

Article

Crop Mapping from Sentinel-1 Polarimetric Time-Series with a Deep Neural Network

Yang Qu ^{1,2,3}, Wenzhi Zhao ^{1,2,*} , Zhanliang Yuan ³ and Jiage Chen ⁴

¹ State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing Science and Engineering, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China; 211804020029@home.hpu.edu.cn

² Beijing Engineering Research Center for Global Land Remote Sensing Products, Institute of Remote Sensing Science and Engineering, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China

³ School of Surveying & Land Information Engineering, Henan Polytechnic University, Henan 454000, China; yuan6400@hpu.edu.cn

⁴ National Geomatics Center of China, Beijing 100830, China; jiagechen@ngcc.cn

* Correspondence: wenzhi.zhao@bnu.edu.cn

Received: 21 July 2020; Accepted: 30 July 2020; Published: 3 August 2020



Abstract: Timely and accurate agricultural information is essential for food security assessment and agricultural management. Synthetic aperture radar (SAR) systems are increasingly available in crop mapping, as they provide all-weather imagery. In particular, the Sentinel-1 sensor provides dense time-series data, thus offering a unique opportunity for crop mapping. However, in most studies, the Sentinel-1 complex backscatter coefficient was used directly which limits the potential of the Sentinel-1 in crop mapping. Meanwhile, most of the existing methods may not be tailored for the task of crop classification in time-series polarimetric SAR data. To solve the above problem, we present a novel deep learning strategy in this research. To be specific, we collected Sentinel-1 time-series data in two study areas. The Sentinel-1 image covariance matrix is used as an input to maintain the integrity of polarimetric information. Then, a depthwise separable convolution recurrent neural network (DSCRNN) architecture is proposed to characterize crop types from multiple perspectives and achieve better classification results. The experimental results indicate that the proposed method achieves better accuracy in complex agricultural areas than other classical methods. Additionally, the variable importance provided by the random forest (RF) illustrated that the covariance vector has a far greater influence than the backscatter coefficient. Consequently, the strategy proposed in this research is effective and promising for crop mapping.

Keywords: crop mapping; synthetic aperture radar; Sentinel-1; time-series; deep learning; covariance matrix

1. Introduction

Many of the problems resulting from the rapid growth of the global population are related to agricultural production [1,2]. In this context, it is necessary to have a comprehensive understanding of crop production information. Timely and accurate agricultural information can achieve a range of important purposes, such as improving agricultural production, ensuring food security, and facilitating ecosystem services valuation [3]. Remote sensing, which provides timely earth observation data with large spatial coverage, could serve as a convenient and reliable method for agricultural monitoring [4]. It is now possible to build a time-series image stack for full-season monitoring and differentiate crop types according to their unique seasonal features [5].

Over the past few decades, optical data has been regarded as the main earth observation strategy for crop monitoring [4]. The photosynthetic and optical properties of the plant leaves are used to distinguish between different crop types [6]. However, the acquisition of optical data depends heavily on clear sky conditions. In areas with high frequency cloud cover, it is difficult to get enough usable images [7], which hugely limits the application of dynamic crop monitoring [8]. Synthetic aperture radar (SAR) can collect data regardless of weather conditions, solving the main problem of optical sensors [9]. With the continuous development of SAR sensors, several studies have demonstrated that SAR data has great potential to distinguish various land cover types [10,11]. However, compared to optical data, SAR data has not been well used in agriculture [6].

Sentinel-1 provides high revisit frequency data and free access to historical archives, which greatly improves the availability of SAR time-series in agricultural monitoring [12]. Some efforts have been devoted to using Sentinel-1 data from the intensive time series for crop mapping and monitoring [13,14]. Nevertheless, these studies directly input the amplitude of the Sentinel-1 image (converted to dB scale), while neglecting the phase information. Phase information is unique to the SAR image, and it plays an important role in some retrieval applications (e.g., target recognition and classification). In particular, the phase information of the off-diagonal elements in the coherence/covariance matrix can characterize different land cover types [15]. Unfortunately, rich information and complex-valued data format make polarimetric synthetic aperture radar (PolSAR) image interpretations difficult.

Up to now, various methods have been developed to process and analyze PolSAR data. Some methods are based on the scattering mechanism of PolSAR data, such as Cameron decomposition, H/a/ α decomposition [16], Freeman decomposition [17], and so forth. These methods have strong physical interpretability. Unfortunately, Sentinel-1 is a dual-polarized SAR, and there are few methods applicable to dual-polarization information decomposition. In complex scenarios, such as agricultural land, it is not easy to use only one decomposition method to distinguish all crop types [18]. There are also some methods based on machine learning techniques for crop mapping, such as random forest (RF) [19], support vector machine (SVM), and AdaBoost [20]. These methods have strong universality, but the ability of feature extraction is limited [21]. For information extraction in a complex agricultural area, satisfactory results may not always be obtained. In short, most existing PolSAR feature extractors have limitations for Sentinel-1 images crop classification. Thus, it is urgent to develop a proper way of feature representations to make full use of the polarization information in Sentinel-1 data.

