



Article

A Dual Attention Convolutional Neural Network for Crop Classification Using Time-Series Sentinel-2 Imagery

Seydi Teymoor Seydi ¹, Meisam Amani ^{2,*} and Arsalan Ghorbanian ³¹ School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran 14399-57131, Iran; seydi.teymoor@ut.ac.ir² Wood Environment & Infrastructure Solutions, Ottawa, ON K2E 7L5, Canada³ Department of Photogrammetry and Remote Sensing, Faculty of Geodesy and Geomatics Engineering, K. N. Toosi University of Technology, Tehran 19967-15433, Iran; a.ghorbanian@email.kntu.ac.ir

* Correspondence: meisam.amani@woodplc.com

Abstract: Accurate and timely mapping of crop types and having reliable information about the cultivation pattern/area play a key role in various applications, including food security and sustainable agriculture management. Remote sensing (RS) has extensively been employed for crop type classification. However, accurate mapping of crop types and extents is still a challenge, especially using traditional machine learning methods. Therefore, in this study, a novel framework based on a deep convolutional neural network (CNN) and a dual attention module (DAM) and using Sentinel-2 time-series datasets was proposed to classify crops. A new DAM was implemented to extract informative deep features by taking advantage of both spectral and spatial characteristics of Sentinel-2 datasets. The spectral and spatial attention modules (AMs) were respectively applied to investigate the behavior of crops during the growing season and their neighborhood properties (e.g., textural characteristics and spatial relation to surrounding crops). The proposed network contained two streams: (1) convolution blocks for deep feature extraction and (2) several DAMs, which were employed after each convolution block. The first stream included three multi-scale residual convolution blocks, where the spectral attention blocks were mainly applied to extract deep spectral features. The second stream was built using four multi-scale convolution blocks with a spatial AM. In this study, over 200,000 samples from six different crop types (i.e., alfalfa, broad bean, wheat, barley, canola, and garden) and three non-crop classes (i.e., built-up, barren, and water) were collected to train and validate the proposed framework. The results demonstrated that the proposed method achieved high overall accuracy and a Kappa coefficient of 98.54% and 0.981, respectively. It also outperformed other state-of-the-art classification methods, including RF, XGBOOST, R-CNN, 2D-CNN, 3D-CNN, and CBAM, indicating its high potential to discriminate different crop types.

Keywords: crop mapping; deep learning; convolutional neural networks (CNN); attention modules (AM); dual attention CNN; Sentinel-2; multi-temporal



Citation: Seydi, S.T.; Amani, M.; Ghorbanian, A. A Dual Attention Convolutional Neural Network for Crop Classification Using Time-Series Sentinel-2 Imagery. *Remote Sens.* **2022**, *14*, 498. <https://doi.org/10.3390/rs14030498>

Academic Editors: Iman Tahmasbian, Kourosh Khoshelham and Shahla Hosseini Bai

Received: 21 December 2021

Accepted: 19 January 2022

Published: 21 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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

Considering the prospect of human population growth, which is expected to reach 8.7 billion by 2030, the food supply system is subjected to escalating pressure [1,2]. Additionally, climate change effects and catastrophic natural disasters (e.g., drought and flood) are already hampering agricultural production and threatening food security from local to global scales [3,4]. Accordingly, it is vital to obtain authentic information about the location, extent, type, health, and yield of crops to ensure food security, poverty reduction, and water resource management [5]. Additionally, it is more appealing to incorporate efficient approaches that facilitate the requirement of sustainability and climate change adaption [6,7]. Thus, it is crucial to employ efficient approaches, such as advanced machine learning along with remote sensing (RS) data, to ensure high-quality information is derived about crops in order to achieve specified goals [8].

RS has long been recognized as a trustworthy approach for extracting specialized information about agricultural products [9–12]. This is owing to the frequent, broad-scale, and spatially consistent data acquisition of RS systems. In particular, RS allows timely monitoring of croplands to extract different information concerning the crop phenological status [13,14], health [15], types [16–18], and yield estimation [19,20] over small- to large-scale areas, based on different characteristics of satellite images (e.g., spatial, temporal, and spectral resolutions). These practices have been performed using different sources of RS data, including multi-spectral [21,22], synthetic aperture radar (SAR) [23,24], light detection and ranging (LiDAR) [25,26], hyperspectral [27,28], thermal [29,30], and digital elevation model (DEM) [31,32].

Along with the advancement of RS, image processing techniques and machine learning algorithms have also been significantly promoted [33–35]. Accordingly, machine learning algorithms offer the potential to exploit the information content of RS data through automated frameworks [36,37]. In this regard, many scholars incorporated RS data and machine learning algorithms for crops mapping and monitoring. For instance, Zhang et al. [38] implemented a random forest (RF) algorithm to classify croplands in China and Canada. To this end, textural features and vegetation indices were extracted from RapidEye images and were added to the spectral bands. The results revealed that the integration of all spectral, textural, and vegetation indices features could considerably enhance the classification results. Additionally, Mandal et al. [39] employed a support vector machine (SVM) classifier along with time-series RADARSAT-2 C-band quad-pol data to discriminate different crops in Vijayawada, India. Temporal signatures of backscattering intensities were initially derived, and then informative features were selected by adopting kernel principal component analysis (PCA). It was reported that selecting discriminative temporal features improved the classification results by 7%. Moreover, Ma ponya et al. [40] evaluated the potential of five machine learning algorithms, including SVM, RF, decision tree, K nearest neighbour, and maximum likelihood, for crop mapping using multi-temporal Sentinel-2 images. Four different scenarios were considered for the classification tasks, and the results indicated the superiority of RF and SVM over other classifiers, especially when a subset of hand-selected images (i.e., knowledge-based) were utilized. Furthermore, Saini and Ghosh [41] identified the major crop types in Roorkee, India, using four different machine learning algorithms and Sentinel-2 images. It was observed that extreme gradient boosting (XGBOOST) outperformed other classifiers with an overall accuracy of 87%.

Currently, deep learning algorithms are recognized as a breakthrough approach for processing RS data [35,42]. In particular, classification studies using RS data greatly benefit from deep learning approaches because of their flexibility in feature representation, automation through the end-to-end procedure, and automatic extraction features [43–45]. In this regard, different deep learning models (i.e., different structures and networks) have been employed for crop type mapping and monitoring [46–51]. For example, Zhao, Duan, Liu, Sun, and Reymondin [51] compared five deep learning models for crop mapping based on dense time-series Sentinel-2 images. Their results suggested the high capabilities of one-dimensional convolutional neural networks (1D-CNN), long short-term memory CNN (LSTM-CNN), and gated recurrent unit CNN (GRU-CNN) models for crop mapping, respectively. Furthermore, Ji, Zhang, Xu, Shi, and Duan [47] developed a three-dimensional CNN (3D-CNN) model to automatically classify crops using spatio-temporal RS images. The proposed network was enhanced using an active learning strategy to increase the labelling accuracy. The results were compared to a two-dimensional CNN (2D-CNN) classifier, suggesting higher efficiency and accuracy of their proposed approach.

Similar to other countries, RS data have widely been utilized for crop type mapping in Iran. For instance, Akbari et al. [52] implemented the particle swarm optimization algorithm to select informative features from time-series Sentinel-2 images for crop mapping in Ghale-Nou, Tehran, Iran. The selected features were ingested into an RF classifier, and the results showed the high potential of the proposed method for heterogenous crop fields classification. Asgarian et al. [53] also investigated the potential of time-series

Landsat-8 images for crop mapping in fragmented and heterogeneous landscapes of Najaf-Abad, Iran. To this end, long-term in-situ phenological information was combined with satellite images to map annual crop types using intensive decision trees and SVM classifiers. Furthermore, Saadat, et al. [54] employed time-series Sentinel-1 data to map rice in the northern part of Iran. To this end, Gamma Nought, Sigma Nought, and Beta Nought features of Sentinel-1 images in three scenarios were used in the RF classifier. Their results indicated the superiority of Sigma Nought and Gamma Nought Sentinel-1 data in vertical transmittance and horizontal receiving (VH) polarization.

Although many crop mapping frameworks have been proposed by various researchers, they generally have one of the following disadvantages:

- (I) Most crop mapping studies have focused on conventional machine learning methods (e.g., RF and SVM). These algorithms do not usually provide the highest possible accuracies due to several factors, such as climatic conditions and the fluctuations in planting times.
- (II) Many studies have only used spectral-temporal information for crop mapping. However, spatial information should be included in the classification algorithm to produce highly accurate maps.
- (III) Many state-of-the-art deep learning methods for crop mapping have only used the 2D/3D convolution blocks for extracting deep features. All of these extracted deep features are not informative for crop mapping and provide redundant information. In this regard, attention blocks should be implemented to select the most informative features.

The Iranian crop system is under escalating pressure mainly due to the severe water crisis and population growth [55]. Additionally, climate change and the current dramatic drought condition in Iran also exacerbate the existing pressure [56]. Furthermore, the current economic and political sanctions have become a notable issue that would amplify this pressure in Iran [57,58]. Consequently, incorporation of advanced technologies, such as remote sensing and machine/deep learning algorithms, is required to support efficient agricultural practices in Iran. Considering the importance of crop mapping in Iran, a novel deep learning algorithm was developed in this study for accurate crop classification. The classification model has three main steps: (1) data preparation, (2) deep feature extraction based on multi-scale residual kernel convolution and CNN's parameters optimization, and (3) crop type mapping based on an optimized model. The key contributions of this research are as follows:

- (I) Proposing a novel framework for mapping crop types based on a two-stream CNN with a DAM.
- (II) Introducing novel spatial and spectral attention mechanisms (AMs) to extract informative deep features for crop mapping.
- (III) Utilizing multi-scale and residual blocks for increasing the accuracy of the proposed network.
- (IV) Evaluating the sensitivity of the proposed method during the growing season of crops based on a time-series normalized difference vegetation index (NDVI).
- (V) Evaluating the performance of commonly used machine learning and deep learning methods for crop type mapping.

2. Study Area and Datasets

2.1. Study Area

The study area was an agricultural area in the southern portions of the Aq Qala counties, Golestan province. This study area is approximately centered at a latitude and longitude of 37°50' N and 54°40' E, respectively (see Figure 1). The climate of the study area is mainly influenced by the Alborz Mountains and the Caspian Sea. Thus, it has different climates with a diverse rate of precipitation and humidity [59]. For instance, the study area contains semi-arid (northern parts) and humid (southern parts) climates with annual precipitation between 249 and 529 mm [60]. Consequently, it includes both irrigated and

rainfed agricultural systems. The study area is among the most important counties for crop production in the province of Golestan, and various crops (e.g., wheat, alfalfa, and barleys) are cultivated in this region during each growing season, of which wheat is the dominant one. As one of the biggest crop production sources in Golestan, it is essential to establish regular and accurate crop condition monitoring systems and estimate the cultivated crop area with high reliability.

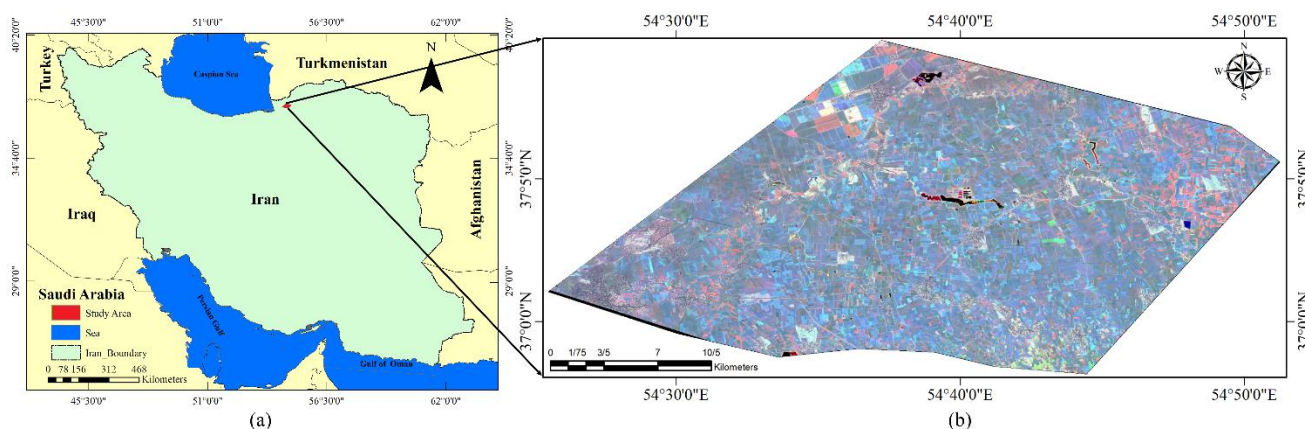


Figure 1. (a) The geographical location of the study area, and (b) a false-color composite NDVI image (R: first month NDVI, G: second month NDVI, and B: third month NDVI) from the study area.

