remote sensing

MDPI

Article

Mapping Crop Types for Beekeepers Using Sentinel-2 Satellite Image Time Series: Five Essential Crops in the Pollination Services

Navid Mahdizadeh Gharakhanlou [1,*], Liliana Perez [1] and Nico Coallier [2]

[1] Laboratoire de Géosimulation Environnementale (LEDGE), Département de Géographie, Université de Montréal, 1375 Avenue Thérèse-Lavoie-Roux, Montréal, QC H2V 0B3, Canada; l.perez@umontreal.ca
[2] Nectar Technologies Inc., 6250 Rue Hutchison #302, Montréal, QC H2V 4C5, Canada; nico@nectar.buzz
* Correspondence: navid.mahdizadeh.gharakhanlou@umontreal.ca

Abstract: Driven by the widespread adoption of deep learning (DL) in crop mapping with satellite image time series (SITS), this study was motivated by the recent success of temporal attention-based approaches in crop mapping. To meet the needs of beekeepers, this study aimed to develop DL-based classification models for mapping five essential crops in pollination services in Quebec province, Canada, by using Sentinel-2 SITS. Due to the challenging task of crop mapping using SITS, this study employed three DL-based models, namely one-dimensional temporal convolutional neural networks (CNNs) (1DTempCNNs), one-dimensional spectral CNNs (1DSpecCNNs), and long short-term memory (LSTM). Accordingly, this study aimed to capture expert-free temporal and spectral features, specifically targeting temporal features using 1DTempCNN and LSTM models, and spectral features using the 1DSpecCNN model. Our findings indicated that the LSTM model (macro-averaged recall of 0.80, precision of 0.80, F1-score of 0.80, and ROC of 0.89) outperformed both 1DTempCNNs (macro-averaged recall of 0.73, precision of 0.74, F1-score of 0.73, and ROC of 0.85) and 1DSpecCNNs (macro-averaged recall of 0.78, precision of 0.77, F1-score of 0.77, and ROC of 0.88) models, underscoring its effectiveness in capturing temporal features and highlighting its suitability for crop mapping using Sentinel-2 SITS. Furthermore, applying one-dimensional convolution (Conv1D) across the spectral domain demonstrated greater potential in distinguishing land covers and crop types than applying it across the temporal domain. This study contributes to providing insights into the capabilities and limitations of various DL-based classification models for crop mapping using Sentinel-2 SITS.

Keywords: satellite image time series (SITS); land cover classification; remote sensing (RS); deep learning (DL); geospatial artificial intelligence (GeoAI); beehives; Quebec; Canada

check for updates

Citation: Mahdizadeh Gharakhanlou, N.; Perez, L.; Coallier, N. Mapping Crop Types for Beekeepers Using Sentinel-2 Satellite Image Time Series: Five Essential Crops in the Pollination Services. *Remote Sens.* **2024**, *16*, 4225. https://doi.org/10.3390/rs16224225

Academic Editor: Haopeng Zhang

Received: 16 October 2024
Revised: 6 November 2024
Accepted: 11 November 2024
Published: 13 November 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Traditionally, crop-type data are gathered through field surveys, which are time-intensive and costly, especially for large-scale mapping [1]. The growing availability of remote sensing (RS) data now allows for the automatic creation of crop maps by leveraging spectral differences between crop types, offering a more efficient and scalable approach [2]. Some studies use single-date RS data to identify crops by selecting imagery based on the crop growth calendar [3]. Although single-date methods involve fewer image-processing tasks, they present challenges in identifying the best time for crop classification [4]. Alternatively, a satellite image time series (SITS) enables continuous monitoring of crop dynamics, allowing for the distinction of crop types based on their distinct spectral signatures and providing detailed insights into the various phenological stages throughout the growing season [5,6]. Effectively utilizing temporal sequences to capture fluctuations in crop growth is a significant challenge in crop mapping using SITS data. This has sparked interest in

leveraging methods capable of using SITS data to extract features that provide insights into vegetation growth stages.

Certain approaches derive phenological features from SITS for further analysis instead of directly using RS data. One approach is to use the original vegetation indice (VI) values during different periods of the year [7]. This approach could have the drawback of overlooking certain crucial information inherent in the time series inputs since the sequential relationship of VI is not expressly taken into account [8]. Accordingly, the temporal sequence of the images' presentation has no bearing on the results; in other words, rearranging the order of values in the series would not alter the model's outcomes or its level of accuracy. An alternative approach is to pre-compute temporal features from the VI time series [9–11]. This approach, due to taking into account the sequential relationship of Vis, may boost the accuracy of classification compared with the original VI values [12]. Employing or crafting temporal features from VIs such as maximum VI, approximations of crucial dates in the phenological stages of the vegetation classes such as the timing of peak VI, and beginning of green-up [13,14] or more complex temporal feature extractors such as pre-defined mathematical equations/models [15–17] have been extensively employed for multi-temporal classification and vegetation phenology investigations. This approach also poses several challenges in extracting temporal features from vegetation dynamics, including (i) the requirement for manual model design and feature extraction, which depends on human expertise and domain knowledge and may not be well suited for specific tasks [8]; (ii) manual feature engineering is time-consuming and is influenced by environmental changes [18]; and (iii) the use of fixed pre-defined models and mathematical assumptions can restrict adaptability when dealing with various temporal patterns [19]. Accordingly, addressing this issue necessitates more adaptable and automated methods that can make the most of the temporal aspect [20]. Deep learning (DL) presents promising solutions for addressing this challenge and demonstrates considerable potential in capturing and representing features from time series data obtained through RS [18].

In recent years, DL models have shown outstanding performance in various computer vision tasks such as image classification, semantic segmentation, object tracking, and object detection [21–23]. Unlike machine learning (ML) models that require empirical feature design and manually crafted features [24], DL models can automatically extract high-level features from high-dimensional images in an expert-free end-to-end fashion. Considering the strengths of DL models, various attempts have been made to apply DL architectures to the classification of image time series [8,25,26], particularly for the aim of crop mapping [8,27,28]. Two main DL architectures used in the classification of SITS are convolutional neural networks (CNNs) [8,15,26] and recurrent neural networks (RNNs) [20,25]. Numerous studies have highlighted the significant potential of these two DL models in the field of crop mapping [26,28–30].

CNNs stand out as a prominent DL architecture extensively employed in RS applications, especially land cover classification [31]. CNNs have been employed in RS research for SITS classification to extract spatial and/or spectral features, accomplished through either 1D convolution along the spectral dimension [32], 2D convolution across spatial dimensions [33], or 3D convolution across both spectral and spatial dimensions [34]. Nonetheless, applying convolutions in the spectral and/or spatial dimensions leads to overlooking the temporal aspect and sequence of images [26]. Applying convolutions across the temporal domain has demonstrated effectiveness in addressing temporal aspects in the SITS classification [26,35]. RNNs are another DL architecture that can link adjacent observations and capture dependencies across time steps to learn sequential relationships [36]. Due to their ability to process sequential data, RNNs are frequently viewed as a suitable choice for understanding the temporal dynamics in SITS and representing changes in land cover patterns [37]. The long short-term memory (LSTM), a popular variant of RNN, is created to overcome the shortcomings of traditional RNNs by enabling them to accurately capture long-term dependencies while circumventing the issue of vanishing gradients. The LSTM has been utilized in RS for various applications, including detecting changes in bi-temporal

images [37,38], classifying time series data to distinguish crops [25,30,39], analyzing land use and land cover [40,41], and studying vegetation dynamics [42].

Accurate crop prediction is crucial for beekeepers, aiding in strategic hive placement for optimal pollination efficiency and comprehensive coverage across diverse crop pollination needs. Accordingly, this study aimed to map five essential crops for pollination services across a region in Quebec, Canada, using Sentinel-2 SITS. To accomplish this aim, this study attempted to provide beekeepers with three DL-based models, namely one-dimensional temporal CNNs (1DTempCNNs), one-dimensional spectral CNNs (1DSpecCNNs), and LSTM. Furthermore, the motivations of this study were investigated to address the following research concerns: (i) assessing the performance of two DL architectures of CNNs and LSTM in crop mapping; (ii) investigating the potential of applying one-dimensional convolution (Conv1D) across the temporal domain (i.e., 1DTempCNNs model) and its efficacy in crop type prediction compared with applying Conv1D across the spectral domain (i.e., 1DSpecCNNs model), (iii) providing an approach for tuning models' architecture by assessing the impact of each hyperparameter in the model's architecture on its accuracy using the ANOVA (analysis of variance) method, and (iv) outlining the advantages and limitations of the provided DL-based models in predicting five essential crops for the beekeepers. The specific highlight of the presented study that differentiated it from preceding ones was its ability to address these concerns.