With the development of deep learning strategies, several solutions are provided for such tasks. The main attraction of deep learning strategies is that they can extract high-level features with an end-to-end learning strategy [5]. The deep learning-based image processing model represented by convolutional neural networks (CNNs) is often used to interpret SAR data, such as SAR image denoising [22], SAR target identification [23], and PolSAR image classification [24]. Zhang et al. [15] designed a complex-valued convolutional neural network (CV-CNN) to fit the complex-valued PolSAR data, where both the amplitude and phase information is used for image classification. In this work, the regular patches extracted from PolSAR image are regarded as the CNN input, thus the geometric features and topological relations within patches have been considered [25]. Moreover, due to the scattering properties of PolSAR data, there is also a coupling relationship between the phase information and directions received through transmitting. Therefore, some related studies proposing an extraction of the phase information by using a depthwise separable convolutional neural network [26] have achieved better results than those obtained through the conventional convolutional networks. Although these networks focused on spatial and polarization feature extraction, the temporal feature remains unexploited. Therefore, these methods may not be very suitable for research in agricultural areas.

The temporal feature is one of the most important indicators for crop classification because of its unique patterns through the temporal domain. For instance, the structural characteristics and water content of crops may vary greatly at different phenological stages with various crops. Recurrent neural networks (RNN) have the ability to analyze sequential data, and are often considered as the preferred choice to learn the temporal relationship for time series signal process [27]. Also, some studies have

demonstrated the advantages of using RNN as a temporal extractor compared with other methods. For example, Ndikumana et al. [14] designed the RNN framework to explore the temporal correlation with Sentinel-1 data for crop classification. Meanwhile, some studies have proposed combined methods by combining cyclic and convolution operations to process spatio-temporal cubes [28]. For instance, Rubwurm and Korner [29] designed a convolutional Recurrent model (convRNN) to tackle land cover classification in the Sentinel-2 time series. Compared to single model methods, the combined models generally provide better performance. Thus, it is necessary to develop a combined model that simultaneously considers the spatial-polarization-temporal features for time series SAR image classification.

In this study, we propose a Sentinel-1 time-series crop mapping strategy to further improve the classification accuracy. To serve this purpose, deep learning strategies were introduced to understand the spatial-temporal patterns and scattering mechanisms of crops. To be specific, we use the Sentinel-1 covariance matrix as the input vector to provide polarization features information for deep network training. Then, a novel depthwise separable convolution recurrent neural network (DSCRNN) architecture is proposed to better extract complex features from the Sentinel-1 time series, which integrates the operations of cyclic and convolution. Moreover, in order to better model the potential correlations from the phase information, the conventional convolution is replaced by the depthwise separable convolutions. The main contributions of this paper are:

1. By using the decomposed covariance matrix, the potential of the Sentinel-1 time series in crop discrimination is fully explored.
2. An effective crop classification method is proposed for time-series polarimetric SAR data by considering the temporal patterns of crop polarimetric and spatial characteristics.

The rest of this paper is organized as follows. Study areas and data are described in Section 2. Section 3 details the specific architecture and method of DSCRNN. Section 4 presents the results of the classification. The discussion and conclusion are presented in Sections 5 and 6, respectively.

2. Study Area and Data

2.1. Study Area

California is the largest agricultural state in the United States of America (U.S.) [30]. This indicates the significance of crop mapping in California. Thus, this study is carried out at two different sites in California, henceforth referred to as study area 1 and study area 2 (Figure 1).

Study area 1 is situated in Imperial, Southern California, at 33°01'N and 115°35'W, covering a region about 10 km × 10 km. The area is in the Colorado Desert, with a tropical desert climate, which is very hot. The area has one of the highest yields of crops such as alfalfa, onions, and lettuce in California. The average mean temperature is higher than 27 °C [31], and the temperature variation is also very large. There was little rain throughout the year, below the mean annual precipitation in the U.S. Six classes were selected for analysis: winter wheat, alfalfa, other hay/non-alfalfa, sugar beets, onions, and lettuce.

Study area 2 is situated in an agricultural district stretching over Solano and Yolo counties of California, Northern California, at 38°26'N and 121°44'W, covering a region about 10 km × 10 km. The area has a Mediterranean climate, characterized by dry hot summers and wet cool winters [32]. The region is flat, the agricultural system is complex, and one of the most productive agricultural areas in the U.S. It has an annual precipitation of about 750 mm, concentrated in the spring and winter [33]. Seven major crop types were selected for analysis: walnut, almond, alfalfa, winter wheat, corn, sunflower, and tomato.