2.2. Sentinel-2 Imagery

In this study, time-series Sentinel-2 optical satellite images were employed for crop type classification. Sentinel-2 is a European satellite developed via the cooperation of the European Commission initiative Copernicus and the European Space Agency [33]. This platform carries the MultiSpectral Instrument (MSI) sensor, a wide-swath multispectral imager that images the Earth's surface using 13 bands with a spectral range from 443 nm to 2190 nm. These bands are taken from visible to shortwave infrared domains of the electromagnetic spectrum in three different spatial resolutions (i.e., 10–60 m) [61]. Sentinel-2 constellation (Sentinel-2A and B) provides global coverage of the Earth's surface every five days, making it suitable for a variety of land monitoring tasks. In total, 13 Sentinel-2 images were used in this study (see Table 1). As is clear from Table 1, the imagery acquired in the first two weeks of February 2018 and the second two weeks of March 2018 were not used because of the cloud cover over the study area on these two dates. Overall, we could effectively distinguish various types of crops in the study area using these time-series images [62,63].

Table 1. Date and description of Sentinel-2 multispectral images that were used for crop mapping.

Data	Date	Description
Dataset-Time-1	November 2017	The first two weeks
Dataset-Time-2	November 2017	The second two weeks
Dataset-Time-3	December 2017	The first two weeks
Dataset-Time-4	December 2017	The second two weeks
Dataset-Time-5	January 2018	The first two weeks
Dataset-Time-6	January 2018	The second two weeks
Dataset-Time-7	February 2018	The first two weeks, high-cloudy, not used
Dataset-Time-8	February 2018	The second two weeks
Dataset-Time-9	March 2018	The first two weeks
Dataset-Time-10	March 2018	The second two weeks, high-cloudy, not used
Dataset-Time-11	April 2018	The first two weeks
Dataset-Time-12	April 2018	The second two weeks
Dataset-Time-13	May 2018	The first two weeks
Dataset-Time-14	May 2018	The second two weeks
Dataset-Time-15	June 2018	The first two weeks

2.3. Reference Samples

Figure 2 illustrates the distribution of the collected in-situ samples over the study area. These samples were collected from ten classes during several field surveys. The field data were collected in 2018 from April to May for all crop classes. A handheld global positioning system (GPS) with a positional accuracy of <5 m was used to record the locations of the samples.

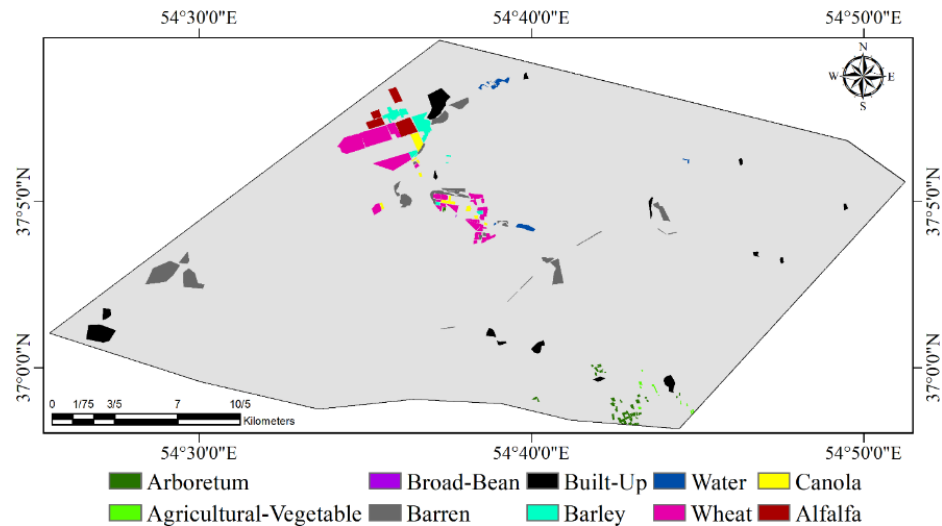


Figure 2. The distribution of the reference samples from ten classes collected over the study areas.

As is clear from Figure 2, more arboretum and agricultural-vegetable areas are located on the right side of the study area while other crops are dispersed over the study area.

Table 2 provides the number of samples for each class. The wheat and broad bean classes had the maximum and minimum numbers of reference samples. There are different approaches, such as manual splitting, random splitting, and non-random splitting for the division of reference samples into training, validation, and test samples [64]. In this regard, random sampling is the most common way to split reference samples, which has extensively been used for classification tasks using remote sensing images [65–67]. Accordingly, in this study, random sampling was employed to divide reference samples into training (3%), validation (0.1%), and test (96.9%) samples.

Table 2. The number of reference samples that were divided into training, validation, and test samples.

ID	Crop Type	All Samples	Training (3%)	Validation (0.1%)	Test (96.9%)
1	Arboretum	9336	306	67	8963
2	Agricultural-Vegetable	1618	53	12	1553
3	Broad Bean	71	3	1	67
4	Barren	58,604	1922	422	56,260
5	Built-Up	43,252	1419	311	41,522
6	Barley	17,363	569	125	16,669
7	Water	5813	191	42	5580
8	Wheat	58,701	1925	423	56,353
9	Canola	8282	271	60	7951
10	Alfalfa	17,995	590	130	17,275
Total		221,035	7249	1593	212,193

3. Method

The general framework of crop type classification based on the proposed method is illustrated in Figure 3. The proposed classification framework was implemented in

three main steps: (1) data preparation and normalized difference vegetation index (NDVI) calculation, (2) model training and parameters tuning, and (3) prediction and accuracy assessment. The detail of each step is discussed in the following subsections.

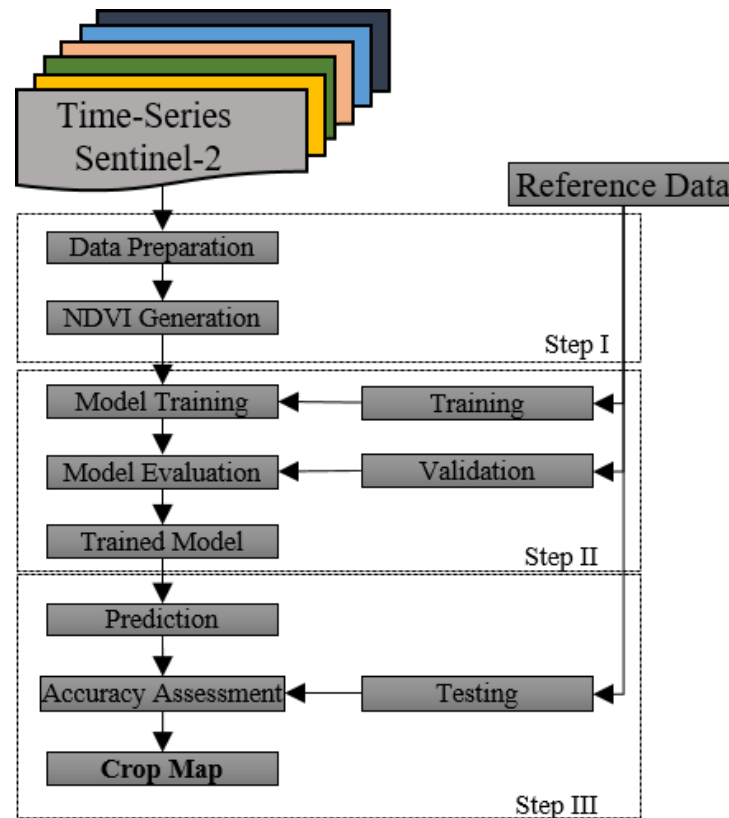


Figure 3. Overview of the proposed framework for crop mapping.

3.1. Data Preparation and Time-Series NDVI Calculation

Sentinel-2 datasets require several preprocessing steps, such as cloud masking and atmospheric correction. In this regard, we selected only non-cloudy images for the analysis. Moreover, the atmospheric correction was implemented using the Sen2cor module [68], which is available in the SNAP software.

Spectral feature extraction is the most common step in RS classification tasks [61]. The feature extraction can be conducted in two main categories: (1) combining spectral bands using simple mathematical operations, such as the spectral indices of NDVI [69,70]; and (2) deriving high order statistical features (i.e., covariance and correlation), such as PCA [71] and factor analysis (FA). Among different spectral indices, NDVI was selected due to its simplicity and its high applicability for crop mapping [72–75]. NDVI was computed based on the red (0.665 μm) and near-infrared (NIR) (0.842 μm) bands (see Equation (1)).

$$\text{NDVI} = \frac{(\text{NIR} - \text{Red})}{(\text{NIR} + \text{Red})} \quad (1)$$

Crops have a dynamic nature because of their growth during their lifetime. Thus, employing time-series datasets is an effective and pertinent solution for mapping crops [76,77]. Consequently, the time-series NDVI features were utilized in this study for crop types classification.

3.2. Proposed Deep Learning Architecture

This study proposed a new dual-stream CNN architecture with both spectral and spatial attention blocks. According to the presented architecture in Figure 4, the proposed

method received input patches of $11 \times 11 \times 13$, and then the patches were fed into two separate streams for deep feature extraction.

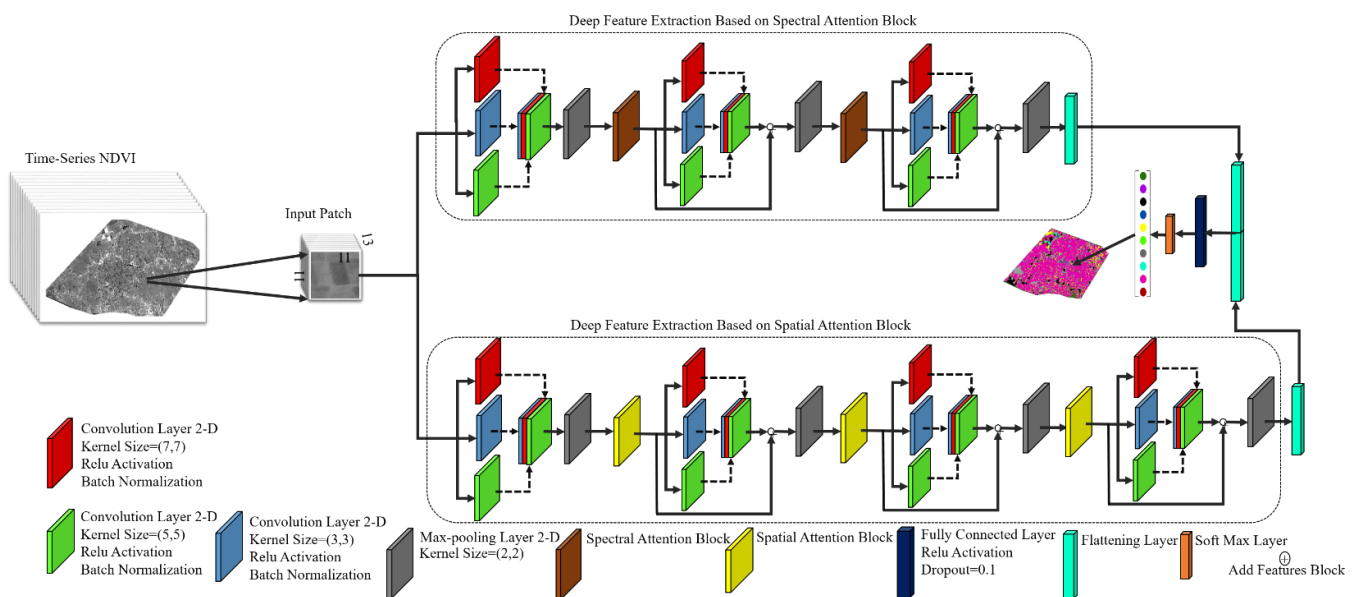


Figure 4. The proposed dual-stream spatial and spectral attention blocks framework for crop mapping.

The first stream explored deep features based on multi-scale residual convolution blocks and spectral attention blocks. This stream focused on deep spectral feature extraction based on spectral AM. In this regard, a shallow multi-layer feature extractor, max-pooling layer, spectral attention blocks, and multi-scale residual blocks were employed. First, the shallow deep features were extracted via a multi-scale convolution block. Then, the spectral attention block was employed to investigate the inter-channel relationship of feature maps. Subsequently, the max-pooling layer was applied to reduce the size of the generated feature maps. The multi-scale residual block was then employed to find more meaningful features. Similarly, the spectral attention block and max-pooling were employed. Finally, the extracted deep features were transferred to the latest multi-scale residual and spectral attention blocks to generate high-level deep features.