2. Materials and Methods

The methodology implemented in this research is described in Figure 1. The initial phase of the research involved gathering and preprocessing the ground reference data in ArcGIS Pro 3.3.0. The ground reference data aided in training the models, finding the optimized architectures of DL models, and evaluating their accuracy. Having reduced the number of samples in the majority classes and made a relative balance among classes, ground reference pixels were partitioned into training, validation, and test subsets. Following the division of ground reference pixels into training, validation, and test subsets and ensuring relatively consistent class distributions across all subsets, we examined whether the dataset partitioning adhered to the assumption of independence and identical distribution (IID). Accordingly, the conformity to the IID assumption was evaluated by comparing the cumulative distribution functions (CDFs) of subsets pairwise and examining the corresponding p-values through the Kolmogorov–Smirnov (KS) statistical test. Subsequently, 10-day interval median composites of Sentinel-2 satellite imagery comprising 10 spectral bands (i.e., blue, green, red, red-edge1, red-edge2, red-edge3, NIR, red-edge4, SWIR1, SWIR2) were created through the Earth Engine Python API on the Google Earth Engine (GEE) cloud-based platform. This process entailed capturing the median pixel values of corresponding pixels in the images for each band within 10-day intervals to accentuate the dynamics of vegetation growth. Then, the values of the spectral bands were extracted for all training, validation, and test pixels from April to October 2021. Subsequently, the models were trained with the training pixels, the optimal architecture for each DL model was determined using the validation pixels, and the accuracy of the models was assessed with the test pixels. Finally, the DL-based model exhibiting the highest accuracy was selected for predicting land cover and crop type throughout the study area.

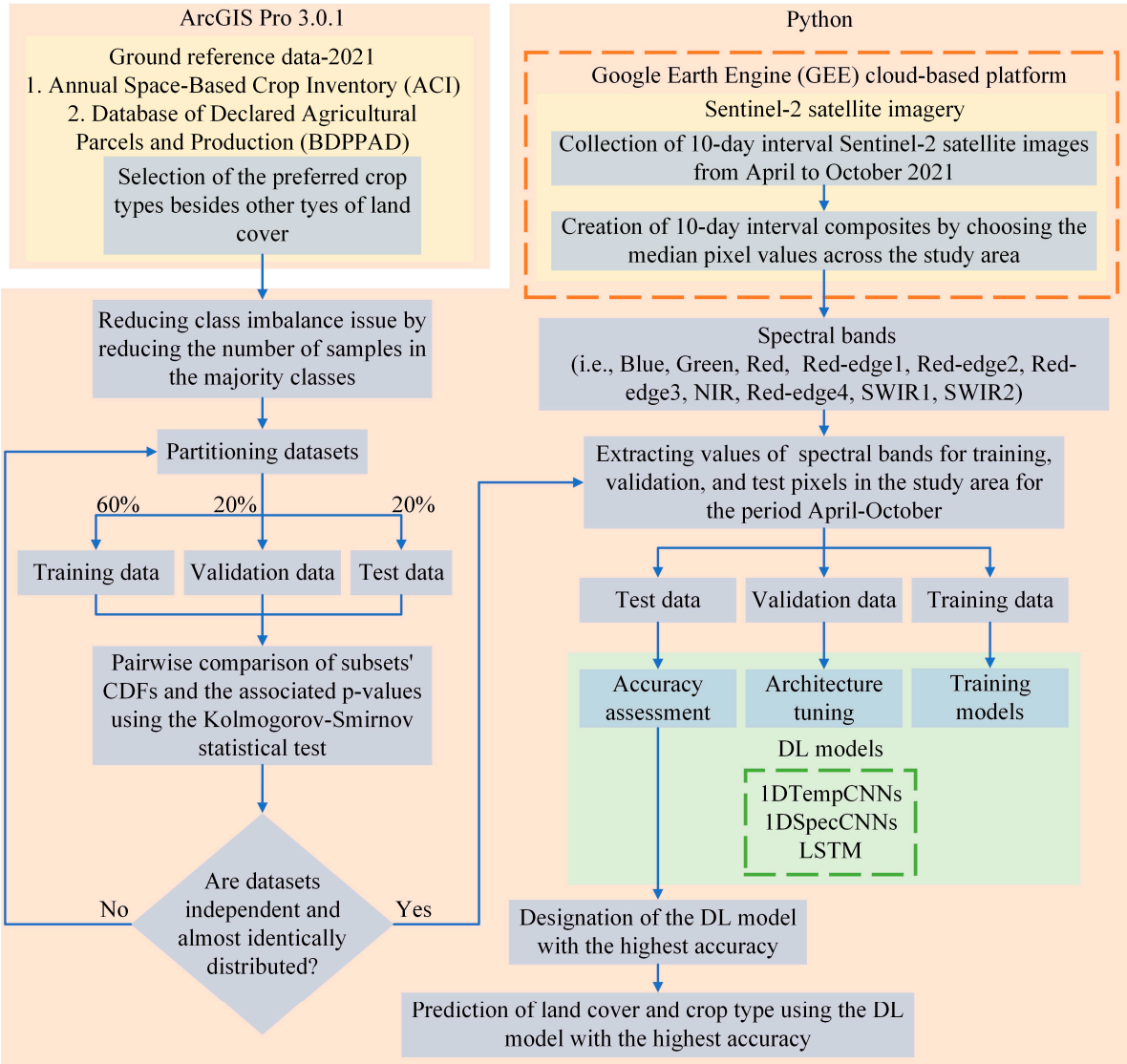


Figure 1. The flowchart of the research methodology.

2.1. Study Area

Our study was undertaken in a region in Quebec province, Canada (Figure 2), emphasizing the diverse cultivation of crucial crops for beekeepers, including apples, corn, soy, canola, and small fruits (e.g., berries, blueberries, strawberries, and raspberries). Quebec's croplands are crucial for supporting honeybee populations and enhancing crop yields [43]. Honeybees rely on these lands for forage, serving an essential role in pollination [43]. Predicting crop types helps beekeepers strategically place hives for optimal forage access, improving hive management and maximizing productivity. Understanding the crop landscape also aids in addressing challenges like pesticide exposure risks, promoting sustainability, and supporting pollinator well-being.

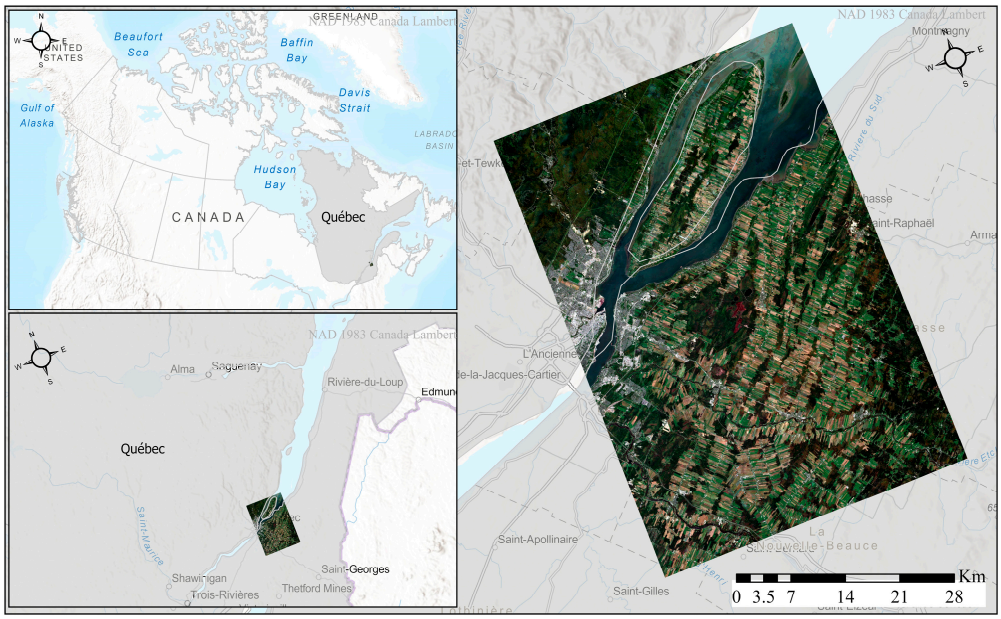


Figure 2. Geographic location of the study area with a true-color median composite of Sentinel-2 satellite imagery generated for 1–10 April 2021.

2.2. Data and Preprocessing of the Data

2.2.1. Satellite Imagery

This study utilized multi-temporal archived Level-2A orthorectified atmospherically corrected surface reflectance Sentinel-2 satellite images obtained from GEE via the Earth Engine Python API. The Sentinel-2 mission encompasses two satellites, namely Sentinel-2A and Sentinel-2B, launched in 2015 and 2017, respectively. Both satellites offer global coverage of the Earth's surface every 10 days, thereby enabling a combined revisit cycle of 5 days for the satellite constellation. The multi-temporal characteristics of Sentinel satellite imagery allow it to effectively capture the unique phenological and temporal patterns of different crops, drawing upon insights emphasized by Veloso et al. [44] regarding the accuracy of crop classification using multi-temporal satellite data.

In this study, Sentinel-2 products underwent preprocessing on the GEE platform, involving acquiring images at 10-day intervals with cloud coverage below 20% from April to October 2021. A median function was applied to each 10-day interval image collection to select the median pixel reflectance values, resulting in 10-day interval composites. The median function enabled the production of cloud-free images across the study area. Additionally, this function eliminated erroneous values caused by the presence of extremely bright, dark, or noisy pixels [45].

2.2.2. Ground Reference Data

Ground reference data for this research came from two sources: the Annual Space-Based Crop Inventory (ACI) provided by Agriculture and Agri-Food Canada (obtained from https://open.canada.ca accessed on 13 December 2023), and the Database of Declared Agricultural Parcels and Production (BDPPAD) from La Financière Agricole du Québec (obtained from https://www.fadq.qc.ca accessed on 11 February 2024). ACI maps out agricultural and non-agricultural land cover across Canada, while BDPPAD offers detailed information on agricultural parcels linked with client records. The data extracted from this source encompassed 13 classes, including 5 essential crops for beekeepers (namely apples, corn, canola, soy, and small fruits such as berries, blueberries, strawberries, and raspberries) along with 8 land cover classes.