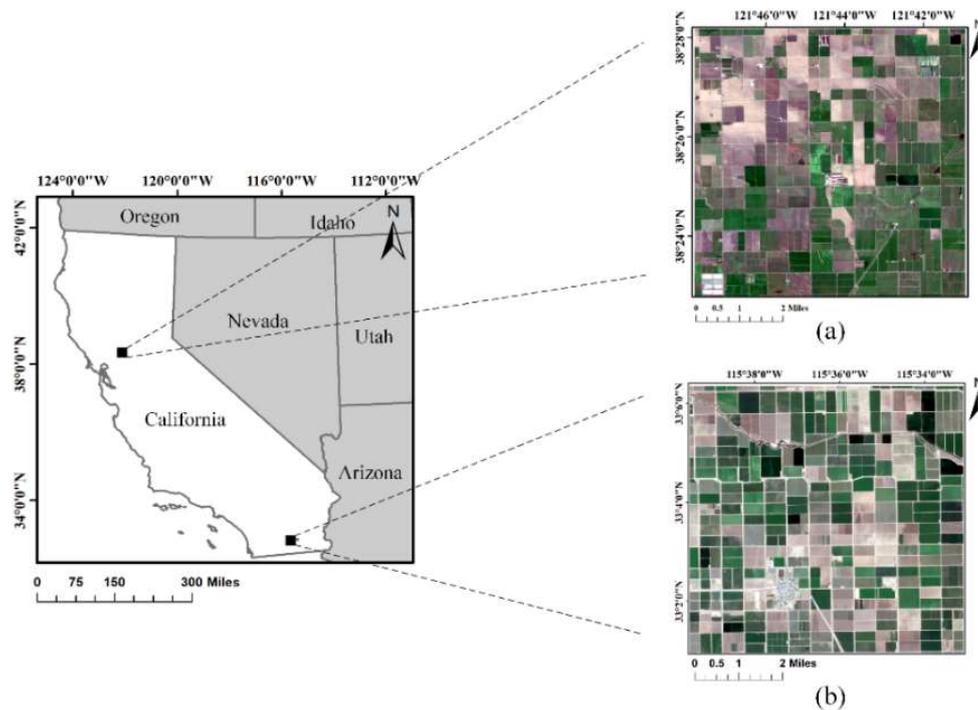


Figure 1. The study areas in California. The crop areas of interest (AOI) in study area 1 and 2 with true color composite of Sentinel-2. (a) study area 2 dated on 2019/02/10 (b) study area 1 on 2018/03/12.

2.2. Data

2.2.1. Sentinel-1 Data

In this study, the Sentinel-1 Interferometric Wide (IW) Single Look Complex (SLC) products were used. All the images were downloaded from the Sentinel-1 Scientific Data Hub. Since the major agricultural practices in both study areas were in spring and summer, we focused our data analysis on these seasons. Figure 2 shows the time distribution of Sentinel-1 images collected in the two study areas. In total, 15 scenes of Sentinel-1A images from 2018 were collected in the study area 1, and 11 Sentinel-1A images from 2019 were collected in study area 2.

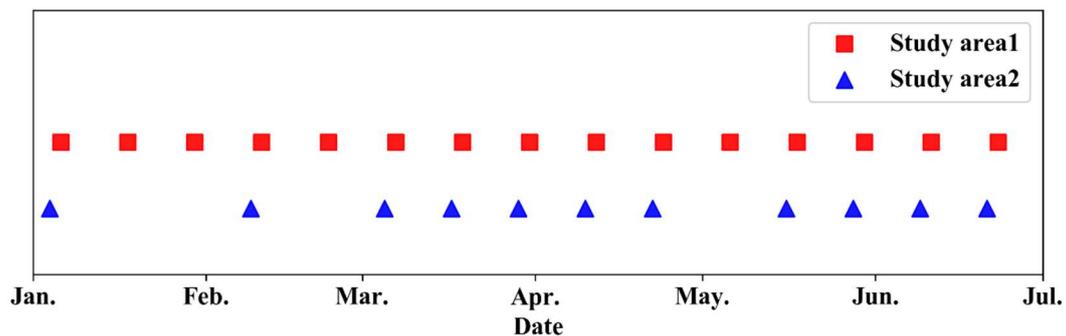


Figure 2. Data acquisition date in two study areas.

The pre-processing of time series Sentinel-1 images was done using the sentinel application platform (SNAP) offered by European Space Agency (ESA). Data preprocessing consists of five steps: (1) terrain observation by progressive scans synthetic aperture radar (TOPSAR) split, (2) calibrate Sentinel-1 data to the complex number, (3) debursting, (4) refined Lee filter, and (5) range-doppler terrain correction for all images using the same digital elevation model data (SRTM DEM 30 m).

Since we hoped that this study would facilitate the fusion of Sentinel-1 and Sentinel-2 features, we projected the data to the UTM reference and re-sampled it to 10 m to co-registration Sentinel-2.

Sentinel-1 backscatter images were generated to investigate the relative importance of input data to crop classification. The steps include: (1) thermal noise removal, (2) apply orbit file, (3) radiometrically calibrated to sigma0, (4) geocoding, and (5) backscatter images (δ) transform to the logarithmic dB scale [34].

2.2.2. Cropland Reference Data

The U.S. Department of Crop (USDA) Cropland Data Layer (CDL) of 2018 and 2019 was used as the reference data for crop classification and to test the experiment. The data is published regularly by the USDA and covers 48 states [35]. CDL has been widely used in all kinds of remote sensing crop research because of its high quality. However, there are some misclassifications in the data [36]. Through visual inspection, it was found that the misclassified pixels of CDL were concentrated at the boundary of the crop fields. Therefore, we performed a manual drawing of the reference data according to the CDL (Figure 3).