The second stream investigated deep features while concentrating on deep spatial features using spatial attention blocks. Similarly, this stream had one multi-scale convolution block and three multi-scale residual blocks. Moreover, after convolution block layers, the spectral attention block and max-pooling layers were employed.

After deep feature extraction based on multi-scale residual blocks and attention blocks, the deep features were flattened using a flattening layer. Then, they were fed to a dense layer, and the decision was made via a soft-max layer.

The main differences between the proposed architecture and other CNN frameworks are:

- (1) Utilizing a double streams framework for investigating spatial/spectral deep feature extraction.
- (2) Proposing a novel AM framework for extraction of informative deep features that have a higher efficiency compared to the convolutional block attention module (CBAM).
- (3) Taking advantage of residual, depth-wise, and separable convolution blocks as well as combining them for deep feature extraction.
- (4) Employing separable (point/depth-wise convolution layers) convolution which has a better performance.

3.2.1. Attention Mechanism (AM)

The AM in deep learning was inspired by the psychological attention mechanisms within the human brain [78–82]. The main idea behind the AM is to direct the focus of the network on extracting meaningful features instead of non-essential features [81]. The efficiency of the AM in deep learning models has been proven in previous literature [78,82–85]. In this regard, this study proposed a novel AM to increase the efficiency of the implemented/developed architecture by considering both spectral and spatial AM. The main idea to incorporate the AM was to explore the relationship between spectral-temporal and spatial-temporal information of input patches for the crop type classification task.

The developed spectral AM concentrated on ‘what’ is meaningful in the given input feature map [83,84,86]. To this end, we introduced a spectral attention block in accordance with the architecture illustrated in Figure 5. Based on this, the input feature map was fed into a convolution block with a kernel size (a,b) that was equal to the length and width of the input feature data. The size of the output feature map was $1 \times 1 \times c$. Moreover, the number of filters was c, which was equal to the number of feature maps of the input data. After reshaping the output of the previous layer, the features were transferred into a multi-layer perceptron (MLP) layer with two dense layers with different neuron sizes. The first and second layers reduced the number of neurons based on the reduction rate and reconstructed the features, respectively. Simultaneously, the separable convolution layer was employed for input data before the multiplication of features with the input feature map. Finally, the output of the first stream and separable convolution layer was fused using multiplication. The separable convolution layer was implemented in two steps: point-wise convolution and depth-wise convolution on the output of the point-wise convolution.

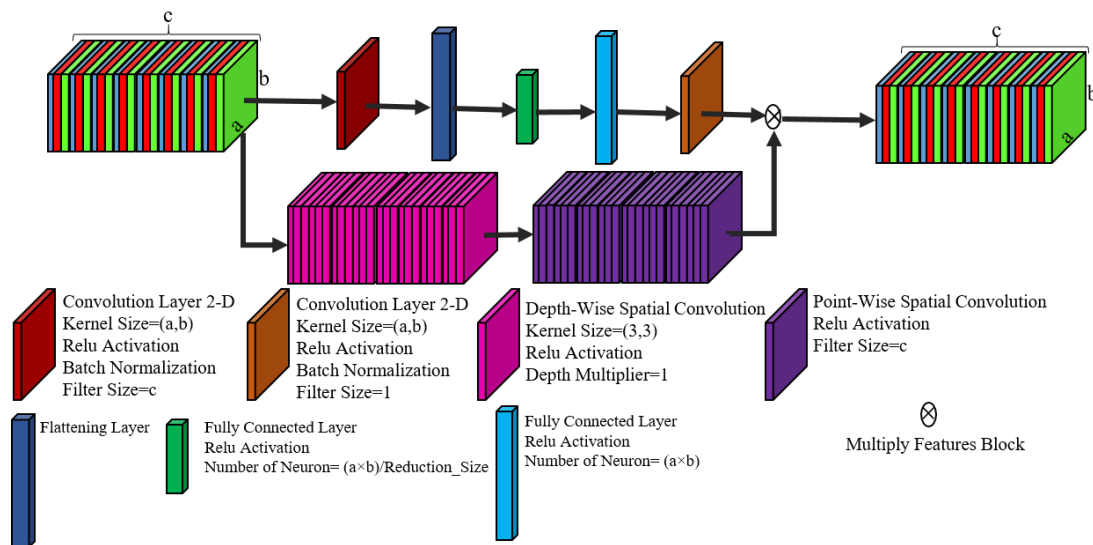


Figure 5. The proposed spectral attention block.

The developed spatial AM considered the inter-spatial relationship of feature maps [84,87,88]. The spatial AM concentrated on ‘where’ a useful region within the input feature map is [86,89]. This AM was implemented similarly to the spectral AM, but with different output sizes of convolution layers (see Figure 6). Based on this, the input feature map was transferred into a convolution block with a kernel size (a,b) with only one kernel convolution and padding. This means that the output size of the feature map was $a \times b \times 1$. After reshaping the output of the previous layer, the features were fed into an MLP with two fully connected layers with different neuron sizes. The first layer reduced the number of neurons based on the reduction rate. Then, the second fully connected layer reconstructed the features. Simultaneously, the separable convolution layer was employed for input data before the multiplication of features with the input feature map. Finally, the output of the first stream and separable convolution layer was fused via multiplication.

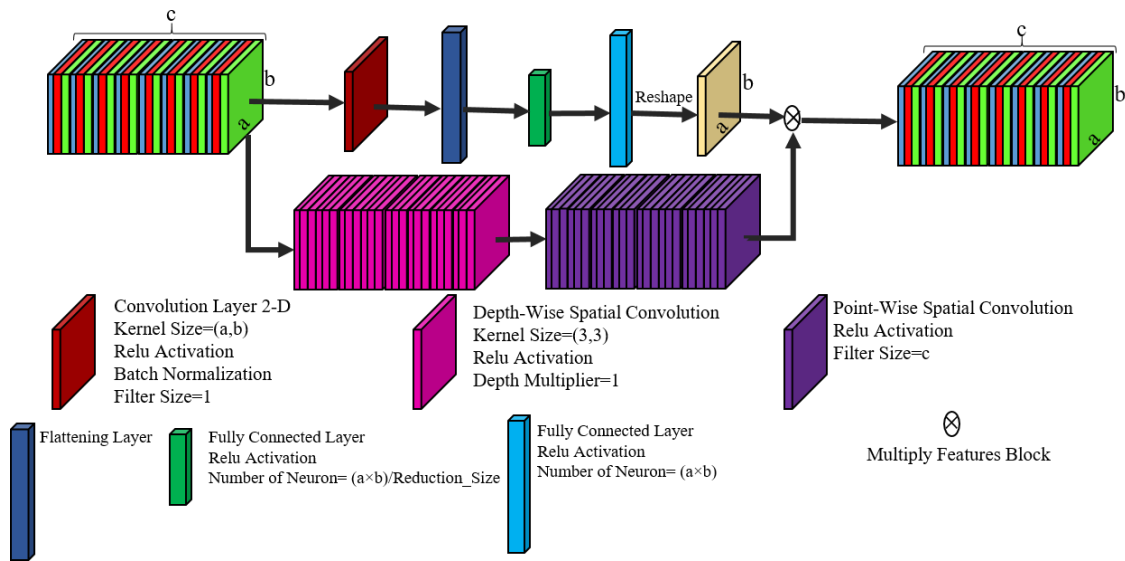


Figure 6. The proposed spatial attention block.

3.2.2. Convolution Layer

Convolution layers are the central core of CNN frameworks, and the main task of these layers is extracting high-level deep features from input imagery [90]. The convolution layers automatically explore spatial and spectral features at the same time. The basic computation of the convolutional layer can be defined as follows (Equation (2)) [91].

$$f^{\mathcal{N}} = \phi(w^l f^{\mathcal{N}-1}) + b^{\mathcal{N}} \quad (2)$$

where x is the input data from layer $\mathcal{N} - 1$, ϕ is an activation function, w and b are the weighted template and bias vector, respectively.

The output of j th feature map for a 2D convolution at the spatial location of (x, y) can be computed using Equation (3) [35].

$$f_{\mathcal{N},j}^{xy} = \phi \left(b_{\mathcal{N},j} + \sum_m \sum_{r=0}^{R_{\mathcal{N}}-1} \sum_{s=0}^{S_{\mathcal{N}}-1} W_{\mathcal{N},j}^{r,s} f_{\mathcal{N}-1,m}^{(x+r)(y+s)} \right) \quad (3)$$

where m is the feature cube connected to the current feature cube in the $(\mathcal{N} - 1)$ th layer, and R and S are the length and width of the filter, respectively.

This research took advantage of both residual and multi-scale blocks. The multi-scale blocks increase the efficiency of the network against the differences in the scale of objects [35]. Moreover, the residual blocks improved the efficiency of the network and helped to prevent gradient vanishing.

3.3. Model Training

Since the unknown parameters of the deep learning architecture cannot be calculated through an analytical solution, the iterative framework was employed to optimize the model parameters [90]. The adaptive moment estimation (Adam) optimizer [80] was used in this study to optimize the model parameters. Furthermore, the cross-entropy (CE) loss function was utilized to calculate the error of the network during the training phase. The training phase was conducted based on the training samples, and then the loss value of the trained model was computed using validation samples. The CE loss function can be calculated using Equation (4):

$$\text{CE-loss} = - \sum_{i=1}^N \Phi_i \log \phi_i \quad (4)$$

where Φ and φ are the true and predicted labels, respectively. Moreover, N refers to the number of classes.

3.4. Accuracy Assessment

The statistical accuracy assessment was performed using independent test samples. The six most common statistical criteria, all extracted from the confusion matrix of the classification, were utilized to evaluate classification results. These criteria were overall accuracy (OA), user accuracy (UA), producer accuracy (PA), Kappa coefficient (KC), omission error (OE), and commission error (CE).

3.5. Comparison with Other Classification Methods

Crop mapping has widely been applied by machine learning and deep learning-based methods [92,93]. The RF and XGBOOST are the most common machine learning methods that have widely been used in many crop mapping applications based on time-series datasets [93,94]. This research implemented these two machine learning-based methods to evaluate their efficiency in comparison with deep learning-based methods. Thus, six different classifiers, including two commonly used machine learning algorithms (i.e., RF [94] and XGBOOST [95]) and four deep learning models (i.e., recurrent-convolutional neural network (R-CNN) [49], 2D-CNN [47], 3D-CNN [47], and convolutional block attention module (CBAM)), were also implemented to produce a more comprehensive evaluation of the performance of the proposed model. R-CNN, developed by Mazzia, Khaliq and Chiaberge [49], combines the LSTM cells and 2D convolution layers for crop mapping based on time-series datasets. Moreover, CBAM [82] combines channel attention and spatial attention after each convolution layer, wherein the channel attention block is employed after the spatial attention block. The inputs of the RF and XGBOOST algorithms were spectral-temporal features with the size of 1×13 where 1 and 13 refer to the spectral (i.e., NDVI) and temporal information. Moreover, the input datasets of the deep learning-based methods were spatial-spectral-temporal information with the size of $11 \times 11 \times 1 \times 13$, where 11×11 was the width and length of spatial information, respectively, 1 was the spectral information (i.e., NDVI), and 13 was the temporal information. It is worth noting that the size of the spatial information for the deep learning-based methods was determined by trial and error. The patch data were also generated by moving a window with the size of 11×11 . The label of this patch corresponded to the central pixel of the patch.

4. Experiments and Results

4.1. Parameter Setting

The proposed method and other classifiers have several parameters that need to be set. As described in the Method Section, the optimum values of the parameters for each classifier were determined based on several trial and error attempts (see Table 3). All parameters of the deep learning-based methods were set identically. It is worth noting that the selection of some of these parameters depended on the processing system.

Table 3. The optimum values of the classifier parameters.

Data	Description
RF	number of estimators = 105, number of features to split each node = 3
XGBoost	number of rounds = 500, max. depth = 5, subsample = 1, min. child weight = 1, lambda = 1, colsample bytree = 0.8.
Deep Learning Models	dropout rate = 0.1, epochs = 500, initial learning = 10^{-4} , mini-batch size = 550, weight initializer = He normal

4.2. Classification Results

The results of crop mapping based on the proposed deep learning method along with other algorithms are illustrated in Figure 7. A high-resolution image from the study area is also provided in Figure 7a for comparison purposes. The results showed that the map produced using XGBOOST (Figure 7b) included salt and pepper errors. Furthermore, the RF classifier (Figure 7c) could not delineate different classes with a high level of accuracy.

In general, deep learning methods (Figure 7d–h) produced better results compared to the XGBOOST and RF models. However, there were still several wrongly classified pixels in the results of the deep learning methods, especially those of the R-CNN and 2D-CNN methods. Overall, the proposed method (Figure 7h) provided the most accurate crop map based on visual interpretation.