2.2.3. Dataset Partition

In classifying land covers and crop types, addressing imbalanced distributions is crucial as it can skew training data and affect classification accuracy. To mitigate this, we balanced class distributions by reducing samples in majority classes, then divided pixels into training, validation, and test sets in proportions of 60%, 20%, and 20%, respectively. The sets, totaling 1,074,097, 358,027, and 358,047 pixels, respectively, were used for model training, hyperparameter and architecture tuning, and evaluation.

The conformity of the dataset partitioning to the IID assumption was evaluated by comparing the CDFs of the subsets using the KS statistical test. The KS test statistics were 1.8×10^{-6} for the training vs. validation datasets, 4.6×10^{-6} for the training vs. test datasets, and 5.72×10^{-6} for the validation vs. test datasets, all resulting in a p-value of 1.0. These low KS statistics indicate minimal differences in the distributions, and the p-value of 1.0 suggests no statistically significant differences among the dataset distributions. Therefore, the null hypothesis cannot be rejected, confirming that the training, validation, and test datasets are likely from the same distribution and adhere to the IID assumption.

2.3. Deep Learning (DL) Models

DL models belong to a category of artificial neural networks (ANNs) distinguished by their ability to discern complex patterns and representations within extensive datasets. Their multilayer architecture empowers them to acquire hierarchical understandings of input data, with each layer progressively capturing more abstract features. DL models are particularly proficient at autonomously extracting meaningful features or representations directly from raw data. Instead of the traditional approach of manual feature engineering, DL models autonomously glean pertinent features from the data during the training phase. This intrinsic capacity negates the necessity for extensive feature engineering, rendering DL models highly adaptable to diverse domains [36]. DL models can have various architectures tailored to specific tasks. This research centers on two widely recognized architectures, CNNs and LSTM, renowned for their remarkable ability to capture features [46,47].

2.3.1. Convolutional Neural Networks (CNNs)

CNNs are advanced neural network architectures tailored for processing structured grid data such as images or time series. They comprise several layers, including convolutional, pooling, and fully connected layers, employing convolutions to autonomously identify and capture significant features from the input data. This enables CNNs to excel in pattern recognition and classification tasks while preserving spatial relationships within the data [48]. The learning process in CNNs unfolds through several key stages. It commences with convolutional layers, where adaptable filters slide over the input image to discern features, resulting in feature maps that highlight where these features are present. These feature maps then undergo non-linear activation functions such as rectified linear unit (i.e., ReLU) to introduce complexity and facilitate the understanding of complex data relationships. Subsequently, pooling layers are introduced to reduce the spatial dimensions of feature maps, enhancing computational efficiency by retaining essential information while discarding less relevant details. After several convolutional and pooling layers, CNNs commonly incorporate one or more fully connected layers, enabling the model to perform classification or regression tasks based on the learned features. CNNs are trained through backpropagation, an optimization technique that iteratively adjusts the model's parameters to minimize the disparity between predicted and actual values. Optimization algorithms like stochastic gradient descent (SGD) or its variants (e.g., Adam) further refine the model's parameters by guiding them toward minimizing the loss function. This training process occurs over multiple epochs, allowing the model to progressively learn and enhance its feature representations [36]. The remaining part outlined two CNN architectures, 1DTempCNN and 1DSpecCNN. The 1DSpecCNN model extracts expert-free temporal and spectral features, focusing on spectral features, whereas 1DTempCNN crafts expert-free temporal and spectral features, focusing on temporal features.

- One-dimensional temporal CNNs (1DTempCNNs)

The 1DTempCNN architecture is designed to capture temporal dynamics in time series data by applying convolutional filters along the temporal dimension of Sentinel-2 satellite imagery. Each temporal observation is treated as a separate input, allowing the network to exploit sequential patterns related to crop development and environmental changes. By focusing on temporal dependencies and variations, the 1DTempCNN aims to extract features indicating seasonal and phenological shifts essential for distinguishing crop types. This architecture enhances the model's ability to recognize time-dependent features.

- One-dimensional spectral CNNs (1DSpecCNNs)

The 1DSpecCNN architecture is intended to extract features from Sentinel-2 imagery by applying convolutional filters over the spectral dimension of Sentinel-2 satellite imagery, representing 10 distinct bands. This approach leverages spectral diversity to identify subtle differences in land cover and crop types. The 1DSpecCNN captures unique features, such as reflectance variations, that indicate specific vegetation or land characteristics by focusing on the spectral dimension.

2.3.2. Long Short-Term Memory (LSTM)

RNNs are a type of ANN uniquely designed for processing sequential data. Their primary strength lies in their ability to understand and learn patterns within sequences, making them highly suitable for tasks such as time series data. LSTM, one of the most common RNN architectures, known for achieving state-of-the-art results in sequential data tasks, excels in capturing long-term dependencies. Its architectural design features a memory cell that retains current states across sequential instances, along with non-linear dependencies regulating the flow of information into and out of the cell [49]. The learning process in the LSTM architecture involves several stages. Initially, the LSTM receives input data, processes it through gate mechanisms, and updates its memory cell state. Gate operations, including the forget, input, and output gates, regulate the flow of information. Non-linear activation functions, such as sigmoid and hyperbolic tangent, introduce complexity and capture patterns in the data. The LSTM generates an output based on the input and updated memory cell state at each time step. During training, techniques like backpropagation through time (BPTT) optimize the model's parameters to minimize prediction errors over multiple time steps. In essence, the LSTM learns by iteratively adjusting its memory cells, gate operations, and output generation to capture long-term dependencies in sequential data [50].

2.4. Architecture Tuning in DL Models

Crafting and refining architectures in DL models are vital for enhancing performance, but this process is more complex due to the adaptability of architectures and the lack of standardized approaches. In this research, architecture tuning of CNNs and LSTM models followed an iterative trial-and-error process, starting with simpler architectures and progressively refining them based on classification performance on the validation set. This involved exploring various combinations of hyperparameters such as the number of layers, filters/cells per layer, and dropout rate. Among the designed architectures, those exhibiting promising classification performance on the validation set were chosen as the basis for further exploration. Accordingly, after establishing the initial architectures for each DL model, we proceeded to enhance the architectures by adjusting one or two hyperparameters, adding layers, rearranging existing layers, or replacing components with more complicated alternatives. This iterative approach allowed the architecture to evolve in size and complexity until no further improvement was observed.

3. Results

3.1. Architectures of DL Models with Optimal Performance

This study established the initial architecture of the 1DTempCNN, 1DSpecCNN, and LSTM models by evaluating the models' accuracy across multiple candidate values for three hyperparameters in each model's architecture (Table 1). The performance of each of the 100 designed architectures was assessed based on the macro-average of the F1-score on the validation dataset. Heatmaps were used to visualize the accuracy of each designed architecture for each model (Figure 3). For 1DTempCNNs, accuracy declined as the number of layers increased, while better performance was often observed with a high number of filters per layer and lower dropout rates. Similarly, for 1DSpecCNNs, increasing the number of layers led to a decrease in accuracy, indicating that deeper architectures may not necessarily enhance classification accuracy on the validation dataset (i.e., the model's generalizability). The 1DSpecCNN model generally performed better with fewer filters per layer, despite fluctuations in accuracy. Lower dropout rates were also linked to higher accuracy for 1DSpecCNN architectures. In the LSTM model, accuracy generally improved up to two layers, with higher performance often seen when using a greater number of cells per layer and lower dropout rates, although some fluctuations in accuracy were noted. It is thus recommended to use relatively shallow architectures with lower dropout rates to improve the model's classification accuracy.

Table 1. Candidate values for three hyperparameters and their corresponding optimal values in the initial architecture configuration of DL models.

Models	Hyperparameters of the Architecture	Candidate Values	Optimal Value
1DTempCNNs	Number of layers	2, 4, 6, 8	2
	Number of filters per layer	32, 64, 128, 256, 512	256
	Dropout rate	0.1, 0.2, 0.3, 0.4, 0.5	0.1
1DSpecCNNs	Number of layers	2, 4, 6, 8	2
	Number of filters per layer	32, 64, 128, 256, 512	64
	Dropout rate	0.1, 0.2, 0.3, 0.4, 0.5	0.1
LSTM	Number of layers	1, 2, 3, 4	2
	Number of cells per layer	32, 64, 128, 256, 512	512
	Dropout rate	0.1, 0.2, 0.3, 0.4, 0.5	0.1

To examine the impact of each hyperparameter in the model's architecture on its accuracy, this study employed the ANOVA method in conjunction with the heatmap (Figure 3). ANOVA is a statistical method used to assess the impact of independent variables on a dependent variable. It evaluates whether the dependent variable's average values vary significantly across each independent variable's levels. For each factor (i.e., independent variable), ANOVA calculates an F-statistic that quantifies the variance attributed to that factor concerning random error. A high F-value and a low p-value suggest that the observed differences in means are likely meaningful and not due to random chance, indicating a significant effect of the factor on the outcome. [51]. Table 2 outlines the results of the ANOVA method for each hyperparameter in the architecture of all three DL-based models. The ANOVA results indicated that the number of layers and the dropout rate significantly impacted model accuracy more than the number of filters per layer in both 1DTempCNN and 1DSpecCNN models. In the LSTM model, all three hyperparameters had a significant impact on performance, with the number of layers and dropout rate having the largest effects, followed by the number of cells per layer with a slightly lower influence.