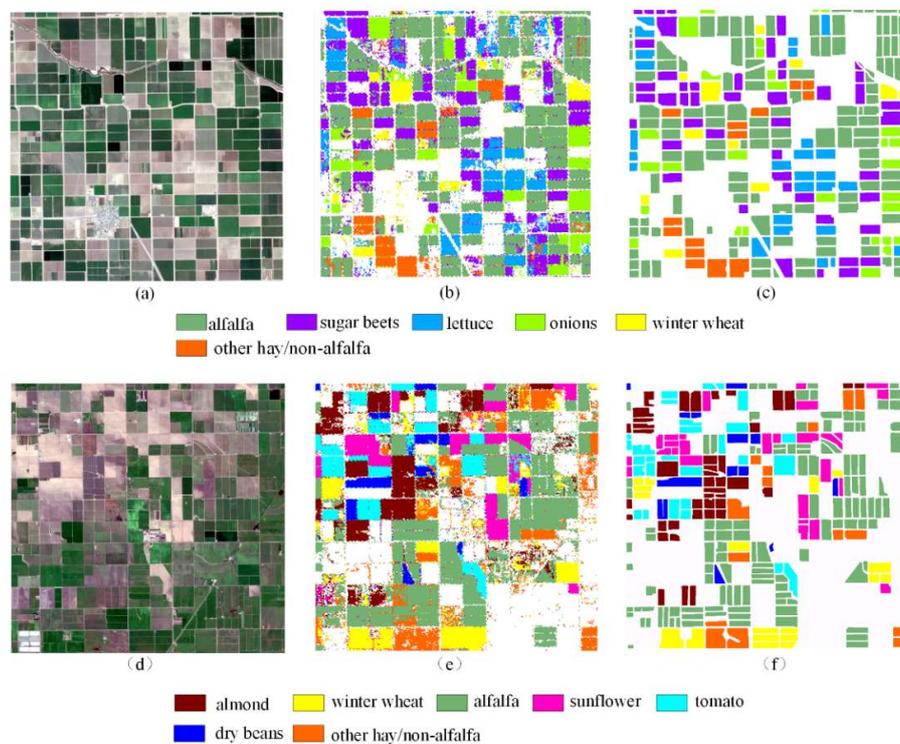


Figure 3. Ground truth maps of the study area. (a) RGB image of study area 1 from Sentinel-2 on 2018/03/12, (d) RGB image of study area 2 from Sentinel-2 on 2019/02/10. (b,e) the CDL data (major crop types) for 2018 and 2019, respectively. (c,f) manually labeled ground reference data.

The process of drawing labeled data consists of three steps. First, the spatial resolution of the CDL is resampled to 10 m before Sentinel-2 images overlaid on the CDL image to determine the crop field boundary. Secondly, the field of each major crop is manually delineated and buffered one pixel inward from the field boundary. Finally, fields of the same crop type are combined into a class. Detailed information about the modified labeled data are reported in Tables 1 and 2.

Table 1. Number of pixels for each of the 6 classes in study area 1.

Class	Number of Pixels
Alfalfa	240,969
Sugar beets	91,176
Lettuce	50,504
Onions	46,053
Winter wheat	20,627
Other hay	33,017

Table 2. Number of pixels for each of the 7 classes in study area 2.

Class	Number of Pixels
Almond	56,435
Winter wheat	34,308
Alfalfa	148,189
Sunflower	40,049
Tomato	44,277
Dry beans	19,718
Other hay	41,834

3. Methods

3.1. Representation of Sentinel-1 Data

PolSAR image can be represented by a 2×2 complex scattering matrix S . However, Sentinel-1 only provides dual-polarization information. Therefore, the expression of S needs to be modified. The backscattering matrix of Sentinel-1 is expressed as:

$$S = \begin{bmatrix} 0 & 0 \\ S_{VH} & S_{VV} \end{bmatrix}, \quad (1)$$

where S_{VH} and S_{VV} are backscattering coefficients under different polarimetric combinations. H and V represent the horizontal and vertical directions of the electromagnetic wave, respectively.

Since the scattering matrix S is an inadequate representation of the scattering characteristics of complex targets [37], the covariance matrix C is used. This is written as:

$$C_{dual} = \begin{bmatrix} \langle S_{VV}S_{VV}^* \rangle & \langle S_{VV}S_{VH}^* \rangle \\ \langle S_{VH}S_{VV}^* \rangle & \langle S_{VH}S_{VH}^* \rangle \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}, \quad (2)$$

where C_{11} , C_{12} , C_{21} , C_{22} are the members of the covariance matrix, and $*$ is the conjugate operation.

It can be seen from Equation (2) that the diagonal value of the matrix C_{dual} is real and the off-diagonal complex value. Since the matrix C_{dual} is a symmetric matrix, this means that $\{C_{11}, C_{12}, C_{22}\}$ contain all the information about C_{dual} . We separate the true and imaginary parts of C_{12} and convert them to real values. Thus, we get a 4-dimensional vector:

$$C_v = \{C_{11}, re(C_{12}), im(C_{12}), C_{22}\}, \quad (3)$$

where re and im represent the real and imaginary parts of complex numbers, respectively.

Finally, in order to accelerate the convergence of the model, each pixel is normalized. The equation is:

$$C_v[i] = \frac{C_v[i] - C_{v-min}[i]}{C_{v-max}[i] - C_{v-min}[i]}, \quad (4)$$

where i represents channel of C_v . C_{v-max} and C_{v-min} are the maximum and minimum values of i th channel, respectively.

