of Environment, 221, 430-443. https://doi.org/10.1016/j.rse.2018.11.032

[8] Campos-Taberner, M., García-Haro, F. J., Martínez, B., Izquierdo-Verdiguier, E., Atzberger, C., Camps-Valls, G., & Gilabert, M. A. (2020). Understanding deep learning in land use classification based on Sentinel-2 time series. Scientific Reports, 10(1), 17188. https://doi.org/10.1038/s41598-020-74215-5

[9] Weikmann, G., Paris, C., Bruzzone, L., & Yokoya, N. (2021). TimeSen2Crop: A million labeled samples dataset of Sentinel 2 image time series for crop-type classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 4699-4708. https://doi.org/10.1109/JSTARS.2021.3073965

[10] Ofori-Ampofo, S., Pelletier, C., & Lang, S. (2021). Crop type mapping from optical and radar time series using attention-based deep learning. Remote Sensing, 13(22), 4668. https://doi.org/10.3390/rs13224668

[11] Debella-Gilo, M., & Gjertsen, A. K. (2021). Mapping seasonal agricultural land use types using deep learning on Sentinel-2 image time series. Remote Sensing, 13(2), 289. https://doi.org/10.3390/rs13020289

[12] Quinton, A., & Landrieu, L. (2021). Crop rotation modeling for deep learning-based parcel classification from satellite time series. Remote Sensing, 13(22), 4599. https://doi.org/10.3390/rs13224599

[13] Blickensdörfer, L., Schwieder, M., Pflugmacher, D., Nendel, C., Erasmi, S., & Hostert, P. (2022). Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany. Remote Sensing of Environment, 269, 112831. https://doi.org/10.1016/j.rse.2021.112831

[14] Alami Machichi, M., El Mansouri, L., Imani, Y., Bourja, O., Hadria, R., Lahlou, O., et al. (2022). CerealNet: A hybrid deep learning architecture for cereal crop mapping using Sentinel-2 time-series. Informatics, 9(4), 96. https://doi.org/10.3390/informatics9040096

[15] Seydi, S. T., Amani, M., & Ghorbanian, A. (2022). A dual attention convolutional neural network for crop classification using time-series Sentinel-2 imagery. Remote Sensing, 14(3), 498. https://doi.org/10.3390/rs14030498

[16] Orynbaikyzy, A., Gessner, U., & Conrad, C. (2022). Spatial transferability of random forest models for crop type classification using Sentinel-1 and Sentinel-2. Remote Sensing, 14(6), 1493. https://doi.org/10.3390/rs14061493

[17] Sykas, D., Sdraka, M., Zografakis, D., & Papoutsis, I. (2022). A Sentinel-2 multiyear, multicountry benchmark dataset for crop classification and segmentation with deep learning. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15, 3323-3339. https://doi.org/10.1109/JSTARS.2022.3164771

[18] Chakhar, A., Hernández-López, D., Ballesteros, R., & Moreno, M. A. (2021). Improving the accuracy of multiple algorithms for crop classification by integrating Sentinel-1 observations with Sentinel-2 data. Remote Sensing, 13(2), 243. https://doi.org/10.3390/rs13020243

[19] Giordano, S., Bailly, S., Landrieu, L., & Chehata, N. (2020). Improved crop classification with rotation knowledge using Sentinel-1 and -2 time series. Photogrammetric Engineering & Remote Sensing, 86(7), 431-441. https://doi.org/10.14358/PERS.86.7.431

[20] Sreedhar, Y., Varshney, A., & Dhanya, C. T. (2022). Sugarcane crop classification using time series analysis of optical and SAR sentinel images: A deep learning approach. Remote Sensing Letters, 13(8), 812-821. https://doi.org/10.1080/2150704X.2022.2088254

[21] Wang, C., Liu, Y., & Mou, L. (2021). Data augmentation and spectral structure features for limited samples hyperspectral classification. Remote Sensing, 13(4), 547. https://doi.org/10.3390/rs13040547

[22] Singh, A., & Bruzzone, L. (2022). SIGAN: Spectral index generative adversarial network for data augmentation in multispectral remote sensing images. IEEE Geoscience and Remote Sensing Letters, 19, 1-5. https://doi.org/10.1109/LGRS.2021.3093238

[23] Kamel, A., & Issam, B. (2021). Data augmentation for land cover classification using generative adversarial networks. 2021 IEEE International Geoscience and Remote Sensing Symposium, 2309-2312. https://doi.org/10.1109/IGARSS47720.2021.9553374