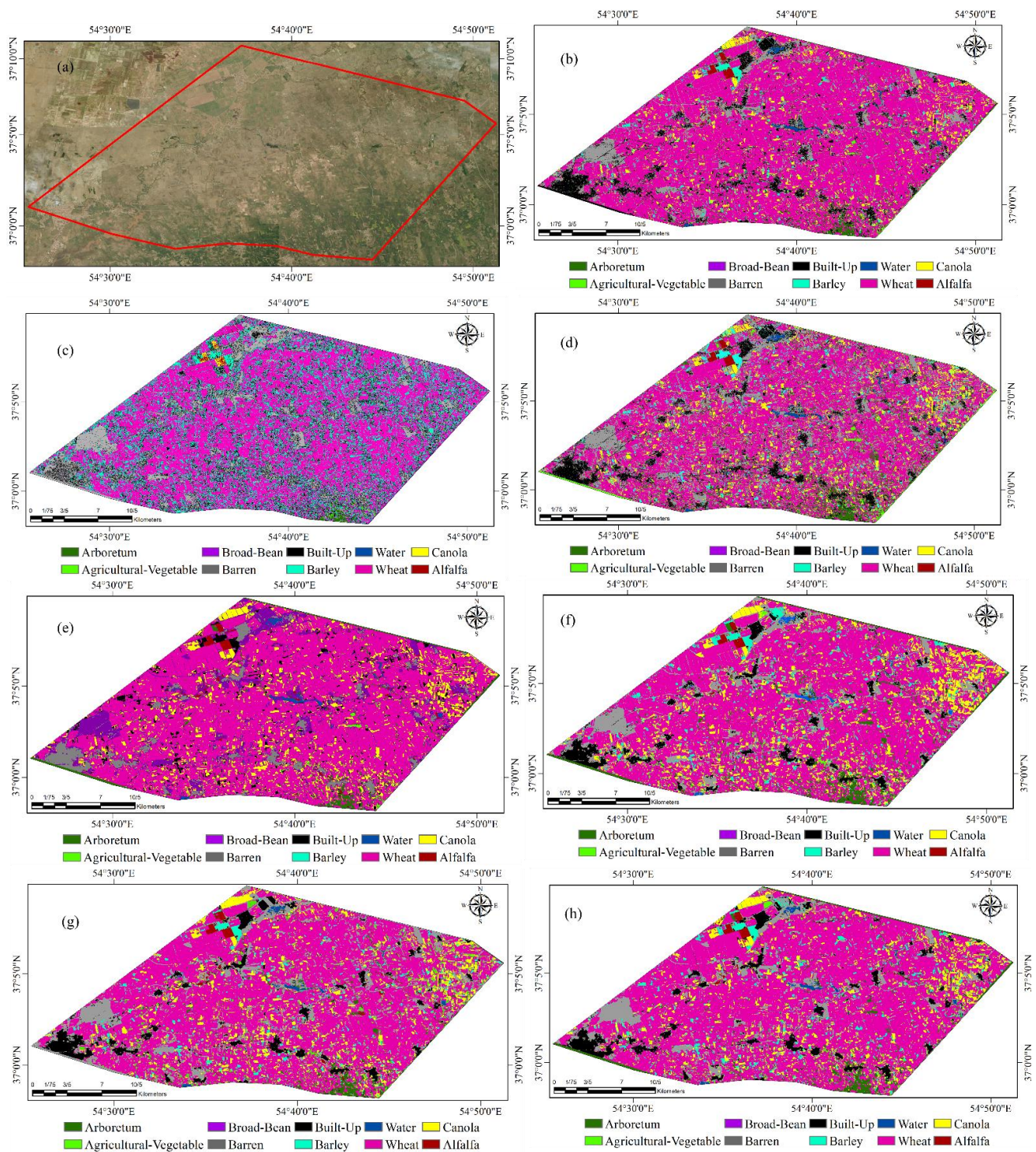


Figure 7. Comparison of the crop maps produced using the proposed algorithm and other classification methods: (a) high resolution satellite image, (b) XGBOOST, (c) RF, (d) R-CNN, (e) 2D-CNN, (f) 3D-CNN, (g) deep learning with an CBAM attention block, and (h) proposed method.

Figure 8 shows the confusion matrices of the proposed and all implemented classification methods. Generally, the proposed deep learning method resulted in the lowest confusion between the classes, indicating its high potential for accurate crop type mapping. Among non-agricultural classes, there was considerable confusion between the barren and built-up classes with other classes, except broad bean. Furthermore, water had the lowest mixing with different classes. Overall, most confusions occurred between the arboretum, barren, built-up, barley, and wheat classes. These confusions were much higher for the XGBOOST, RF, and R-CNN algorithms compared to other methods. For example, the highest confusion was between the barren and built-up classes (11,918 pixels) using the RF algorithm. The R-CNN algorithm also resulted in relatively high confusion between built-up/barren and barren/wheat. However, other deep learning algorithms, especially the proposed method, had higher accuracies. Among the 2D-CNN, 3D-CNN, and CBAM deep learning algorithms, the highest confusion was observed between barley and wheat using the 2D-CNN algorithm. The RF, XGBOSST, 2D-CNN, and RCNN classification methods could not discriminate the broad bean class from other classes mainly due to the lower number of its samples compared to other classes.

The statistical accuracy assessment of the crop maps using different accuracy measures is also summarized in Table 4. Regarding the non-deep learning algorithms, the RF classifier provided the lowest performance (OA = 74% and KC = 0.68), while XGBOOST provided a satisfactory result (OA = 87% and KC = 0.84). However, all the deep learning methods, except the R-CNN algorithm, achieved an OA of more than 90%. In particular, the proposed method provided the highest accuracy in mapping crops with an OA and KC of 98.5% and 0.98, respectively.

4.3. Impacts of the Time-Series NDVI on the Classification Results

The crop type classification provides useful information before harvesting agricultural products. This information can accurately be obtained by employing time-series NDVI datasets in a growing season. In this regard, the sensitivity of the number of the NDVI used in the classification was investigated in this study. For example, Figure 9 and Table 5 present the calculated confusion matrices and accuracy measures when different NDVI datasets were employed for crop type mapping using the proposed deep learning method. It was observed that using more NDVI images (i.e., seven months NDVI dataset) resulted in higher classification accuracies and lower confusions between different classes, closely followed by using six months NDVI datasets. As is clear from Figure 9, adding further information to the proposed method through adding more NDVI images could steadily reduce the uncertainties and confusion between the classes. For instance, the total interchangeable confusion between wheat and canola was continuously reduced by nearly 30% (i.e., from 800 wrongly classified pixels to 87 incorrectly classified pixels, when incorporating more time-series NDVI datasets). Furthermore, barley and wheat had the lowest confusion (113 pixels) when employing seven months of NDVI datasets, while the highest confusion (304 pixels) was associated with using two months of NDVI datasets. Overall, as is clear from Table 5, the highest accuracies were obtained when seven months of NDVI datasets were utilized.

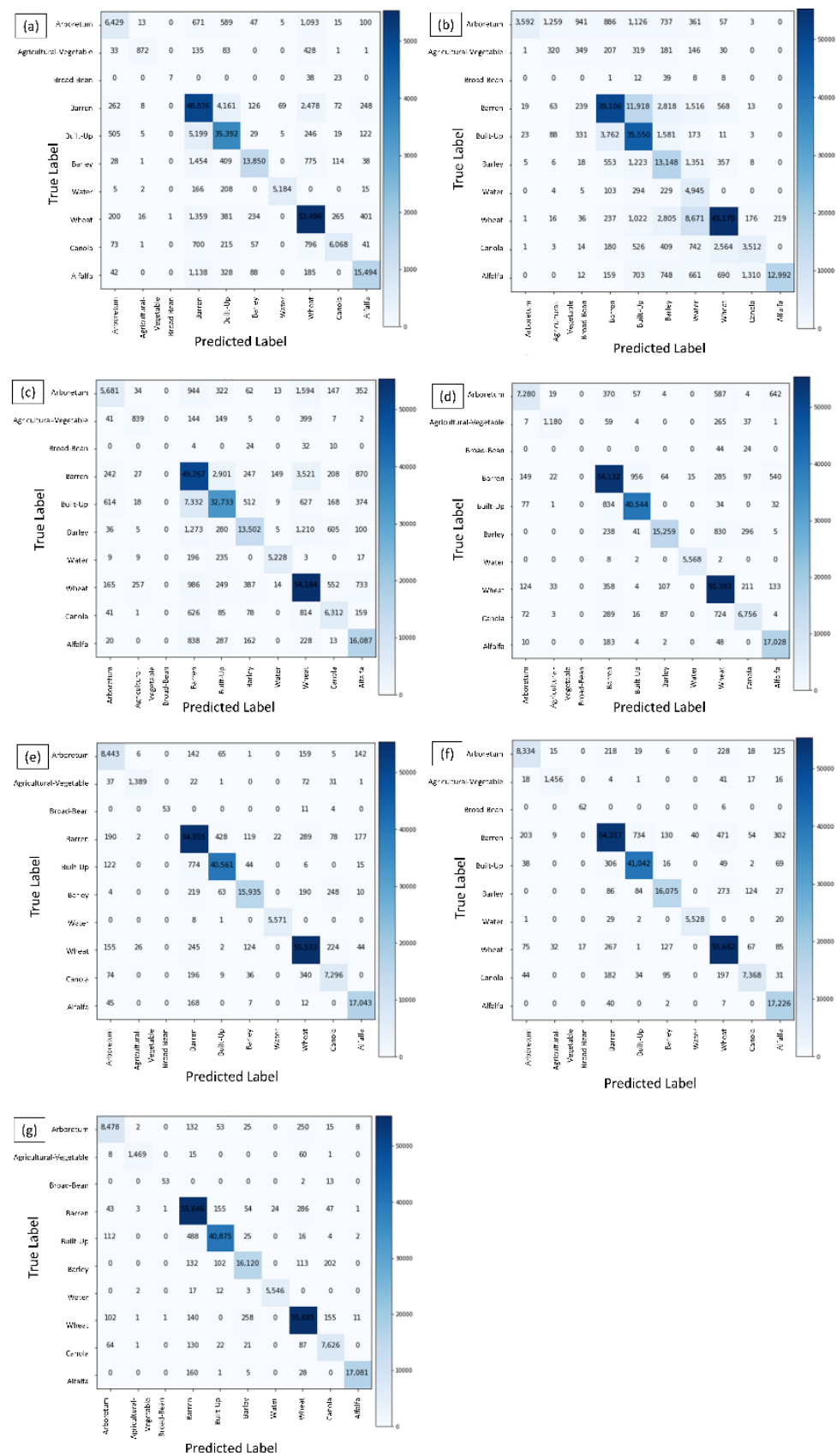


Figure 8. Comparison of the confusion matrices of different classification algorithms for crop mapping: (a) XGBOOST, (b) RF, (c) R-CNN, (d) 2D-CNN, (e) 3D-CNN, (f) deep learning with an CBAM attention block, and (g) proposed method.

Table 4. Comparison of the accuracies of different classification algorithms for crop mapping. The bold values show the highest accuracies (OA: overall accuracy, KC: kappa coefficient, UA: user accuracy, PA: producer accuracy, OE: omission error, CE: commission error, RF: random forest, R-CNN: recurrent convolution neural network, CNN: convolution neural network, CBAM: convolutional block attention module, 2D: 2-dimensional, 3D: 3-dimensional). The bold values show the highest accuracies.