3.2. Architecture of DSCRNN Network

Figure 4 shows the proposed DSCRNN architecture. In order to maintain the integrity of the Sentinel-1 data, the covariance matrix vectors are sliced into patches as the input of the neural network. Then, the patches of T timestamps are fed into the DSCRNN. In this step, the same depthwise separable convolution operation is performed for the patches of each timestamp to obtain the feature sequence. Finally, the attentive LSTM layer is input to produce the crop classification. Next, we introduce the components of the architecture and its advantages.

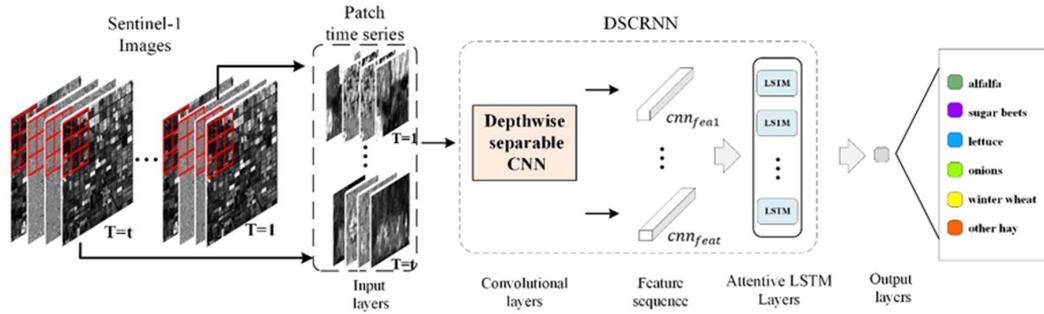


Figure 4. The general view of the proposed depthwise separable convolution recurrent neural network (DSCRNN).

3.2.1. Depthwise Separable Convolution

As shown in Figure 5, the convolution mechanism in the conventional CNNs extracts features from all dimensions of each image, including the spatial dimension and channel dimension [21]. For conventional CNNs, suppose the three-dimensional (3D) tensor $x \in \mathbb{R}^{H \times W \times D}$ input network, where H , W , and D are the height, width, and depth of the input. This is written as:

$$\text{Conv}(x, f)_{(i,j)} = \sum_{h,w,d}^{H,W,D} f_{h,w,d} \cdot x_{i+h,j+w,d}, \quad (5)$$

where f is the trainable parameter, (i, j) is the location of output feature maps, (h, w, d) is an element in x , which is in spatial location (h, w) and in the d 's channel.

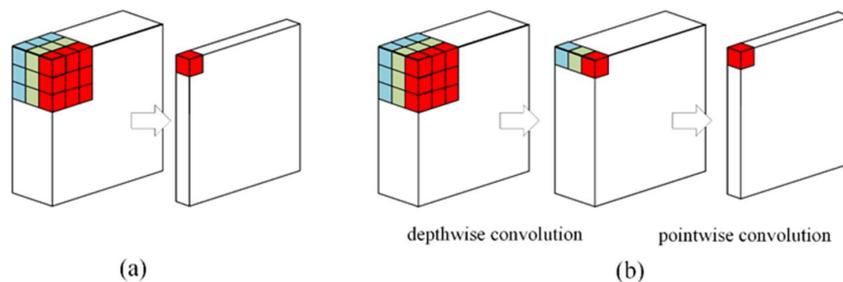


Figure 5. Representation of the comparison between (a) conventional convolutional neural networks (CNNs) and (b) depthwise separable CNN.

Depthwise separable convolution is successfully applied to Xception [26] and MobileNet [38]. Different from the conventional CNNs, the depthwise separable convolution can be divided into depthwise convolution and pointwise convolution. To be specific, the depthwise separable convolution convolves kernels of each filter with each input channel, and pointwise convolution [39]. This is written as:

$$\text{DConv}(x, f)_{(i,j)} = \sum_{h,w}^{H,W} f_{h,w} \cdot x_{i+h,j+w} \quad (6)$$

$$\text{PConv}(x, f)_{(i,j)} = \sum_d^D f_d \quad (7)$$

where DConv is the depthwise convolution, PConv is the pointwise convolution, and f_d represents a convolution filter of size 1. Compared with the conventional CNNs, the parameters of the depthwise separable convolution are significantly reduced [40].

For the data which is closely related between channels, the depthwise separable convolution may yield better results [41]. C_{dual} matrix contains the phase information and amplitude information. This means that the correlations between multi-channel can express the structure information of the crop. Therefore, depthwise separable convolution is more suitable for feature extraction in PolSAR images than conventional CNNs [41].