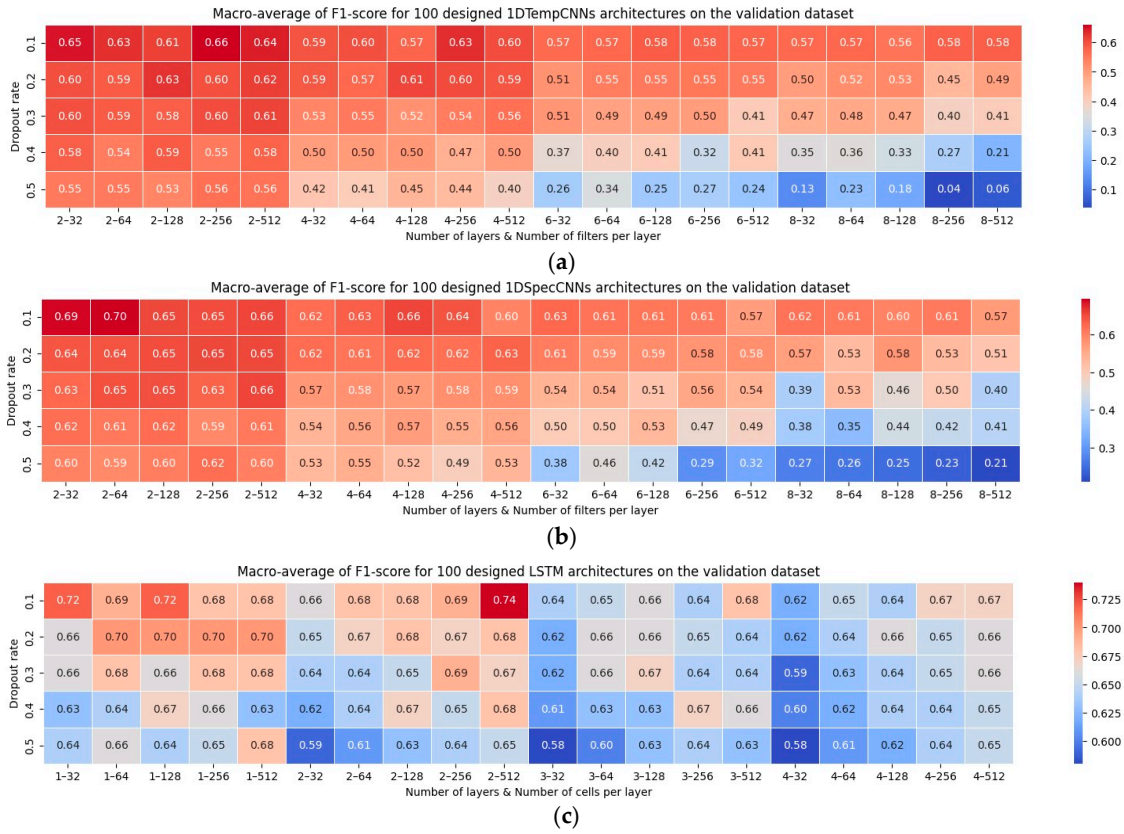


Figure 3. The macro-average of the F1-score for the 100 designed architectures on the validation dataset for (**a**) 1DTempCNN, (**b**) 1DSpecCNN, and (**c**) LSTM models.

Table 2. The ANOVA results for each hyperparameter in the models' architecture.

Models	Hyperparameters of the Architecture	Sum of Squares (SS)	F-Value	*p*-Value
1DTempCNNs	Number of layers	0.586	16.577	9.327×10^{-9}
	Number of filters per layer	0.008	0.112	0.978
	Dropout rate	0.814	21.424	1.283×10^{-12}
1DSpecCNNs	Number of layers	0.482	23.031	2.577×10^{-11}
	Number of filters per layer	0.007	0.139	0.967
	Dropout rate	0.45	15.263	1.131×10^{-9}
LSTM	Number of layers	0.021	10.495	5.0×10^{-6}
	Number of cells per layer	0.018	6.438	1.25×10^{-4}
	Dropout rate	0.024	9.417	2.0×10^{-6}

As explained in Section 2.4, the architecture tuning process of DL models involved an iterative trial-and-error approach, refining architectures based on classification performance on the validation set. After investigating the influence of each hyperparameter in the model's architecture on its accuracy, exploring potential modifications in the models' architecture, and ultimately designing initial architectures, optimal architectures were designated for each model. For the 1DTempCNN model, the optimal architecture (Figure 4) included 1 Conv1D layer with 256 filters (kernel size of 3) and an inception layer with 3 convolutional branches of various filters (256 and 512 filters with kernel sizes of 1, 3, and 5) following the first Conv1D layer. Each of the Conv1D layer and the inception layer was followed by a dropout layer with a dropout rate of 0.1, and 2 dense layers with 512 and

13 units at the end of the architecture. For the 1DSpecCNN model, the optimal architecture (Figure 5) comprised a Conv1D layer with 64 filters and a kernel size of 3, followed by an inception layer with 3 convolutional branches of different filters (32 and 64 filters with kernel sizes 1, 3, and 5). As with the optimal 1DTempCNN architecture, each Conv1D and inception layer in the optimal 1DSpecCNN architecture was followed by a dropout layer with a dropout rate of 0.1, and 2 dense layers with 512 and 13 units at the end of the architecture. Finally, for the LSTM model, the optimal architecture (Figure 6) consisted of 2 LSTM layers (512 and 128 cells), each followed by a dropout layer with a dropout rate of 0.1, and 2 dense layers with 512 and 13 units at the end of the architecture.

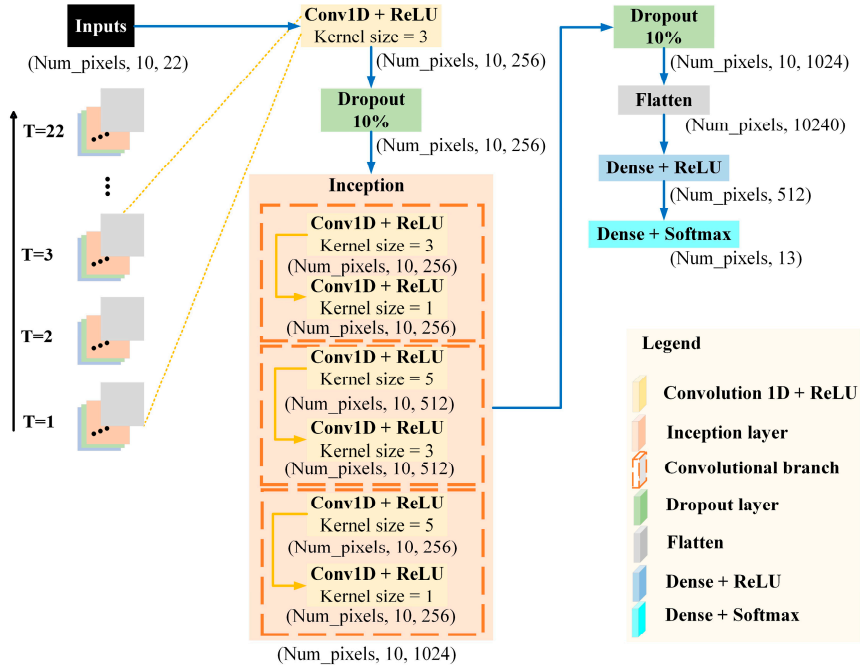


Figure 4. The 1DTempCNN architecture with optimal performance.

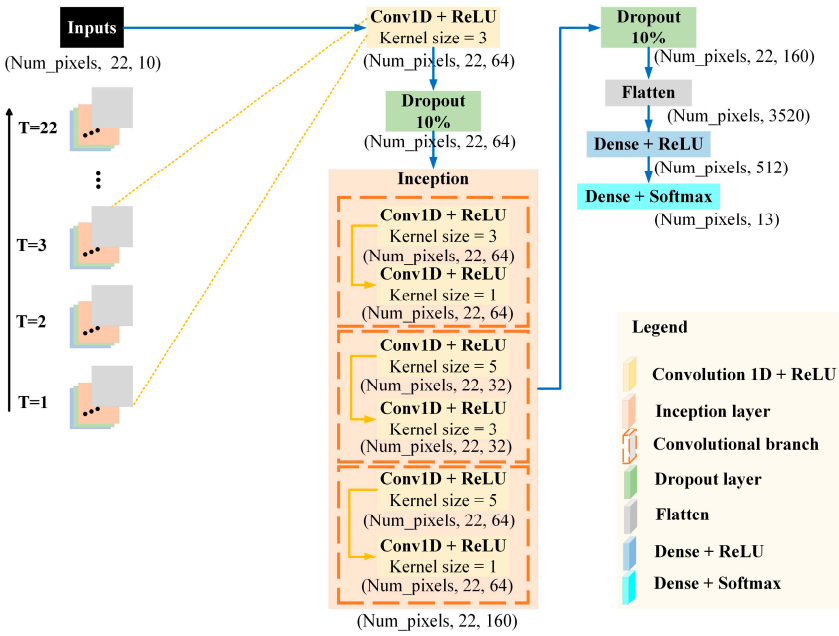


Figure 5. The 1DSpecCNN architecture with optimal performance.