Method	Index	Arboretum	Agricultural-Vegetable	Broad-Bean	Barren	Built-Up	Barley	Water	Wheat	Canola	Alfalfa	
RF	PA	98.63	18.19	0.00	86.53	67.47	57.93	26.62	90.97	69.89	98.34	
	UA	40.08	20.61	0.00	69.51	85.62	78.88	88.62	76.61	44.17	75.21	
	OE	1.37	81.81	100	13.47	32.53	42.07	73.38	9.03	30.11	1.66	
	CE	59.92	79.39	100	30.49	14.38	21.12	11.38	23.39	55.83	24.79	
	OA						73.68					
	KC						0.678					
XGBOOST	PA	84.85	94.99	87.50	81.86	84.74	95.97	98.50	89.86	92.26	94.13	
	UA	71.74	56.15	10.29	86.80	85.24	83.09	92.90	94.93	76.32	89.69	
	OE	15.15	5.01	12.50	18.14	15.26	4.03	1.50	10.14	7.74	5.87	
	CE	28.26	43.85	89.71	13.20	14.76	16.91	7.10	5.07	23.68	10.31	
	OA						87.48					
	KC						0.844					
R-CNN	PA	82.95	70.50	0.00	79.97	87.90	90.14	96.49	86.54	78.68	86.05	
	UA	62.09	52.90	0.00	85.78	77.22	79.35	91.77	94.19	77.77	91.22	
	OE	17.05	29.50	0.00	20.03	12.10	9.86	3.51	13.46	21.32	13.95	
	CE	37.91	47.10	100	14.22	22.78	20.65	8.23	5.81	22.23	8.78	
	OA						84.87					
	KC						0.810					
2D-CNN	PA	94.31	93.80	0.00	95.86	97.40	98.30	99.73	95.16	90.99	92.62	
	UA	81.22	78.98	0.00	96.22	97.64	91.54	99.78	96.28	84.97	98.57	
	OE	5.69	6.20	0.00	4.14	2.60	1.70	0.27	4.84	9.01	7.38	
	CE	18.78	24.02	100	3.78	2.36	8.46	0.22	1.72	15.03	1.43	
	OA						95.73					
	KC						0.947					
3D-CNN	PA	93.09	97.61	100	96.87	98.62	97.97	99.61	98.09	92.52	97.77	
	UA	94.20	89.44	77.94	97.68	97.69	95.60	99.84	98.54	91.76	98.66	
	OE	6.91	2.39	0.00	3.13	1.38	2.03	0.39	1.91	7.48	2.23	
	CE	5.80	10.56	22.06	2.32	2.31	4.40	0.16	1.46	8.24	1.34	
	OA						97.45					
	KC						0.968					
CBAM	PA	95.65	96.30	78.48	97.96	97.91	97.71	99.28	97.77	96.31	96.23	
	UA	92.98	93.75	91.18	96.55	98.84	96.44	99.07	98.81	92.67	99.72	
	OE	4.35	3.70	21.52	2.04	2.09	2.29	0.72	2.23	3.69	3.77	
	CE	7.02	6.25	8.82	3.45	1.16	3.56	0.93	1.19	7.33	0.28	
	OA						97.59					
	KC						0.970					
Proposed Method	PA	94.64	95.63	100	98.77	98.50	98.76	99.82	99.02	94.46	99.46	
	UA	96.47	95.75	76.47	98.46	99.33	97.14	99.82	98.73	96.42	99.15	
	OE	5.36	4.37	0.00	1.23	1.50	1.24	0.18	0.98	5.54	0.54	
	CE	3.53	4.25	23.53	1.54	0.67	2.86	0.18	1.27	3.58	0.85	
	OA						98.54					
	KC						0.981					

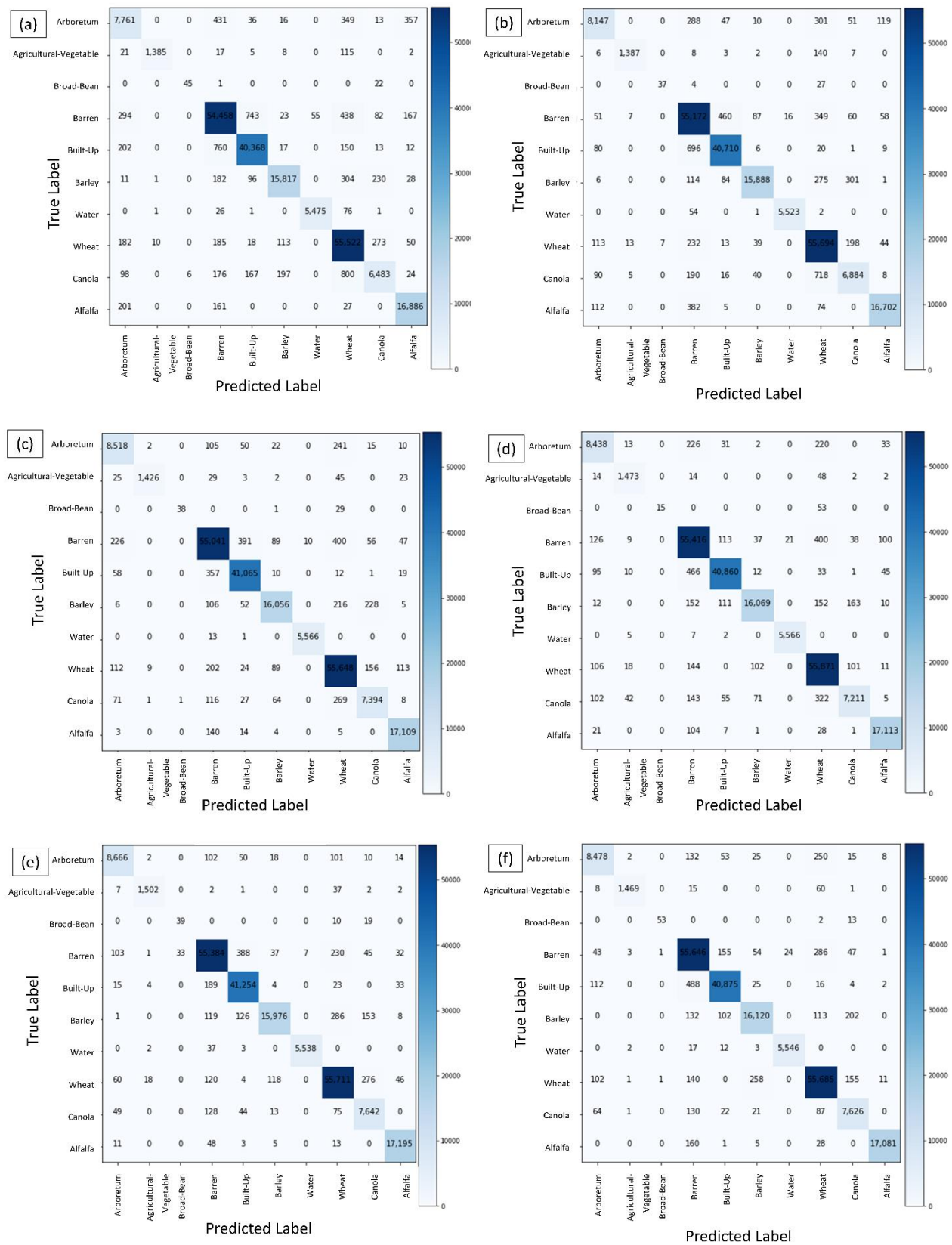


Figure 9. The effects of increasing the number of NDVI datasets in the confusion between different classes obtained by the proposed framework: (a) two, (b) three, (c) four, (d) five, (e) six, and (f) seven months after planting.

Table 5. The effects of increasing the number of NDVI datasets on the classification accuracies obtained using the proposed method. The bold values show the highest accuracies (OA: overall accuracy, KC: kappa coefficient, UA: user accuracy, PA: producer accuracy, OE: omission error, CE: commission error).

Time	Index	Arboretum	Agricultural-Vegetable	Broad Bean	Barren	Built-Up	Barley	Water	Wheat	Canola	Alfalfa
Two months after planting	PA	88.49	99.14	88.24	96.56	97.43	97.69	99.01	96.09	91.09	96.36
	UA	86.59	89.18	66.18	96.80	97.22	94.89	98.12	98.53	81.54	97.75
	OE	11.51	0.86	11.76	3.44	2.57	2.31	0.99	3.91	8.91	3.65
	CE	13.41	10.82	66.18	96.80	97.22	94.89	98.12	98.53	81.54	97.75
	OA	96.23									
	KC	0.953									
Three months after planting	PA	94.68	98.23	84.09	96.56	98.48	98.85	99.71	96.69	91.75	98.59
	UA	90.90	89.31	54.41	98.07	98.04	95.31	98.98	98.83	86.58	96.68
	OE	5.32	1.77	15.91	3.44	1.52	1.15	0.29	3.31	8.25	1.41
	CE	9.10	10.69	45.59	1.93	1.96	4.69	1.02	1.17	13.42	3.32
	OA	97.15									
	KC	0.964									
Four months after planting	PA	94.45	99.17	97.44	98.10	98.65	98.28	99.82	97.86	94.19	98.70
	UA	95.0	91.82	55.88	97.83	98.90	96.32	99.75	98.75	92.99	99.04
	OE	5.55	0.83	2.56	1.90	1.35	1.72	0.18	2.14	5.81	1.30
	CE	4.96	91.82	55.88	97.83	98.90	96.32	99.75	98.75	92.99	99.04
	OA	97.96									
	KC	0.974									
Five months after planting	PA	94.66	93.82	100	97.78	99.23	98.62	99.62	97.80	95.93	98.81
	UA	94.14	94.85	22.06	98.50	98.41	96.40	99.75	99.14	90.69	99.06
	OE	5.34	6.18	0.00	2.22	0.77	1.38	0.38	2.20	4.07	1.19
	CE	5.86	5.15	7.94	1.50	1.59	3.60	0.25	0.86	9.31	0.94
	OA	98.04									
	KC	0.975									
Six months after planting	PA	97.24	98.23	54.17	98.67	98.52	98.79	99.87	98.63	93.80	99.22
	UA	96.69	96.72	57.35	98.44	99.35	95.84	99.25	98.86	96.11	99.54
	OE	2.76	1.77	45.83	1.33	1.48	1.21	0.13	1.37	6.20	0.78
	CE	3.31	3.28	42.65	1.56	0.65	4.16	0.75	1.14	3.89	0.46
	OA	98.45									
	KC	0.980									
Seven months after planting	PA	94.64	95.63	100	98.77	98.50	98.76	99.82	99.02	94.46	99.46
	UA	96.47	95.75	76.47	98.46	99.33	97.14	99.82	98.73	96.42	99.15
	OE	5.36	4.37	0.00	1.23	1.50	1.24	0.18	0.98	5.54	0.54
	CE	3.53	4.25	23.53	1.54	0.67	2.86	0.18	1.27	3.58	0.85
	OA	98.54									
	KC	0.981									

4.4. Ablation Analysis

The ablation analysis is a crucial step for evaluating the performance of different aspects of an artificial intelligence method. The main purpose of this analysis was obtaining an insight into the effects of removing a part of the system on the general performance of the model. In this study, we investigated the impacts of ablation analysis on the efficiency of the proposed crop type mapping framework through three scenarios (S): (S#1): without AM, (S#2): without spectral attention block, and (S#3) without spatial attention block. The results of these scenarios were also compared with the proposed method when all the functions were used (i.e., S#4). Figure 10 shows the confusion matrices of four different scenarios of the ablation analysis. Although the obtained classification results were relatively

similar, the results indicated the higher potential of the proposed method empowered with the AM mechanism, especially in comparison to S#1. For example, the proposed architecture considerably reduced the confusion between barley and wheat, which was over 1000 pixels in S#1 and reached 113 in the proposed architecture. Moreover, the proposed method successfully reduced the slight mutual confusion between canola/alfalfa and arboretum/alfalfa, respectively. Finally, as is clear, the effect of spatial attention was more than spectral attention in the classification results.