3.2.2. Attentive Long Short-Term Memory Neural Network

LSTM is one representative RNN architecture with the ability to maintain a temporal state between continuous input data and it learns from long-term context dependencies [42]. Compared with RNN, the inner structure of the hidden layer in LSTM is more complex [43]. An LSTM block consists of a memory cell state, forget gate, input gate, and output gate. The specific steps of LSTM at time t are as follows:

The previous cell state of C_{t-1} is passed to the forget gate F_t , and the sigmoid activation function is used to determine the proportion of discarded information. This can be represented as:

$$F_t = \text{Sigmoid}(W_{F_x}x_t + W_{F_h}h_{t-1} + \text{bias}_F) \quad (8)$$

$$\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (9)$$

Then, through the input gate I_t , it decides the percentage of the new information \tilde{C}_t is stored in cell state C_t for input x , where the input gate I_t should be updated. This is written as:

$$I_t = \text{Sigmoid}(W_{I_x}x_t + W_{I_h}h_{t-1} + \text{bias}_I) \quad (10)$$

$$\tilde{C}_t = \text{tanh}(W_{C_x}x_t + W_{C_h}h_{t-1} + \text{bias}_C) \quad (11)$$

$$\text{tanh}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (12)$$

Update the present cell state C_t based on multiplying the cell state C_{t-1} of the previous step by F_t and the updated information \tilde{C}_t by I_t . This can be represented as follow:

$$C_j = F_t * C_{t-1} + I_t * \tilde{C}_t \quad (13)$$

Finally, confirm the new hidden state h in the output gate O_t , where the new cell state C_j is used. This can be written as:

$$O_t = \text{Sigmoid}(W_{O_x}x_t + W_{O_h}h_{t-1} + \text{bias}_O), \quad (14)$$

$$h_t = O_t * \text{tanh}(C_t), \quad (15)$$

where $W_{F_x}, W_{F_h}, W_{I_x}, W_{I_h}, W_{O_x}, W_{O_h}, W_{C_x}, W_{C_h}$ are the weight matrices and bias is the trainable bias term.

Finally, we couple LSTM with an attention mechanism that can connect the information extracted by the recursive neural network model in different time-lapse. Intuitively, the attention mechanism

supports the model to pay attention to specific time stamps and discard useless contextual information. This is written as:

$$rnn_{feat} = \sum_{j=1}^N softmax(\tanh(x_t, f)) \times h_j. \tag{16}$$

where x_t is the input vector at time t , h_j is the output vector at time j , and f is the set of all trainable parameters. The purpose of this step is to learn a set of weights to measure the importance of the temporal information.

As discussed in Section 3.2.1, a 3D tensor is an input into the convolution network to obtain a feature vector cnn_{feat} . In this way, the Sentinel-1 time-series data can be regarded as a 4D tensor $x \in \mathbb{R}^{H \times W \times D \times T}$ where T is the temporal dimension. This means that for each individual patch, the output feature returned by the depthwise separable convolution model is $Seq_{feat}^{cnn}(cnn_{feat1}, cnn_{feat2}, \dots, cnn_{feat})$, where the feature map cnn_{feat} represents the feature vectors of depthwise separable convolution at time step t , and each output feature has the same dimensions. Then, the convolved feature sequence Seq_{feat}^{cnn} is inputted into the attentive LSTM layer, and output a feature vector rnn_{feat} . Finally, with the feature vector rnn_{feat} , labels can be further assigned using the softmax classifier, as follow:

$$p_i = softmax(rnn_{feat})_i = \frac{e^{rnn_{feat}}}{\sum_{j=1}^n e^{rnn_{feat}}}, \tag{17}$$

where p_i is the probability that rnn_{feat} belonging to class i .

3.3. Competing Methods

In order to evaluate the performance of DSCRNN, several classification methods such as SVM, RF, Conv1D, and LSTM were selected for comparison. In addition, in order to investigate the interplay among the different components of DSCRNN, we disentangle the different parts of our framework. They are simply recorded as Network (Net) A to Net C in the order of appearance for convenience: Net A is based on a conventional CNN as shown in Figure 6a, Net B is based on depthwise separable convolution as shown in Figure 6b, and Net C is based on a CRNN (using attentive LSTM and conventional CNN) as shown in Figure 6c.

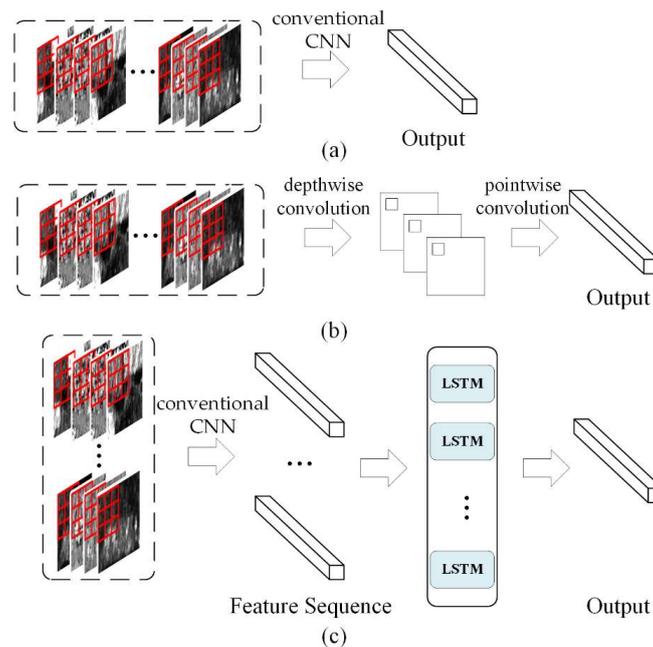


Figure 6. Classification of (a) Net A, (b) Net B, and (c) Net C.