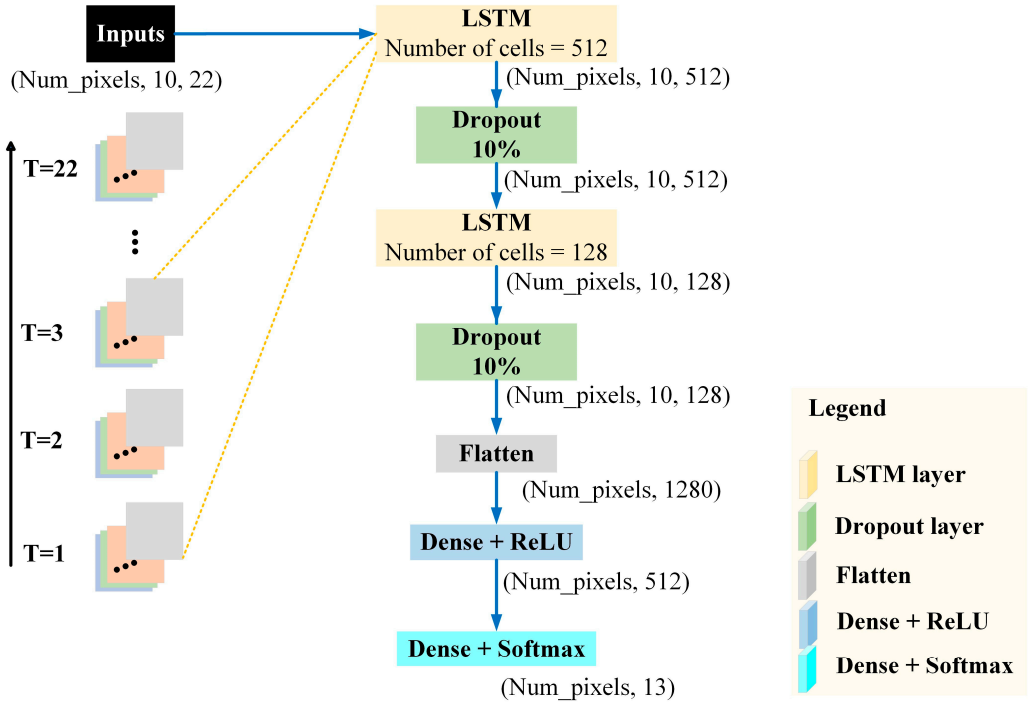


Figure 6. The LSTM architecture with optimal performance.

3.2. Performance Assessment of Models

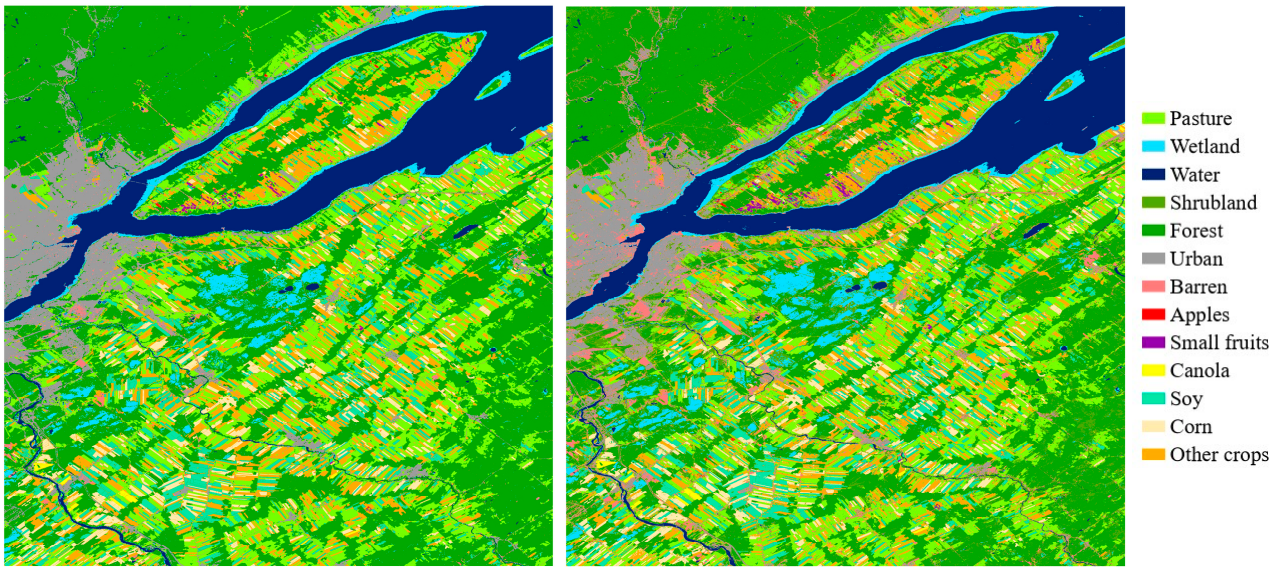
To quantitatively evaluate and compare the performance of DL-based models in land cover and crop type classification and verify their alignment with real-world conditions, four accuracy metrics were employed: recall, precision, F1-score, and receiver operating characteristics curve (ROC). Table 3 delineates the macro-averaged accuracies of the models, indicating their performance in predicting land covers and crop types on the test dataset.

Regarding the results in Table 3, the LSTM model outperformed both 1DTempCNN and 1DSpecCNN models. Notably, applying Conv1D across the spectral domain showed greater potential for distinguishing land covers and crop types than applying it across the temporal domain, as evidenced by the performance of 1DTempCNN and 1DSpecCNN models on the test dataset. Furthermore, the per-class accuracy assessment (Table 4) indicated that, despite notable differences in the DL models' potential to differentiate between classes, they performed exceptionally well in identifying five essential crops for beekeepers compared with other land covers. Figure 7 illustrates the LSTM model's predictions for land cover and crop types across the entire study area, alongside the corresponding ground reference labels.

Centering on distinguishing five essential crop types for beekeepers from other crops and land covers using the top-performing model, errors in the classification, denoted by off-diagonal values in the confusion matrix of the LSTM model (Figure A1, Appendix A), generally fall into several categories as follows: (i) "crop as non-crop" errors resulting from misclassifying crops as other land covers, (ii) "non-crop as crop" errors stemming from classifying other land covers as crops, (iii) "essential crops as other crops" errors resulting from misclassifying five essential crops as other crops, (iv) "other crops as essential crops" errors related to misclassifying other crops as five essential crops, and (v) "inter-essential crops" errors corresponding to misclassifying between five essential crops for beekeepers. Regarding the findings, the top-performing DL-based model (i.e., LSTM) resulted in 10,000 "crop as non-crop", 7401 "non-crop as crop", 2245 "essential crops as other crops", 5285 "other crops as essential crops", and 1367 "inter-essential crops" misclassified pixels out of a total of 358,047 test pixels.

Table 3. The macro-averaged accuracy values of DL-based models on the test dataset using four accuracy criteria.

Models	Accuracy Criteria	Values
LSTM	Recall	0.80
	Precision	0.80
	F1-score	0.80
	ROC	0.89
1DTempCNNs	Recall	0.73
	Precision	0.74
	F1-score	0.73
	ROC	0.85
1DSpecCNNs	Recall	0.78
	Precision	0.77
	F1-score	0.77
	ROC	0.88



(**a**) (**b**)

Figure 7. (**a**) The ground reference map; and (**b**) the LSTM-provided map of land cover and crop type across the entire study area.

Table 4. Per-class accuracy assessment of models in land cover and crop type classification for the test dataset.

Models	Accuracy Criteria	Classes												
		Pasture	**Wetland**	**Water**	**Shrubland**	**Forest**	**Urban**	**Barren**	**Apples**	**Small Fruits**	**Canola**	**Soy**	**Corn**	**Other Crops**
LSTM	Recall	0.75	0.76	0.98	0.63	0.72	0.64	0.60	**0.93**	**0.94**	**0.98**	**0.93**	**0.94**	0.65
	Precision	0.77	0.81	0.97	0.55	0.71	0.67	0.63	**0.84**	**0.89**	**0.97**	**0.89**	**0.88**	0.77
	F1-score	0.76	0.79	0.97	0.58	0.72	0.65	0.61	**0.89**	**0.92**	**0.98**	**0.91**	**0.91**	0.71
	ROC	0.86	0.88	0.98	0.78	0.85	0.81	0.78	**0.97**	**0.97**	**0.99**	**0.96**	**0.96**	0.81
1DTempCNNs	Recall	0.76	0.71	0.95	0.59	0.70	0.66	0.45	**0.75**	**0.79**	**0.94**	**0.90**	**0.86**	0.46
	Precision	0.68	0.80	0.98	0.48	0.67	0.56	0.61	**0.76**	**0.83**	**0.93**	**0.79**	**0.83**	0.75
	F1-score	0.72	0.75	0.97	0.53	0.69	0.61	0.52	**0.75**	**0.81**	**0.94**	**0.84**	**0.84**	0.57
	ROC	0.86	0.85	0.98	0.76	0.84	0.81	0.71	**0.87**	**0.89**	**0.97**	**0.94**	**0.92**	0.72
1DSpecCNNs	Recall	0.75	0.74	0.97	0.63	0.72	0.61	0.55	**0.87**	**0.90**	**0.96**	**0.93**	**0.91**	0.55
	Precision	0.73	0.81	0.97	0.49	0.68	0.66	0.59	**0.74**	**0.87**	**0.97**	**0.86**	**0.87**	0.79
	F1-score	0.74	0.77	0.97	0.55	0.70	0.63	0.57	**0.80**	**0.88**	**0.96**	**0.89**	**0.89**	0.65
	ROC	0.86	0.86	0.98	0.77	0.85	0.79	0.76	**0.94**	**0.95**	**0.98**	**0.95**	**0.95**	0.77

Note: Bold values indicate the model's accuracy for the five target crop classes.