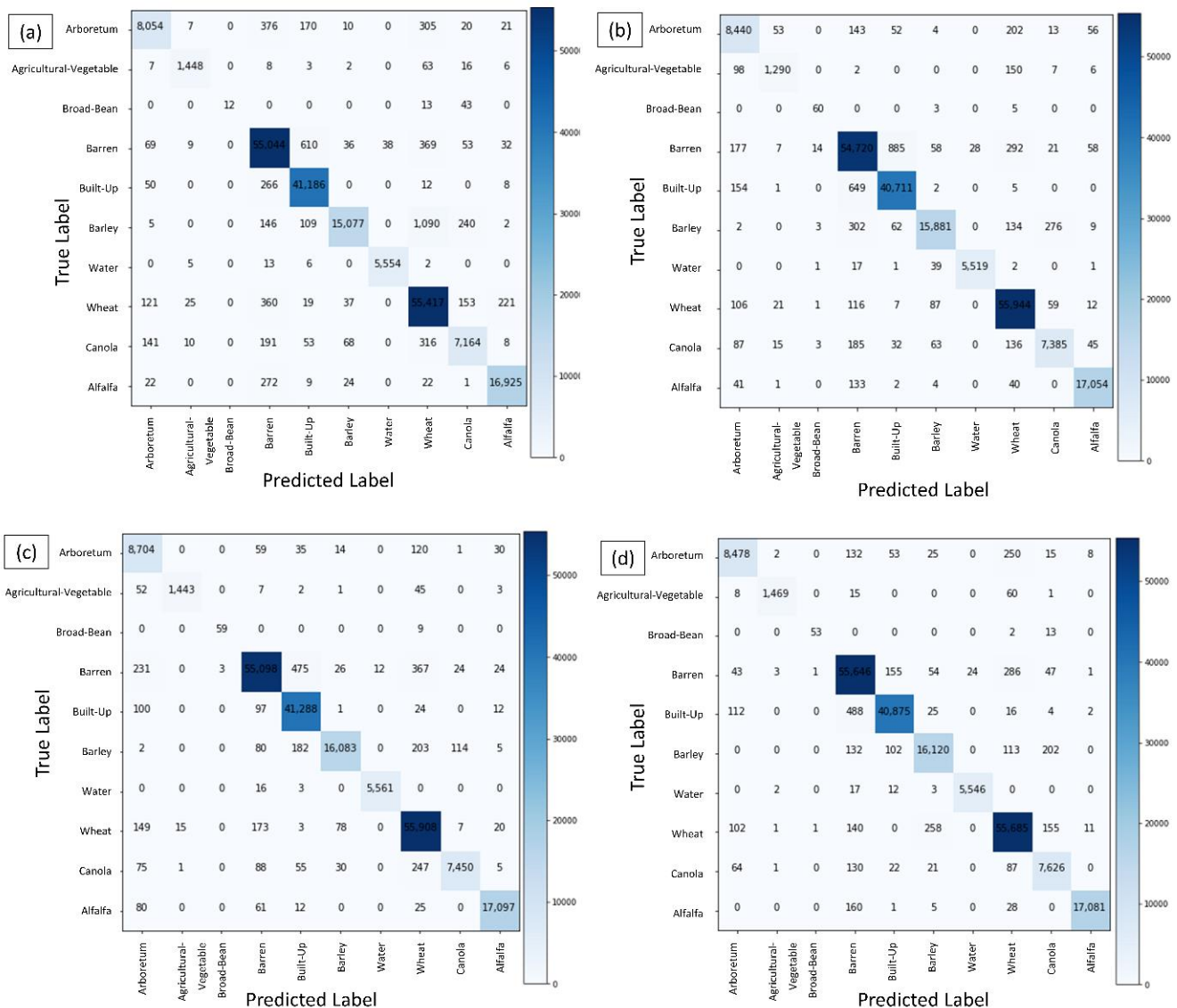


Figure 10. Ablation analysis of the proposed method: (a) without attention block, (b) without spectral attention block, (c) without spatial attention block, and (d) proposed method with all functions.

5. Discussion

5.1. Accuracy

In this study, a new crop mapping framework was proposed using Sentinel-2 time-series NDVI datasets. The results of crop mapping using different classifiers showed that the deep learning-based methods had relatively high potential. For example, the statistical methods (i.e., RF and XGBOOST) provided accuracies lower than 87%, while deep learning methods generally produced crop maps with more than 95% accuracy. Overall, the proposed method had the lowest errors in terms of OE and CE (under 5% in almost all classes).

Imbalanced reference samples are among the common problems in supervised learning frameworks [96,97]. Due to several limitations in this study, the size of the reference samples was not balanced for all classes. For example, the broad bean class had the lowest number of reference samples (i.e., 72 pixels). Nevertheless, the proposed method was able to classify this class with a UA of more than 76% and PA of 100%. This indicated the robustness of the proposed network against the imbalanced reference samples.

Figure 11 shows zoomed in patches of the classified results using the proposed method. Based on the results, the proposed method provided promising results for both crop and non-crop class types. For instance, the proposed method accurately delineated built-up areas with very few missed classifications. Additionally, the proposed method correctly classified arboretum areas in Figure 11c,d.

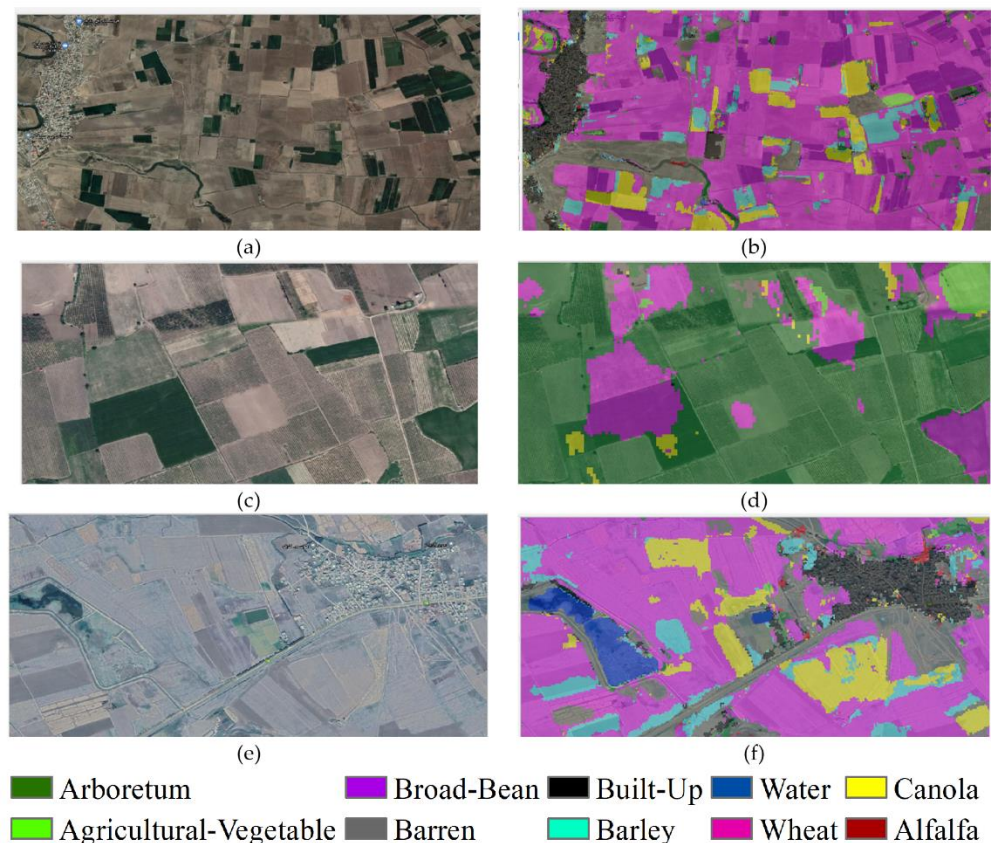


Figure 11. Comparison of the results of crop mapping (b,d,f) using the proposed method with very high resolution (VHR) imagery (a,c,e) in different areas. The left column is VHR imagery, and the right column is classified maps with the background of the VHR imagery.

5.2. Sensitivity Analysis

The effect of the number of NDVI datasets on the crop classification was also investigated in this study (see Section 4.3). Based on Table 5, the lowest accuracy was related to the two-month NDVI datasets (OA = 96%), and the highest accuracy was associated with the case of using NDVI datasets of all months (OA = 98.5%). As a result, although the agricultural crops could be detected with the NDVI datasets after two months of planting, increasing the number of NDVI datasets from other months of the growing season could potentially improve the accuracy. Table 6 shows the performance of the proposed method compared to other state-of-the-arts deep learning methods.

Table 6. Comparison of the performance of the proposed method with other crop mapping methods (OA: overall accuracy).

Reference	OA%	Method
Zhong, Hu, and Zhou [45]	85.54	Deep learning
Xu et al. [98]	98.3	Deep learning
Wei et al. [99]	85.01	Deep learning
Tamiminia et al. [100]	80.48	Kernel-based clustering
Hamidi et al. [101]	95.26	Deep learning
Kwak and Park [102]	96.37	Deep learning
Proposed	98.54	Deep learning

5.3. Proposed Architecture and Deep Feature Extraction

Informative feature extraction is of the most critical factors in classification tasks. These features can be obtained based on combining spectral and spatial features. The results of pixel-based crop type mapping based on RF and XGBOOST algorithms showed that these methods had lower capability than deep learning-based approaches, mainly due to the employment of only spectral features. This indicated the impact of extracting informative spatial features for accurate crop type classification.

Suitable architecture is a key factor for extracting deep features based on CNN methods. In this regard, we designed a new framework for extracting deep features based on multiscale-residual block convolutions. Furthermore, spectral and spatial AMs were implemented to increase the efficiency of the proposed framework [89]. The results of crop type mapping demonstrated the high capability of the proposed algorithm to extract informative deep features, which could enhance the performance of the proposed method compared to other advanced crop mapping techniques.

The stability of deep learning-based methods is the most important factor in classification. To this end, the efficiency of the proposed method was evaluated in ten different epochs, the results of which are provided in Table 7. Based on the results, the proposed method had a high stability in different runs because the OA did not considerably change (i.e., 98.49 ± 0.04).

Table 7. The accuracy of the proposed method through different iterations (OA: overall accuracy).

Index	Value
OA	98.50, 98.51, 98.52, 98.54, 98.40, 98.46, 98.48, 98.52, 98.47, 98.49
Mean	98.49
Standard Deviation	± 0.04

Although the semantic segmentation-based methods, such as deeplabV3+ and U-Net, have achieved promising results in crop mapping [99,103], they require a high amount of sample datasets. This is because all pixels of the image dataset must be labeled through field visits, which is time-consuming and resource-intensive. However, the proposed method required nearly 7000 training samples, the collection of which was applicable in comparison with semantic segmentation-based methods.

The AM increases the performance of deep learning methods in processing tasks [79,81]. The CBAM is the most well-known attention block among other types of AMs [82]. Based on the results, the proposed AM outperformed the CBAM mechanism in all classes. This indicated the high potential of the AM in extracting deep features. In fact, the AM improved the accuracy of the proposed deep learning framework by concentrating the network on informative deep feature extracting.

6. Conclusions

Timely and accurate crop mapping is one of the most important components for managing and making decisions to support food security. In this regard, this study presented a novel deep learning-based technique for crop type mapping. We evaluated the efficiency of the proposed method on seven and three crop and non-crop classes, respectively. This research used the time-series NDVI for mapping crop types mainly because of the dynamic nature of crops. The results of crop type mapping were also compared with other advanced supervised learning techniques. The statistical and visual analyses indicated that the proposed deep learning model produced excellent performance in comparison to different state-of-the-art classification methods. Furthermore, the efficiency of the proposed AM was proven in the crop type classification task, as it resulted in achieving higher classification accuracy than the CBAM architecture. Moreover, we assessed the efficiency of the proposed manner using different NDVI datasets and observed the high potential of the proposed method by achieving high accuracies with different NDVI datasets (i.e., OA = 96% to 98%). The highest accuracy was related to when seven-month NDVI datasets were employed.

Author Contributions: Conceptualization, S.T.S., A.G. and M.A.; methodology, S.T.S.; writing—original draft preparation, S.T.S. and A.G.; writing—review and editing, S.T.S., M.A. and A.G.; visualization, S.T.S. and M.A.; supervision, M.A.; funding acquisition, M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. These datasets can be found here: [<https://scihub.copernicus.eu/>] (accessed on 15 November 2021).

Acknowledgments: The authors would like to thank the European Space Agency (ESA) for providing the Sentinel-2 Level-1C products.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. United Nations, Department of Economic and Social Affairs, Population Division. *World Population Prospects: The 2015 Revision; Key Findings and Advance Tables*; United Nations: New York, NY, USA, 2015.
2. Waldner, F.; Canto, G.S.; Defourny, P. Automated annual cropland mapping using knowledge-based temporal features. *ISPRS J. Photogramm. Remote Sens.* **2015**, *110*, 1–13. [[CrossRef](#)]
3. Khan, M.A.; Tahir, A.; Khurshid, N.; Ahmed, M.; Boughanmi, H. Economic effects of climate change-induced loss of agricultural production by 2050: A case study of Pakistan. *Sustainability* **2020**, *12*, 1216. [[CrossRef](#)]
4. Shi, W.; Wang, M.; Liu, Y. Crop yield and production responses to climate disasters in China. *Sci. Total Environ.* **2021**, *750*, 141147. [[CrossRef](#)]
5. Shelestov, A.; Lavreniuk, M.; Kussul, N.; Novikov, A.; Skakun, S. Exploring Google Earth Engine platform for big data processing: Classification of multi-temporal satellite imagery for crop mapping. *Front. Earth Sci.* **2017**, *5*, 17. [[CrossRef](#)]
6. Agovino, M.; Casaccia, M.; Ciommi, M.; Ferrara, M.; Marchesano, K. Agriculture, climate change and sustainability: The case of EU-28. *Ecol. Indic.* **2019**, *105*, 525–543. [[CrossRef](#)]
7. Anwar, M.R.; Li Liu, D.; Macadam, I.; Kelly, G. Adapting agriculture to climate change: A review. *Theor. Appl. Climatol.* **2013**, *113*, 225–245. [[CrossRef](#)]
8. Amani, M.; Kakooei, M.; Moghimi, A.; Ghorbanian, A.; Ranjgar, B.; Mahdavi, S.; Davidson, A.; Fisette, T.; Rollin, P.; Brisco, B. Application of google earth engine cloud computing platform, sentinel imagery, and neural networks for crop mapping in canada. *Remote Sens.* **2020**, *12*, 3561. [[CrossRef](#)]
9. Bégué, A.; Arvor, D.; Bellon, B.; Betbeder, J.; De Abelleira, D.; Ferraz, R.P.D.; Lebourgeois, V.; Lelong, C.; Simões, M.; Verón, S.R. Remote sensing and cropping practices: A review. *Remote Sens.* **2018**, *10*, 99. [[CrossRef](#)]
10. Karthikeyan, L.; Chawla, I.; Mishra, A.K. A review of remote sensing applications in agriculture for food security: Crop growth and yield, irrigation, and crop losses. *J. Hydrol.* **2020**, *586*, 124905. [[CrossRef](#)]
11. Orynbaikyzy, A.; Gessner, U.; Conrad, C. Crop type classification using a combination of optical and radar remote sensing data: A review. *Int. J. Remote Sens.* **2019**, *40*, 6553–6595. [[CrossRef](#)]