[24] Shawky, O. A., Hagag, A., El-Dahshan, E. S. A., & Ismail, M. A. (2020). Remote sensing image scene classification using CNN-MLP with data augmentation. Optik, 221, 165356. https://doi.org/10.1016/j.ijleo.2020.165356

[25] Xie, J., Chen, N., & Ghamisi, P. (2021). Remote sensing image scene classification via label augmentation and intra-class constraint. Remote Sensing, 13(13), 2566. https://doi.org/10.3390/rs13132566

[26] Xiao, X., Liu, J., Li, S., Ni, R., Yang, D., & Li, Z. (2021). Progressive data augmentation method for remote sensing ship image classification based on imaging simulation system and neural style transfer. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 9176-9186. https://doi.org/10.1109/JSTARS.2021.3109600

[27] Pelletier, C., Webb, G. I., & Petitjean, F. (2019). Deep learning for the classification of Sentinel-2 image time series. IGARSS 2019 IEEE International Geoscience and Remote Sensing Symposium, 461-464. https://doi.org/10.1109/IGARSS.2019.8900123

[28] Garnot, V. S. F., Landrieu, L., Giordano, S., & Chehata, N. (2019). Time-space tradeoff in deep learning models for crop classification on satellite multi-spectral image time series. IGARSS 2019 IEEE International Geoscience and Remote Sensing Symposium, 6247-6250. https://doi.org/10.1109/IGARSS.2019.8900517

[29] Perez-Carabaza, S., Syrris, V., Kempeneers, P., & Soille, P. (2021). Crop classification from Sentinel-2 time series with temporal convolutional neural networks. 2021 IEEE International Geoscience and Remote Sensing Symposium, 6500-6503. https://doi.org/10.1109/IGARSS47720.2021.9554358

[30] Kussul, N., Lavreniuk, M., & Shumilo, L. (2020). Deep recurrent neural network for crop classification task based on Sentinel-1 and Sentinel-2 imagery. IGARSS 2020 IEEE International Geoscience and Remote Sensing Symposium, 6914-6917. https://doi.org/10.1109/IGARSS39084.2020.9324699

[31] Marinelli, D., Paris, C., & Bruzzone, L. (2019). An approach to tree detection based on the fusion of multitemporal LiDAR data. IEEE Geoscience and Remote Sensing Letters, 16(11), 1771-1775. https://doi.org/10.1109/LGRS.2019.2908314

[32] Ghamisi, P., Gloaguen, R., Atkinson, P. M., Benediktsson, J. A., Rasti, B., Yokoya, N., et al. (2019). Multisource and multitemporal data fusion in remote sensing: A comprehensive review of the state of the art. IEEE Geoscience and Remote Sensing Magazine, 7(1), 6-39. https://doi.org/10.1109/MGRS.2018.2890023
[33] Paoletti, M. E., Haut, J. M., Plaza, J., & Plaza, A. (2019). Deep learning classifiers for hyperspectral imaging: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 158, 279-317. https://doi.org/10.1016/j.isprsjprs.2019.09.006
[34] Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. ISPRS Journal of Photogrammetry and Remote Sensing, 152, 166-177. https://doi.org/10.1016/j.isprsjprs.2019.04.015
[35] Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. Nature, 566(7743), 195-204. https://doi.org/10.1038/s41586-019-0912-1
[36] Li, S., Song, W., Fang, L., Chen, Y., Ghamisi, P., & Benediktsson, J. A. (2019). Deep learning for hyperspectral image classification: An overview. IEEE Transactions on Geoscience and Remote Sensing, 57(9), 6690-6709. https://doi.org/10.1109/TGRS.2019.2907932
[37] Sedona, R., Cavallaro, G., Jitsev, J., Strube, A., Riedel, M., & Benediktsson, J. A. (2019). Remote sensing big data classification with high performance distributed deep learning. Remote Sensing, 11(24), 3056. https://doi.org/10.3390/rs11243056
[38] Rostami, M., Kolouri, S., Eaton, E., & Kim, K. (2019). Deep transfer learning for few-shot SAR image classification. Remote Sensing, 11(11), 1374. https://doi.org/10.3390/rs11111374
[39] Li, Y., Zhang, H., & Zhu, Q. (2019). Learning deep networks under noisy labels for remote sensing image scene classification. IGARSS 2019 IEEE International Geoscience and Remote Sensing Symposium, 3025-3028. https://doi.org/10.1109/IGARSS.2019.8900497
[40] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2020). Generative adversarial networks. Communications of the ACM, 63(11), 139-144. https://doi.org/10.1145/3422622