4. Discussion

Earth observation satellites equipped with diverse sensors gather data with varying spatial, spectral, and temporal resolutions, along with different resistance to atmospheric conditions. Over the past decade, Earth observation (EO) imagery has become the primary data source for classifying land cover and mapping crops due to its efficiency and rapid data acquisition [52]. Furthermore, there has been a significant increase in the adoption of methods aimed at improving the classification accuracy of EO imagery for crop mapping [27]. These methods have evolved from relying solely on individual images [53] to incorporating time series imagery [54]. Integrating time series of EO satellite imagery enhances performance by capturing changes in crop phenology throughout the growing season [55,56]. Traditional classification methods often overlook temporal connections within image time series [27], highlighting the need for more advanced methods capable of accounting for these relationships. Geospatial artificial intelligence (GeoAI), utilizing powerful learning methods such as ML and DL, has emerged as a dominant approach for classifying EO imagery by addressing temporal relationships within image time series. This study, aligning with the objectives of beekeepers, aimed to employ GeoAI models to capture temporal relationships within Sentinel-2 SITS, thereby enhancing crop mapping accuracy for predicting five essential crops in pollination services.

Two main classification approaches are utilized for analyzing EO satellite imagery in crop mapping: pixel-based [57,58] and object-based [59,60]. In pixel-based classification, each pixel in the image is classified individually based on its spectral properties, operating at the pixel level. Conversely, object-based classification involves grouping adjacent pixels into homogeneous and meaningful objects or segments, considering higher level entities. These image objects are then assigned target classes using supervised or unsupervised classification techniques [61]. Both approaches have their strengths and weaknesses, and the choice between them depends on specific objectives, data characteristics, and available resources. In alignment with recent research focusing on crop mapping using DL methods [20,26,28], this study focused on applying the pixel-based classification approach.

Given the ability of DL models to automatically extract high-level features in an expert-free end-to-end fashion, this study aimed to capture expert-free temporal and spectral features, specifically targeting temporal features using 1DTempCNN and LSTM models, and spectral features using the 1DSpecCNN model. Our findings indicated that applying Conv1D across the spectral domain proved more effective for distinguishing land covers and crop types than applying it across the temporal domain. Additionally, our findings, consistent with those of Rußwurm and Korner [62], demonstrated that the LSTM model outperformed the CNN models, underscoring its effectiveness in capturing temporal features and highlighting its suitability for crop mapping using multi-temporal Sentinel-2 satellite imagery. Furthermore, our findings, consistent with the earlier studies [62–64], demonstrated the LSTM model's potential to provide a more accurate solution for crop recognition.

DL models tend to favor majority classes, as every sample equally influences internal parameters during training. Previous research has addressed imbalanced data challenges through various strategies at both data and algorithm levels. At the data level, methods like oversampling [40] and undersampling [65] have been proposed, involving random sampling with or without replacement to effectively tackle class imbalances. At the algorithmic level, methods such as cost-sensitive learning have been employed to circumvent dataset manipulation by assigning penalties to samples based on their importance in the training set [66]. These penalties can be manually assigned through class or sample weights [67], systematically determined [68], estimated from occurrence probabilities [69], or dynamically adjusted during training based on sample difficulty [70]. The issue of class imbalance was addressed in this study through the adoption of strategies at both the data and algorithm levels. This study employed a tailored approach of random undersampling to reduce the number of samples in majority classes by a specific ratio, alongside utilizing the "Compute Class Weight" function provided in the scikit-learn library [71].

The IID assumption of data, fundamental to both ML and DL models, simplifies modeling processes and enables reliable inferences about the underlying population from observed data. However, in practical situations, meeting the IID assumption may not always be feasible or realistic. Spatio-temporal datasets often exhibit spatial and/or temporal dependencies and variations that contradict this assumption [72]. For instance, time series data may show temporal dependencies among samples collected at different intervals, while spatial data may demonstrate spatial dependencies among samples from neighboring locations, both challenging the independence assumption. Additionally, class-imbalanced datasets, typical in classification tasks, violate the assumption of identical distribution as sample distribution varies across different classes. Although ML and DL algorithms may yield satisfactory outcomes on a given dataset, it is essential to thoroughly evaluate the performance and generalizability of these models when applied to an independent dataset, particularly in cases where the IID assumption may be violated. Our findings underscored the limited generalizability of DL models due to real-world datasets not adhering to the IID assumption. By applying the trained DL models to various independent datasets, this research demonstrated the limited generalizability of the DL models, highlighting the necessity for specifically tailored methods to effectively handle spatial autocorrelation and heterogeneity (intrinsic to spatial data) and temporal dependencies in the time series data, which contradict the IID assumption.

The primary limitation of this study was associated with the ground reference data. While this research utilized ground reference data from two different sources, both sources had their own inaccuracies, resulting in errors being introduced during the model training phase. Supervised EO-based crop mapping relies heavily on accurate ground reference data, which are challenging to obtain due to restricted access to certain regions, the need for consistent and up-to-date surveys, and potential inaccuracies in observations. These challenges can affect the precision and reliability of EO-generated crop maps, particularly in areas with insufficient or inconsistent ground reference data.

5. Conclusions

EO provides large-scale high-resolution data for monitoring Earth's environment, which, combined with GeoAI modeling, enhances the analysis of geospatial big data for applications like land cover classification and crop mapping. This study focused on using EO imagery for crop mapping, aiding agricultural sustainability by assisting beekeepers in making informed decisions about hive placement to ensure diverse diets and reduce pesticide exposure. This strategy improves bee pollination efficiency, increasing crop yields and reducing bee mortality. This research aimed to map thirteen land cover classes, with particular emphasis on five essential crops in the pollination services—apples, corn, soy, canola, and small fruits—in Quebec, Canada, using Sentinel-2 SITS. To accomplish this aim, this study employed three DL-based models, namely 1DTempCNN, 1DSpecCNN, and LSTM. In addressing the research concerns, the findings of this research indicated that (i) the LSTM model outperformed the 1DTempCNN and 1DSpecCNN models in land cover and crop type classification; (ii) applying Conv1D across the spectral domain was more effective for distinguishing land covers and crop types than across the temporal domain; (iii) using a systematic approach for the architecture tuning of models through assessing the impact of hyperparameters on models' accuracy helped in optimizing models' configurations; and (iv) although the developed DL-based models yielded satisfactory outcomes on all training, validation, and test datasets, our findings underscored the constrained generalizability of models attributed to real-world datasets not adhering to the IID assumption. This research contributes to evaluating various DL-based models for crop mapping, offering valuable insights for advancing future research and practical applications in agriculture and environmental management. Furthermore, this study underscores both the strengths and the limitations of using GeoAI modeling with Sentinel-2 SITS to address crop mapping challenges.

Author Contributions: Conceptualization, N.M.G. and L.P.; methodology, N.M.G. and L.P.; software, N.M.G.; validation, N.M.G.; formal analysis, N.M.G. and L.P.; investigation, N.M.G. and L.P.; resources, N.M.G.; data curation, N.M.G.; writing—original draft preparation, N.M.G.; writing—review and editing, N.M.G., L.P. and N.C.; visualization, N.M.G.; supervision, L.P.; project administration, L.P.; funding acquisition, L.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the Institut de Valorisation des Données (IVADO): REG0 0006; the Mathematics of Information Technology and Complex Systems (Mitacs): IT34553; and the Natural Sciences and Engineering Research Council (NSERC): RGPIN/05396–2016 awarded to L.P.

Data Availability Statement: Datasets used in this research were publicly accessible at the time of writing this paper. Furthermore, all the code to replicate our findings can be found in the following link: https://github.com/Nmg1994/Crop_mapping/tree/main.

Acknowledgments: The authors gratefully acknowledge the support of Compute Canada for providing advanced research computing (ARC) resources that facilitated the execution of the computational tasks in this research.

Conflicts of Interest: Author Nico Coallier was employed by the company Nectar Technologies Inc. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A. Confusion Matrix of the LSTM Model