12. Weiss, M.; Jacob, F.; Duveiller, G. Remote sensing for agricultural applications: A meta-review. *Remote Sens. Environ.* **2020**, *236*, 111402. [\[CrossRef\]](#)
13. Di, Y.; Zhang, G.; You, N.; Yang, T.; Zhang, Q.; Liu, R.; Doughty, R.B.; Zhang, Y. Mapping Croplands in the Granary of the Tibetan Plateau Using All Available Landsat Imagery, A Phenology-Based Approach, and Google Earth Engine. *Remote Sens.* **2021**, *13*, 2289. [\[CrossRef\]](#)
14. Ren, S.; An, S. Temporal Pattern Analysis of Cropland Phenology in Shandong Province of China Based on Two Long-Sequence Remote Sensing Data. *Remote Sens.* **2021**, *13*, 4071. [\[CrossRef\]](#)
15. Mutanga, O.; Dube, T.; Galal, O. Remote sensing of crop health for food security in Africa: Potentials and constraints. *Remote Sens. Appl. Soc. Environ.* **2017**, *8*, 231–239. [\[CrossRef\]](#)
16. Cai, Y.; Guan, K.; Peng, J.; Wang, S.; Seifert, C.; Wardlow, B.; Li, Z. A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach. *Remote Sens. Environ.* **2018**, *210*, 35–47. [\[CrossRef\]](#)
17. Johnson, D.M.; Mueller, R. Pre-and within-season crop type classification trained with archival land cover information. *Remote Sens. Environ.* **2021**, *264*, 112576. [\[CrossRef\]](#)
18. Kenduiywo, B.K.; Bargiel, D.; Soergel, U. Crop-type mapping from a sequence of Sentinel 1 images. *Int. J. Remote Sens.* **2018**, *39*, 6383–6404. [\[CrossRef\]](#)
19. Donohue, R.J.; Lawes, R.A.; Mata, G.; Gobbett, D.; Ouzman, J. Towards a national, remote-sensing-based model for predicting field-scale crop yield. *Field Crops Res.* **2018**, *227*, 79–90. [\[CrossRef\]](#)
20. Kern, A.; Barcza, Z.; Marjanović, H.; Árendás, T.; Fodor, N.; Bónis, P.; Bognár, P.; Lichtenberger, J. Statistical modelling of crop yield in Central Europe using climate data and remote sensing vegetation indices. *Agric. For. Meteorol.* **2018**, *260*, 300–320. [\[CrossRef\]](#)
21. Son, N.-T.; Chen, C.-F.; Chen, C.-R.; Guo, H.-Y. Classification of multitemporal Sentinel-2 data for field-level monitoring of rice cropping practices in Taiwan. *Adv. Space Res.* **2020**, *65*, 1910–1921. [\[CrossRef\]](#)
22. Zhang, H.; Kang, J.; Xu, X.; Zhang, L. Accessing the temporal and spectral features in crop type mapping using multi-temporal Sentinel-2 imagery: A case study of Yí'an County, Heilongjiang province, China. *Comput. Electron. Agric.* **2020**, *176*, 105618. [\[CrossRef\]](#)
23. Dey, S.; Mandal, D.; Robertson, L.D.; Banerjee, B.; Kumar, V.; McNairn, H.; Bhattacharya, A.; Rao, Y. In-season crop classification using elements of the Kennaugh matrix derived from polarimetric RADARSAT-2 SAR data. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *88*, 102059. [\[CrossRef\]](#)
24. Planque, C.; Lucas, R.; Punalekar, S.; Chognard, S.; Hurford, C.; Owers, C.; Horton, C.; Guest, P.; King, S.; Williams, S. National crop mapping using sentinel-1 time series: A knowledge-based descriptive algorithm. *Remote Sens.* **2021**, *13*, 846. [\[CrossRef\]](#)
25. Prins, A.J.; Van Niekerk, A. Regional Mapping of Vineyards Using Machine Learning and LiDAR Data. *Int. J. Appl. Geospatial Res. (IJAGR)* **2020**, *11*, 1–22. [\[CrossRef\]](#)
26. ten Harkel, J.; Bartholomeus, H.; Kooistra, L. Biomass and crop height estimation of different crops using UAV-based LiDAR. *Remote Sens.* **2020**, *12*, 17. [\[CrossRef\]](#)
27. Meng, S.; Wang, X.; Hu, X.; Luo, C.; Zhong, Y. Deep learning-based crop mapping in the cloudy season using one-shot hyperspectral satellite imagery. *Comput. Electron. Agric.* **2021**, *186*, 106188. [\[CrossRef\]](#)
28. Moriya, É.A.S.; Imai, N.N.; Tommaselli, A.M.G.; Berveglieri, A.; Santos, G.H.; Soares, M.A.; Marino, M.; Reis, T.T. Detection and mapping of trees infected with citrus gummosis using UAV hyperspectral data. *Comput. Electron. Agric.* **2021**, *188*, 106298. [\[CrossRef\]](#)
29. Chandel, A.K.; Molaei, B.; Khot, L.R.; Peters, R.T.; Stöckle, C.O. High resolution geospatial evapotranspiration mapping of irrigated field crops using multispectral and thermal infrared imagery with metric energy balance model. *Drones* **2020**, *4*, 52. [\[CrossRef\]](#)
30. James, K.; Nichol, C.J.; Wade, T.; Cowley, D.; Gibson Poole, S.; Gray, A.; Gillespie, J. Thermal and Multispectral Remote Sensing for the Detection and Analysis of Archaeologically Induced Crop Stress at a UK Site. *Drones* **2020**, *4*, 61. [\[CrossRef\]](#)
31. Kyere, I.; Astor, T.; Graß, R.; Wachendorf, M. Agricultural crop discrimination in a heterogeneous low-mountain range region based on multi-temporal and multi-sensor satellite data. *Comput. Electron. Agric.* **2020**, *179*, 105864. [\[CrossRef\]](#)
32. 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. [\[CrossRef\]](#)
33. Hasanlou, M.; Shah-Hosseini, R.; Seydi, S.T.; Karimzadeh, S.; Matsuoka, M. Earthquake Damage Region Detection by Multitemporal Coherence Map Analysis of Radar and Multispectral Imagery. *Remote Sens.* **2021**, *13*, 1195. [\[CrossRef\]](#)
34. Seydi, S.; Rastiveis, H. A Deep Learning Framework for Roads Network Damage Assessment Using Post-Earthquake Lidar Data. *Int. Arch. Photogramm. Remote Sens. Photogramm. Spat. Inf. Sci.* **2019**, *42*, 955–961. [\[CrossRef\]](#)
35. Seydi, S.T.; Hasanlou, M.; Amani, M.; Huang, W. Oil Spill Detection Based on Multi-Scale Multi-Dimensional Residual CNN for Optical Remote Sensing Imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 10941–10952. [\[CrossRef\]](#)
36. Lary, D.J.; Alavi, A.H.; Gandomi, A.H.; Walker, A.L. Machine learning in geosciences and remote sensing. *Geosci. Front.* **2016**, *7*, 3–10. [\[CrossRef\]](#)
37. Maxwell, A.E.; Warner, T.A.; Fang, F. Implementation of machine-learning classification in remote sensing: An applied review. *Int. J. Remote Sens.* **2018**, *39*, 2784–2817. [\[CrossRef\]](#)

38. Zhang, H.; Li, Q.; Liu, J.; Shang, J.; Du, X.; McNairn, H.; Champagne, C.; Dong, T.; Liu, M. Image classification using rapideye data: Integration of spectral and textual features in a random forest classifier. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 5334–5349. [\[CrossRef\]](#)
39. Mandal, D.; Kumar, V.; Rao, Y.S. An assessment of temporal RADARSAT-2 SAR data for crop classification using KPCA based support vector machine. *Geocarto Int.* **2020**, 1–13. [\[CrossRef\]](#)
40. Maponya, M.G.; Van Niekerk, A.; Mashimbye, Z.E. Pre-harvest classification of crop types using a Sentinel-2 time-series and machine learning. *Comput. Electron. Agric.* **2020**, *169*, 105164. [\[CrossRef\]](#)
41. Saini, R.; Ghosh, S.K. Crop classification in a heterogeneous agricultural environment using ensemble classifiers and single-date Sentinel-2A imagery. *Geocarto Int.* **2019**, *36*, 2141–2159. [\[CrossRef\]](#)
42. Seydi, S.T.; Hasanlou, M.; Chanussot, J. DSMNN-Net: A Deep Siamese Morphological Neural Network Model for Burned Area Mapping Using Multispectral Sentinel-2 and Hyperspectral PRISMA Images. *Remote Sens.* **2021**, *13*, 5138. [\[CrossRef\]](#)
43. Koirala, A.; Walsh, K.B.; Wang, Z.; McCarthy, C. Deep learning—Method overview and review of use for fruit detection and yield estimation. *Comput. Electron. Agric.* **2019**, *162*, 219–234. [\[CrossRef\]](#)
44. Wan, X.; Zhao, C.; Wang, Y.; Liu, W. Stacked sparse autoencoder in hyperspectral data classification using spectral-spatial, higher order statistics and multifractal spectrum features. *Infrared Phys. Photogramm. Technol.* **2017**, *86*, 77–89. [\[CrossRef\]](#)
45. Zhong, L.; Hu, L.; Zhou, H. Deep learning based multi-temporal crop classification. *Remote Sens. Environ.* **2019**, *221*, 430–443. [\[CrossRef\]](#)
46. Bhosle, K.; Musande, V. Evaluation of CNN model by comparing with convolutional autoencoder and deep neural network for crop classification on hyperspectral imagery. *Geocarto Int.* **2020**, 1–15. [\[CrossRef\]](#)
47. Ji, S.; Zhang, C.; Xu, A.; Shi, Y.; Duan, Y. 3D convolutional neural networks for crop classification with multi-temporal remote sensing images. *Remote Sens.* **2018**, *10*, 75. [\[CrossRef\]](#)
48. Li, Z.; Chen, G.; Zhang, T. A CNN-Transformer Hybrid Approach for Crop Classification Using Multitemporal Multisensor Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 847–858. [\[CrossRef\]](#)
49. Mazzia, V.; Khaliq, A.; Chiaberge, M. Improvement in land cover and crop classification based on temporal features learning from Sentinel-2 data using recurrent-convolutional neural network (R-CNN). *Appl. Sci.* **2020**, *10*, 238. [\[CrossRef\]](#)
50. Yang, S.; Gu, L.; Li, X.; Jiang, T.; Ren, R. Crop classification method based on optimal feature selection and hybrid CNN-RF networks for multi-temporal remote sensing imagery. *Remote Sens.* **2020**, *12*, 3119. [\[CrossRef\]](#)
51. 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. [\[CrossRef\]](#)
52. Akbari, E.; Darvishi Bolorani, A.; Neysani Samany, N.; Hamzeh, S.; Soufizadeh, S.; Pignatti, S. Crop mapping using random forest and particle swarm optimization based on multi-temporal Sentinel-2. *Remote Sens.* **2020**, *12*, 1449. [\[CrossRef\]](#)
53. Asgarian, A.; Soffianian, A.; Pourmanafi, S. Crop type mapping in a highly fragmented and heterogeneous agricultural landscape: A case of central Iran using multi-temporal Landsat 8 imagery. *Comput. Electron. Agric.* **2016**, *127*, 531–540. [\[CrossRef\]](#)
54. Saadat, M.; Hasanlou, M.; Homayouni, S. Rice Crop Mapping Using SENTINEL-1 Time Series Images (case Study: Mazandaran, Iran). *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *42*, 897–904. [\[CrossRef\]](#)
55. Rezaei, E.E.; Ghazaryan, G.; Moradi, R.; Dubovyk, O.; Siebert, S. Crop harvested area, not yield, drives variability in crop production in Iran. *Environ. Res. Lett.* **2021**, *16*, 064058. [\[CrossRef\]](#)
56. Maghrebi, M.; Noori, R.; Bhattarai, R.; Mundher Yaseen, Z.; Tang, Q.; Al-Ansari, N.; Danandeh Mehr, A.; Karbassi, A.; Omidvar, J.; Farnoush, H. Iran's Agriculture in the Anthropocene. *Earth's Future* **2020**, *8*, e2020EF001547. [\[CrossRef\]](#)
57. Karandish, F. Socioeconomic benefits of conserving Iran's water resources through modifying agricultural practices and water management strategies. *Ambio* **2021**, *50*, 1824–1840. [\[CrossRef\]](#)
58. Momm, H.G.; ElKadiri, R.; Porter, W. Crop-type classification for long-term modeling: An integrated remote sensing and machine learning approach. *Remote Sens.* **2020**, *12*, 449. [\[CrossRef\]](#)
59. Boali, H.; Asgari, H.; Mohammadian Behbahani, A.; Salmanmahiny, A.; Naimi, B. Provide early desertification warning system based on climate and groundwater criteria (Study area: Aq Qala and Gomishan counties). *Geogr. Dev. Iran. J.* **2021**, *19*, 285–306. [\[CrossRef\]](#)
60. Nasrollahi, N.; Kazemi, H.; Kamkar, B. Feasibility of ley-farming system performance in a semi-arid region using spatial analysis. *Ecol. Indic.* **2017**, *72*, 239–248. [\[CrossRef\]](#)
61. Seydi, S.T.; Akhoondzadeh, M.; Amani, M.; Mahdavi, S. Wildfire damage assessment over Australia using sentinel-2 imagery and MODIS land cover product within the google earth engine cloud platform. *Remote Sens.* **2021**, *13*, 220. [\[CrossRef\]](#)
62. Pan, L.; Xia, H.; Zhao, X.; Guo, Y.; Qin, Y. Mapping winter crops using a phenology algorithm, time-series Sentinel-2 and Landsat-7/8 images, and Google Earth Engine. *Remote Sens.* **2021**, *13*, 2510. [\[CrossRef\]](#)
63. Lambert, M.-J.; Traoré, P.C.S.; Blaes, X.; Baret, P.; Defourny, P. Estimating smallholder crops production at village level from Sentinel-2 time series in Mali's cotton belt. *Remote Sens. Environ.* **2018**, *216*, 647–657. [\[CrossRef\]](#)
64. Morais, C.L.; Santos, M.C.; Lima, K.M.; Martin, F.L. Improving data splitting for classification applications in spectrochemical analyses employing a random-mutation Kennard-Stone algorithm approach. *Bioinformatics* **2019**, *35*, 5257–5263. [\[CrossRef\]](#) [\[PubMed\]](#)
65. Butcher, B.; Smith, B.J. *Feature Engineering and Selection: A Practical Approach for Predictive Models*; Kuhn, M., Johnson, K., Eds.; Chapman & Hall/CRC Press: Boca Raton, FL, USA, 2019; ISBN 978-1-13-807922-9.