3.4. Dataset Partition

In crop classification tasks, the labeled data is usually very limited. Therefore, according to the modified ground truth data, we randomly select 1% of all available samples for each crop type as the training set. For a model with a single-pixel as the input (e.g., RF, Conv1D, and LSTM), the time series of the pixel corresponding to the labeled data is used as the input vector. For the DSCRNN and variant model, we take a sample (labeled data) as the center point, and then we segment a square patch with the size of 18×18 as input. The remaining non-overlapping samples for each study area are used for testing (Table 3). It is important to note that there is no overlap between training and testing [44].

Table 3. The detail information of two study areas used in experiments.

	Training Samples	Testing Samples
Study area1	4837	16,124
Study area2	3707	10,131

3.5. Experimental Designs

In this study, a carefully designed DSCRNN model is applied. In order to make full use of the rich information of the Sentinel-1 time series, the covariance matrix is used as an input vector to train the model. The main crop vectors from the manually drawn reference data are identified to guide the sample extraction for Sentinel-1 data. Take the study area 1 as an example; the sizes of input samples are set to $15 \times 18 \times 18 \times 4$. For the first depthwise separable convolution layer, the inputs are converted to $15 \times 16 \times 16 \times 32$. In this step, firstly, four $3 \times 3 \times 1$ convolutional kernels are used to convolute every single channel of the input map to obtain four $15 \times 16 \times 16$ feature maps. Secondly, thirty-two $1 \times 1 \times 4$ convolution kernels are convoluted and generate $15 \times 16 \times 16 \times 32$ outputs. Then, the second depthwise separable convolution layer has sixty-four convolutional kernels to obtain a $15 \times 14 \times 14 \times 64$ output map, which is then downsampled to $15 \times 7 \times 7 \times 64$ with the max pooling operations. After the max pooling layer, the output feature map is flattened. Next, the temporal features can be extracted by the attentive LSTM with 150 hidden units. Finally, a softmax classifier outputs the 6 different labels. In the training stage, the Adam optimizer [45] was used and fixed as: learning rate = 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 1 \times 10^{-7}$ [5], and the batch size was set to 200. In all the following experiments, the patch size is set to 18.

The Conv1D model consists of one convolution layer with 512 filters, one max-pooling layer, and one fully connection layer. The LSTM model consists of two hidden layers with 150 units for each layer. Net A, Net B, and Net C all have a similar architecture to DSCRNN, except for the differences in some key components in DSCRNN (i.e., depthwise separable convolution and attentive LSTM). The neural net models are implemented using the Python TensorFlow library, while other models are implemented using the Python Scikit-learn library [46].

Considering leveraging the RF and SVM classifier, we optimize RF via tuning the number of trees in the forest and the maximum depth of each tree, and the SVM by adjusting C and gamma. We employ a “grid search” strategy to select the optimal parameters of the classifier: the classifier is repeatedly trained many times to select the optimal combination of parameter values. The number of trees values in the set {200,400,600,800,1000}, the maximum depth of each tree in the range {20,40,60,80,100}, C in the set {0.001,0.01,0.1,1,10,100,1000,3000,5000,10,000}, and gamma in the set {0.1, 1,2,5,10}. The average accuracy (AA) [47], overall accuracy (OA), kappa coefficient (kappa), and F1-score are used as the criteria for evaluating the performance of the models.

4. Results

4.1. Classification Results

In this section, we describe and discuss the experimental results obtained on the two study date sets introduced in Section 2. We evaluate the performance of DSCRNN and then compare it with several competing methods.

4.1.1. Results on Study Area 1

The results of crop classification in study area 1 obtained by different methods are shown in Figure 7. Obviously, for competitive methods such as SVM, RF, Conv1D, and LSTM, there is a lot of speckle noise in the classification maps, which results in a low accuracy for crop mapping. However, the spatial feature-based methods, especially DSCRNN, produced a noise-proof classification map. Table 4 lists the detailed accuracy assessment of these classification methods. It can be easily noticed that the overall classification accuracy of DSCRNN (0.9603) is higher than that of LSTM (0.8744), RF (0.8911), Conv1D (0.8998), Net A (0.9092), and the differences between the Net B (0.9477), Net C (0.9486), and DSCRNN are not significant. However, it can be found that the AA, OA, Kappa, and F1-score of DSCRNN are slightly higher than Net B and Net C. From the tables, the overall accuracy of Net B (OA: 0.9477) is significantly higher than that of Net A (OA: 0.9092), which indicates that the depthwise separable convolution can achieve a better classification than conventional CNN in study area 1. This is because depthwise separable convolution can improve the ability to extract information from the phase of Sentinel-1 images. Similarly, the OA of Net C is slightly higher than Net A. This result confirms the importance of introducing time series temporal information into the SAR image classification.