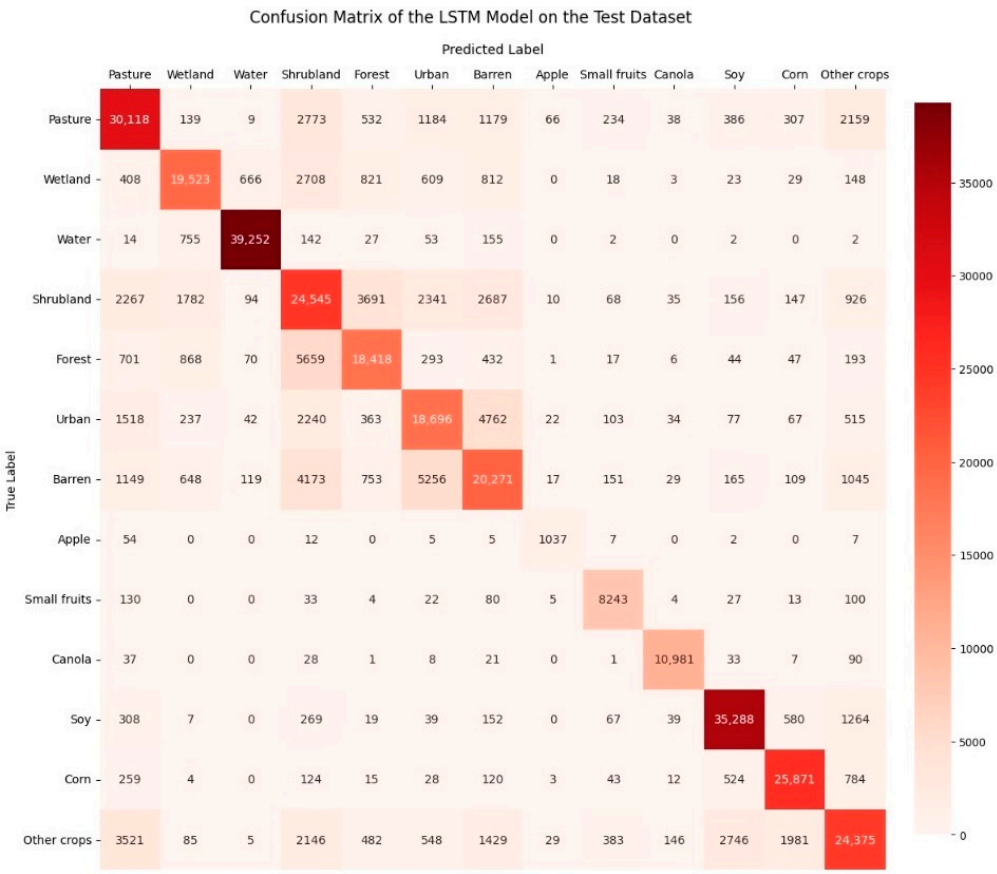


Figure A1. Confusion matrix of the top-performing DL model (i.e., LSTM) in predicting land cover and crop type on the test dataset.

References

1. Pott, L.P.; Amado, T.J.C.; Schwalbert, R.A.; Corassa, G.M.; Ciampitti, I.A. Satellite-based data fusion crop type classification and mapping in Rio Grande do Sul, Brazil. *ISPRS J. Photogramm. Remote Sens.* **2021**, *176*, 196–210. [CrossRef]
2. Waldner, F.; Fritz, S.; Di Gregorio, A.; Defourny, P. Mapping priorities to focus cropland mapping activities: Fitness assessment of existing global, regional and national cropland maps. *Remote Sens.* **2015**, *7*, 7959–7986. [CrossRef]
3. Mathur, A.; Foody, G.M. Multiclass and binary SVM classification: Implications for training and classification users. *IEEE Geosci. Remote Sens. Lett.* **2008**, *5*, 241–245. [CrossRef]
4. Wang, Y.; Zhang, Z.; Feng, L.; Ma, Y.; Du, Q. A new attention-based CNN approach for crop mapping using time series Sentinel-2 images. *Comput. Electron. Agric.* **2021**, *184*, 106090. [CrossRef]
5. Rogan, J.; Franklin, J.; Roberts, D.A. A comparison of methods for monitoring multitemporal vegetation change using Thematic Mapper imagery. *Remote Sens. Environ.* **2002**, *80*, 143–156. [CrossRef]
6. Xie, Y.; Sha, Z.; Yu, M. Remote sensing imagery in vegetation mapping: A review. *J. Plant Ecol.* **2008**, *1*, 9–23. [CrossRef]
7. Wardlow, B.D.; Egbert, S.L. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the US Central Great Plains. *Remote Sens. Environ.* **2008**, *112*, 1096–1116.
8. Zhong, L.; Hu, L.; Zhou, H. Deep learning based multi-temporal crop classification. *Remote Sens. Environ.* **2019**, *221*, 430–443. [CrossRef]
9. Jia, K.; Liang, S.; Wei, X.; Yao, Y.; Su, Y.; Jiang, B.; Wang, X. Land cover classification of Landsat data with phenological features extracted from time series MODIS NDVI data. *Remote Sens.* **2014**, *6*, 11518–11532. [CrossRef]
10. Valero, S.; Morin, D.; Inglada, J.; Sepulcre, G.; Arias, M.; Hagolle, O.; Dedieu, G.; Bontemps, S.; Defourny, P.; Koetz, B. Production of a dynamic cropland mask by processing remote sensing image series at high temporal and spatial resolutions. *Remote Sens.* **2016**, *8*, 55. [CrossRef]
11. Pittman, K.; Hansen, M.C.; Becker-Reshef, I.; Potapov, P.V.; Justice, C.O. Estimating global cropland extent with multi-year MODIS data. *Remote Sens.* **2010**, *2*, 1844–1863. [CrossRef]
12. Simonneaux, V.; Duchemin, B.; Helson, D.; Er-Raki, S.; Olioso, A.; Chehbouni, A. The use of high-resolution image time series for crop classification and evapotranspiration estimate over an irrigated area in central Morocco. *Int. J. Remote Sens.* **2008**, *29*, 95–116. [CrossRef]
13. Walker, J.; De Beurs, K.; Wynne, R. Dryland vegetation phenology across an elevation gradient in Arizona, USA, investigated with fused MODIS and Landsat data. *Remote Sens. Environ.* **2014**, *144*, 85–97. [CrossRef]
14. Walker, J.; De Beurs, K.; Henebry, G. Land surface phenology along urban to rural gradients in the US Great Plains. *Remote Sens. Environ.* **2015**, *165*, 42–52. [CrossRef]
15. Geerken, R.A. An algorithm to classify and monitor seasonal variations in vegetation phenologies and their inter-annual change. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 422–431. [CrossRef]
16. Galford, G.L.; Mustard, J.F.; Melillo, J.; Gendrin, A.; Cerri, C.C.; Cerri, C.E. Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil. *Remote Sens. Environ.* **2008**, *112*, 576–587. [CrossRef]
17. Siachalou, S.; Mallinis, G.; Tsakiri-Strati, M. A hidden Markov models approach for crop classification: Linking crop phenology to time series of multi-sensor remote sensing data. *Remote Sens.* **2015**, *7*, 3633–3650. [CrossRef]
18. Zhu, X.X.; Tuia, D.; Mou, L.; Xia, G.-S.; Zhang, L.; Xu, F.; Fraundorfer, F. Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 8–36. [CrossRef]
19. Zhong, L.; Hawkins, T.; Biging, G.; Gong, P. A phenology-based approach to map crop types in the San Joaquin Valley, California. *Int. J. Remote Sens.* **2011**, *32*, 7777–7804. [CrossRef]
20. Xu, J.; Zhu, Y.; Zhong, R.; Lin, Z.; Xu, J.; Jiang, H.; Huang, J.; Li, H.; Lin, T. DeepCropMapping: A multi-temporal deep learning approach with improved spatial generalizability for dynamic corn and soybean mapping. *Remote Sens. Environ.* **2020**, *247*, 111946. [CrossRef]
21. Ma, A.; Wan, Y.; Zhong, Y.; Wang, J.; Zhang, L. SceneNet: Remote sensing scene classification deep learning network using multi-objective neural evolution architecture search. *ISPRS J. Photogramm. Remote Sens.* **2021**, *172*, 171–188. [CrossRef]
22. Wang, L.; Li, R.; Zhang, C.; Fang, S.; Duan, C.; Meng, X.; Atkinson, P.M. UNetFormer: A UNet-like transformer for efficient semantic segmentation of remote sensing urban scene imagery. *ISPRS J. Photogramm. Remote Sens.* **2022**, *190*, 196–214. [CrossRef]
23. Lyu, Y.; Yang, M.Y.; Vosselman, G.; Xia, G.-S. Video object detection with a convolutional regression tracker. *ISPRS J. Photogramm. Remote Sens.* **2021**, *176*, 139–150. [CrossRef]
24. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
25. Teimouri, N.; Dyrmann, M.; Jørgensen, R.N. A novel spatio-temporal FCN-LSTM network for recognizing various crop types using multi-temporal radar images. *Remote Sens.* **2019**, *11*, 990. [CrossRef]
26. Pelletier, C.; Webb, G.I.; Petitjean, F. Temporal convolutional neural network for the classification of satellite image time series. *Remote Sens.* **2019**, *11*, 523. [CrossRef]
27. Mohammadi, S.; Belgiu, M.; Stein, A. Improvement in crop mapping from satellite image time series by effectively supervising deep neural networks. *ISPRS J. Photogramm. Remote Sens.* **2023**, *198*, 272–283. [CrossRef]
28. Zhao, H.; Duan, S.; Liu, J.; Sun, L.; Reymondin, L. Evaluation of five deep learning models for crop type mapping using sentinel-2 time series images with missing information. *Remote Sens.* **2021**, *13*, 2790. [CrossRef]