66. 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. [\[CrossRef\]](#)
67. Ghorbanian, A.; Zaghian, S.; Asiyabi, R.M.; Amani, M.; Mohammadzadeh, A.; Jamali, S. Mangrove ecosystem mapping using Sentinel-1 and Sentinel-2 satellite images and random forest algorithm in Google Earth Engine. *Remote Sens.* **2021**, *13*, 2565. [\[CrossRef\]](#)
68. Main-Knorn, M.; Pflug, B.; Louis, J.; Debaecker, V.; Müller-Wilm, U.; Gascon, F. Sen2Cor for sentinel-2. In Proceedings of the Image and Signal Processing for Remote Sensing XXIII, Warsaw, Poland, 11–14 September 2017; p. 1042704.
69. Pettorelli, N. *The Normalized Difference Vegetation Index*; Oxford University Press: Oxford, UK, 2013.
70. Townshend, J.R.; Justice, C. Analysis of the dynamics of African vegetation using the normalized difference vegetation index. *Int. J. Remote Sens.* **1986**, *7*, 1435–1445. [\[CrossRef\]](#)
71. Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37–52. [\[CrossRef\]](#)
72. Jakubauskas, M.E.; Legates, D.R.; Kastens, J.H. Crop identification using harmonic analysis of time-series AVHRR NDVI data. *Comput. Electron. Agric.* **2002**, *37*, 127–139. [\[CrossRef\]](#)
73. Pan, Z.; Huang, J.; Zhou, Q.; Wang, L.; Cheng, Y.; Zhang, H.; Blackburn, G.A.; Yan, J.; Liu, J. Mapping crop phenology using NDVI time-series derived from HJ-1 A/B data. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *34*, 188–197. [\[CrossRef\]](#)
74. Skakun, S.; Franch, B.; Vermote, E.; Roger, J.-C.; Becker-Reshef, I.; Justice, C.; Kussul, N. Early season large-area winter crop mapping using MODIS NDVI data, growing degree days information and a Gaussian mixture model. *Remote Sens. Environ.* **2017**, *195*, 244–258. [\[CrossRef\]](#)
75. 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. [\[CrossRef\]](#)
76. Li, F.; Ren, J.; Wu, S.; Zhao, H.; Zhang, N. Comparison of regional winter wheat mapping results from different similarity measurement indicators of NDVI time series and their optimized thresholds. *Remote Sens.* **2021**, *13*, 1162. [\[CrossRef\]](#)
77. Wu, Z.; Liu, Y.; Han, Y.; Zhou, J.; Liu, J.; Wu, J. Mapping farmland soil organic carbon density in plains with combined cropping system extracted from NDVI time-series data. *Sci. Total Environ.* **2021**, *754*, 142120. [\[CrossRef\]](#)
78. Ghaffarian, S.; Valente, J.; Van Der Voort, M.; Tekinerdogan, B. Effect of Attention Mechanism in Deep Learning-Based Remote Sensing Image Processing: A Systematic Literature Review. *Remote Sens.* **2021**, *13*, 2965. [\[CrossRef\]](#)
79. Li, M.; Wang, Y.; Wang, Z.; Zheng, H. A deep learning method based on an attention mechanism for wireless network traffic prediction. *Ad Hoc Netw.* **2020**, *107*, 102258. [\[CrossRef\]](#)
80. Li, X.; Zhang, W.; Ding, Q. Understanding and improving deep learning-based rolling bearing fault diagnosis with attention mechanism. *Signal Process.* **2019**, *161*, 136–154. [\[CrossRef\]](#)
81. Niu, Z.; Zhong, G.; Yu, H. A review on the attention mechanism of deep learning. *Neurocomputing* **2021**, *452*, 48–62. [\[CrossRef\]](#)
82. Woo, S.; Park, J.; Lee, J.-Y.; Kweon, I.S. Cbam: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 3–19.
83. Huang, G.; Zhu, J.; Li, J.; Wang, Z.; Cheng, L.; Liu, L.; Li, H.; Zhou, J. Channel-attention U-Net: Channel attention mechanism for semantic segmentation of esophagus and esophageal cancer. *IEEE Access.* **2020**, *8*, 122798–122810. [\[CrossRef\]](#)
84. Li, H.; Qiu, K.; Chen, L.; Mei, X.; Hong, L.; Tao, C. SCAAttNet: Semantic segmentation network with spatial and channel attention mechanism for high-resolution remote sensing images. *IEEE Geosci. Remote Sens. Lett.* **2020**, *18*, 905–909. [\[CrossRef\]](#)
85. Tong, W.; Chen, W.; Han, W.; Li, X.; Wang, L. Channel-attention-based DenseNet network for remote sensing image scene classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 4121–4132. [\[CrossRef\]](#)
86. Zhou, T.; Canu, S.; Ruan, S. Automatic COVID-19 CT segmentation using U-Net integrated spatial and channel attention mechanism. *Int. J. Imaging Syst. Technol.* **2021**, *31*, 16–27. [\[CrossRef\]](#)
87. Mohanty, A.; Gitelman, D.R.; Small, D.M.; Mesulam, M.M. The spatial attention network interacts with limbic and monoaminergic systems to modulate motivation-induced attention shifts. *Cereb. Cortex* **2008**, *18*, 2604–2613. [\[CrossRef\]](#)
88. Mou, L.; Zhao, Y.; Chen, L.; Cheng, J.; Gu, Z.; Hao, H.; Qi, H.; Zheng, Y.; Frangi, A.; Liu, J. CS-Net: Channel and spatial attention network for curvilinear structure segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Shenzhen, China, 13–17 October 2019; pp. 721–730.
89. Sun, H.; Zheng, X.; Lu, X.; Wu, S. Spectral-spatial attention network for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 3232–3245. [\[CrossRef\]](#)
90. Seydi, S.T.; Hasanlou, M.; Amani, M. A new end-to-end multi-dimensional CNN framework for land cover/land use change detection in multi-source remote sensing datasets. *Remote Sens.* **2020**, *12*, 2010. [\[CrossRef\]](#)
91. Seydi, S.T.; Hasanlou, M. A New Structure for Binary and Multiple Hyperspectral Change Detection Based on Spectral Unmixing and Convolutional Neural Network. *Measurement* **2021**, *186*, 110137. [\[CrossRef\]](#)
92. Dobrinić, D.; Medak, D.; Gašparović, M. Integration of Multitemporal SENTINEL-1 and SENTINEL-2 Imagery for Land-Cover Classification Using Machine Learning Methods. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *43*, 91–98. [\[CrossRef\]](#)
93. Zhang, W.; Liu, H.; Wu, W.; Zhan, L.; Wei, J. Mapping rice paddy based on machine learning with Sentinel-2 multi-temporal data: Model comparison and transferability. *Remote Sens.* **2020**, *12*, 1620. [\[CrossRef\]](#)
94. Tatsumi, K.; Yamashiki, Y.; Torres, M.A.C.; Taïpe, C.L.R. Crop classification of upland fields using Random forest of time-series Landsat 7 ETM+ data. *Comput. Electron. Agric.* **2015**, *115*, 171–179. [\[CrossRef\]](#)

95. Chuc, M.D.; Anh, N.H.; Thuy, N.T.; Hung, B.Q.; Thanh, N.T.N. Paddy rice mapping in red river delta region using landsat 8 images: Preliminary results. In Proceedings of the 2017 9th International Conference on Knowledge and Systems Engineering (KSE), Hue, Vietnam, 19–21 October 2017; pp. 209–214.
96. Naboureh, A.; Ebrahimi, H.; Azadbakht, M.; Bian, J.; Amani, M. RUESVMs: An Ensemble Method to Handle the Class Imbalance Problem in Land Cover Mapping Using Google Earth Engine. *Remote Sens.* **2020**, *12*, 3484. [[CrossRef](#)]
97. Naboureh, A.; Li, A.; Bian, J.; Lei, G.; Amani, M. A hybrid data balancing method for classification of imbalanced training data within google earth engine: Case studies from mountainous regions. *Remote Sens.* **2020**, *12*, 3301. [[CrossRef](#)]
98. Xu, J.; Yang, J.; Xiong, X.; Li, H.; Huang, J.; Ting, K.; Ying, Y.; Lin, T. Towards interpreting multi-temporal deep learning models in crop mapping. *Remote Sens. Environ.* **2021**, *264*, 112599. [[CrossRef](#)]
99. Wei, S.; Zhang, H.; Wang, C.; Wang, Y.; Xu, L. Multi-temporal SAR data large-scale crop mapping based on U-Net model. *Remote Sens.* **2019**, *11*, 68. [[CrossRef](#)]
100. Tamiminia, H.; Homayouni, S.; McNairn, H.; Safari, A. A particle swarm optimized kernel-based clustering method for crop mapping from multi-temporal polarimetric L-band SAR observations. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *58*, 201–212. [[CrossRef](#)]
101. Hamidi, M.; Safari, A.; Homayouni, S. An auto-encoder based classifier for crop mapping from multitemporal multispectral imagery. *Int. J. Remote Sens.* **2021**, *42*, 986–1016. [[CrossRef](#)]
102. Kwak, G.-H.; Park, N.-W. Two-stage Deep Learning Model with LSTM-based Autoencoder and CNN for Crop Classification Using Multi-temporal Remote Sensing Images. *Korean J. Remote Sens.* **2021**, *37*, 719–731.
103. Virnodkar, S.; Pachghare, V.K.; Murade, S. A Technique to Classify Sugarcane Crop from Sentinel-2 Satellite Imagery Using U-Net Architecture. In *Progress in Advanced Computing Intelligent Engineering*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 322–330.