29. Kussul, N.; Lavreniuk, M.; Skakun, S.; Shelestov, A. Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 778–782. [CrossRef]
30. Sun, Z.; Di, L.; Fang, H. Using long short-term memory recurrent neural network in land cover classification on Landsat and Cropland data layer time series. *Int. J. Remote Sens.* **2019**, *40*, 593–614. [CrossRef]
31. Fırat, H.; Asker, M.E.; Hanbay, D. Classification of hyperspectral remote sensing images using different dimension reduction methods with 3D/2D CNN. *Remote Sens. Appl. Soc.* **2022**, *25*, 100694. [CrossRef]
32. Hu, W.; Huang, Y.; Wei, L.; Zhang, F.; Li, H. Deep convolutional neural networks for hyperspectral image classification. *J. Sens.* **2015**, *2015*, 258619. [CrossRef]
33. Mou, L.; Ghamisi, P.; Zhu, X.X. Unsupervised spectral–spatial feature learning via deep residual Conv–Deconv network for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *56*, 391–406. [CrossRef]
34. Li, Y.; Xiao, B.; Sun, L.; Wang, X.; Gao, Y.; Wang, Y. Phonon spectrum, IR and Raman modes, thermal expansion tensor and thermal physical properties of M2TiAlC2 (M= Cr, Mo, W). *Comput. Mater. Sci.* **2017**, *134*, 67–83. [CrossRef]
35. Wang, Z.; Yan, W.; Oates, T. Time series classification from scratch with deep neural networks: A strong baseline. In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017.
36. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
37. Mou, L.; Bruzzone, L.; Zhu, X.X. Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2018**, *57*, 924–935. [CrossRef]
38. Lyu, H.; Lu, H.; Mou, L. Learning a transferable change rule from a recurrent neural network for land cover change detection. *Remote Sens.* **2016**, *8*, 506. [CrossRef]
39. Rußwurm, M.; Körner, M. Multi-temporal land cover classification with sequential recurrent encoders. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 129. [CrossRef]
40. Wang, H.; Zhao, X.; Zhang, X.; Wu, D.; Du, X. Long time series land cover classification in China from 1982 to 2015 based on Bi-LSTM deep learning. *Remote Sens.* **2019**, *11*, 1639. [CrossRef]
41. Ienco, D.; Gaetano, R.; Dupaquier, C.; Maurel, P. Land cover classification via multitemporal spatial data by deep recurrent neural networks. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1685–1689. [CrossRef]
42. Reddy, D.S.; Prasad, P.R.C. Prediction of vegetation dynamics using NDVI time series data and LSTM. *Model. Earth Syst. Environ.* **2018**, *4*, 409–419. [CrossRef]
43. Gharakhanlou, N.M.; Perez, L. From data to harvest: Leveraging ensemble machine learning for enhanced crop yield predictions across Canada amidst climate change. *Sci. Total Environ.* **2024**, *951*, 175764.
44. Veloso, A.; Mermoz, S.; Bouvet, A.; Le Toan, T.; Planells, M.; Dejoux, J.-F.; Ceschia, E. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. *Remote Sens. Environ.* **2017**, *199*, 415–426. [CrossRef]
45. Ghorbanian, A.; Kakooei, M.; Amani, M.; Mahdavi, S.; Mohammadzadeh, A.; Hasanlou, M. Improved land cover map of Iran using Sentinel imagery within Google Earth Engine and a novel automatic workflow for land cover classification using migrated training samples. *ISPRS J. Photogramm. Remote Sens.* **2020**, *167*, 276–288. [CrossRef]
46. Ismail Fawaz, H.; Forestier, G.; Weber, J.; Idoumghar, L.; Muller, P.-A. Deep learning for time series classification: A review. *Data Min. Knowl. Discov.* **2019**, *33*, 917–963. [CrossRef]
47. Crisóstomo de Castro Filho, H.; Abílio de Carvalho Júnior, O.; Ferreira de Carvalho, O.L.; Pozzobon de Bem, P.; dos Santos de Moura, R.; Olino de Albuquerque, A.; Rosa Silva, C.; Guimaraes Ferreira, P.H.; Fontes Guimarães, R.; Trancoso Gomes, R.A. Rice crop detection using LSTM, Bi-LSTM, and machine learning models from Sentinel-1 time series. *Remote Sens.* **2020**, *12*, 2655. [CrossRef]
48. Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J. Recent advances in convolutional neural networks. *Pattern Recognit.* **2018**, *77*, 354–377. [CrossRef]
49. Greff, K.; Srivastava, R.K.; Koutník, J.; Steunebrink, B.R.; Schmidhuber, J. LSTM: A search space odyssey. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *28*, 2222–2232. [CrossRef]
50. Siami-Namini, S.; Tavakoli, N.; Namin, A.S. A comparison of ARIMA and LSTM in forecasting time series. In Proceedings of the 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 17–20 December 2018.
51. Armstrong, R.A.; Eperjesi, F.; Gilmartin, B. The application of analysis of variance (ANOVA) to different experimental designs in optometry. *Ophthalmic Physiol. Opt.* **2002**, *22*, 248–256. [CrossRef] [PubMed]
52. Jia, K.; Wu, B.; Li, Q. Crop classification using HJ satellite multispectral data in the North China Plain. *J. Appl. Remote Sens.* **2013**, *7*, 073576. [CrossRef]
53. Yang, C.; Everitt, J.H.; Murden, D. Evaluating high resolution SPOT 5 satellite imagery for crop identification. *Comput. Electron. Agric.* **2011**, *75*, 347–354. [CrossRef]
54. Chang, J.; Hansen, M.C.; Pittman, K.; Carroll, M.; DiMiceli, C. Corn and soybean mapping in the United States using MODIS time-series data sets. *Agron. J.* **2007**, *99*, 1654–1664. [CrossRef]
55. Gómez, C.; White, J.C.; Wulder, M.A. Optical remotely sensed time series data for land cover classification: A review. *ISPRS J. Photogramm. Remote Sens.* **2016**, *116*, 55–72. [CrossRef]
56. Long, J.A.; Lawrence, R.L.; Greenwood, M.C.; Marshall, L.; Miller, P.R. Object-oriented crop classification using multitemporal ETM+ SLC-off imagery and random forest. *GIScience Remote Sens.* **2013**, *50*, 418–436. [CrossRef]

57. Müller, H.; Rufin, P.; Griffiths, P.; Siqueira, A.J.B.; Hostert, P. Mining dense Landsat time series for separating cropland and pasture in a heterogeneous Brazilian savanna landscape. *Remote Sens. Environ.* **2015**, *156*, 490–499. [CrossRef]
58. Novelli, A.; Aguilar, M.A.; Nemmaoui, A.; Aguilar, F.J.; Tarantino, E. Performance evaluation of object based greenhouse detection from Sentinel-2 MSI and Landsat 8 OLI data: A case study from Almería (Spain). *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 403–411. [CrossRef]
59. Castillejo-González, I.L.; López-Granados, F.; García-Ferrer, A.; Peña-Barragán, J.M.; Jurado-Expósito, M.; de la Orden, M.S.; González-Audicana, M. Object-and pixel-based analysis for mapping crops and their agro-environmental associated measures using QuickBird imagery. *Comput. Electron. Agric.* **2009**, *68*, 207–215. [CrossRef]
60. Matton, N.; Sepulcre Canto, G.; Waldner, F.; Valero, S.; Morin, D.; Inglada, J.; Arias, M.; Bontemps, S.; Koetz, B.; Defourny, P. An automated method for annual cropland mapping along the season for various globally-distributed agrosystems using high spatial and temporal resolution time series. *Remote Sens.* **2015**, *7*, 13208–13232. [CrossRef]
61. Lebourgeois, V.; Dupuy, S.; Vintrou, É.; Ameline, M.; Butler, S.; Bégué, A. A combined random forest and OBIA classification scheme for mapping smallholder agriculture at different nomenclature levels using multisource data (simulated Sentinel-2 time series, VHRS and DEM). *Remote Sens.* **2017**, *9*, 259. [CrossRef]
62. Rußwurm, M.; Korner, M. Temporal vegetation modelling using long short-term memory networks for crop identification from medium-resolution multi-spectral satellite images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 22–25 July 2017.
63. Gafurov, A.; Mukharamova, S.; Saveliev, A.; Yermolaev, O. Advancing Agricultural Crop Recognition: The Application of LSTM Networks and Spatial Generalization in Satellite Data Analysis. *Agriculture* **2023**, *13*, 1672. [CrossRef]
64. Zhang, X.; Cai, Z.; Hu, Q.; Yang, J.; Wei, H.; You, L.; Xu, B. Improving crop type mapping by integrating LSTM with temporal random masking and pixel-set spatial information. *ISPRS J. Photogramm. Remote Sens.* **2024**, *218*, 87–101. [CrossRef]
65. van Duynhoven, A.; Dragićević, S. Analyzing the effects of temporal resolution and classification confidence for modeling land cover change with long short-term memory networks. *Remote Sens.* **2019**, *11*, 2784. [CrossRef]
66. Lu, J.; Ren, K.; Li, X.; Zhao, Y.; Xu, Z.; Ren, X. From reanalysis to satellite observations: Gap-filling with imbalanced learning. *GeoInformatica* **2022**, *26*, 397–428. [CrossRef]
67. Kang, M.; Liu, Y.; Wang, M.; Li, L.; Weng, M. A random forest classifier with cost-sensitive learning to extract urban landmarks from an imbalanced dataset. *Int. J. Geogr. Inf. Sci.* **2022**, *36*, 496–513. [CrossRef]
68. Zhang, H.; Song, Y.; Xu, S.; He, Y.; Li, Z.; Yu, X.; Liang, Y.; Wu, W.; Wang, Y. Combining a class-weighted algorithm and machine learning models in landslide susceptibility mapping: A case study of Wanzhou section of the Three Gorges Reservoir, China. *Comput. Geosci.* **2022**, *158*, 104966. [CrossRef]
69. Van den Broeck, W.A.; Goedemé, T.; Loopmans, M. Multiclass land cover mapping from historical orthophotos using domain adaptation and spatio-temporal transfer learning. *Remote Sens.* **2022**, *14*, 5911. [CrossRef]
70. Zhang, X.; Zhou, Y.n.; Luo, J. Deep learning for processing and analysis of remote sensing big data: A technical review. *Big Earth Data* **2022**, *6*, 527–560. [CrossRef]
71. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
72. Atluri, G.; Karpatne, A.; Kumar, V. Spatio-temporal data mining: A survey of problems and methods. *ACM Comput. Surv.* **2018**, *51*, 1–41. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.