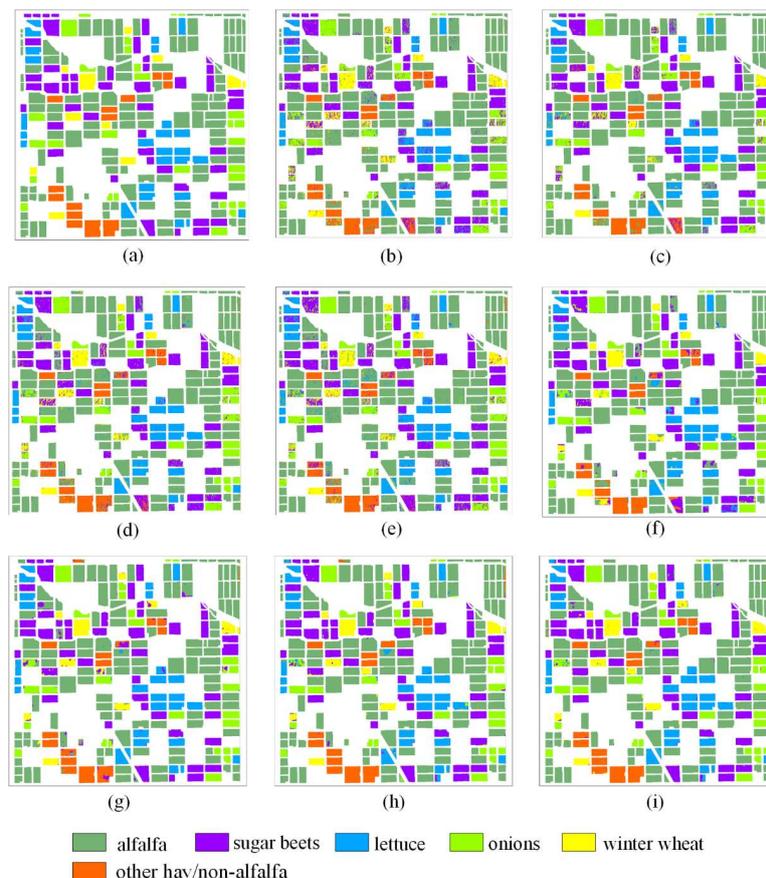


Figure 7. Maps of study area 1: (a) Ground truth map, (b) SVM, (c) RF, (d) Conv1D, (e) LSTM, (f) Net A, (g) Net B, (h) Net C, and (i) DSCRNN.

Table 4. Results of study area 1.

Class	SVM	RF	Conv 1D	LSTM	Net A	Net B	Net C	DSCR NN
alfalfa	0.8695	0.8457	0.9004	0.8859	0.9436	0.9535	0.9509	0.9537
sugar beets	0.8858	0.9652	0.9380	0.9127	0.9164	0.9455	0.9821	0.9838
lettuce	0.9164	0.9700	0.8667	0.7978	0.8346	0.9977	0.9535	0.9687
onions	0.9158	0.9611	0.9410	0.9377	0.9960	0.9766	0.9443	0.9818
winter wheat	0.8169	0.9827	0.8161	0.8284	0.6948	0.8441	0.9449	0.9947
other hay	0.7459	0.9016	0.8021	0.6958	0.8273	0.8434	0.8270	0.8539
AA	0.8584	0.9377	0.8774	0.8430	0.8688	0.9268	0.9338	0.9561
OA	0.8735	0.8911	0.8998	0.8744	0.9092	0.9477	0.9486	0.9603
Kappa	0.8069	0.8297	0.8487	0.8110	0.8757	0.9284	0.9296	0.9456
F1_score	0.8709	0.8876	0.8982	0.8733	0.9086	0.9477	0.9483	0.9601

4.1.2. Results on Study Area 2

The classification results of study area 2 with various classification methods are shown in Figure 8. The characteristics of the different crop types in study area 2 are very similar, which elevates the difficulty of performing an accurate classification. Therefore, the classification performances of all methods decrease compared with those of study area 1. Table 5 reports the classification accuracy of each method; it can be seen that DSCRNN still has a high overall classification accuracy (0.9389). It is worth noting that the Kappa of DSCRNN is 2.79% and 2.59% higher than that of Net B and Net C. This shows that the combination of depthwise separation convolution and CRNN makes more improvements in crop classification than using CRNN or depthwise separation convolution alone. In the comparison with Net A and Net B, it can be seen that for the same architecture (CNN and depthwise separable convolution), depthwise separable convolution improves the accuracy of alfalfa (0.9096 vs. 0.9575) significantly. Similarly, in the comparison between Net A and Net C, we can find that the context information for time series remarkable improvement in recognizing other hay (0.7766 vs. 0.8912). In addition, DSCRNN has a good recognition performance for alfalfa and other hay (0.9634 and 0.9549, respectively), which benefits from the utilization of the phase and temporal information of time series data.

Table 5. Classification results of study area 2.

Class	SVM	RF	Conv 1D	LSTM	Net A	Net B	Net C	DSCR NN
almond	0.7726	0.6887	0.8702	0.8418	0.8876	0.8456	0.8443	0.8736
winter wheat	0.6871	0.8487	0.8428	0.7229	0.9383	0.9274	0.9608	0.9403
alfalfa	0.8357	0.8122	0.8690	0.8771	0.9096	0.9575	0.9807	0.9634
sunflower	0.8950	0.8532	0.8859	0.9029	0.9254	0.9652	0.9716	0.9524
tomato	0.7467	0.6818	0.7836	0.8088	0.8744	0.8249	0.7913	0.9706
dry beans	0.7710	0.9492	0.7435	0.8177	0.8750	0.9126	0.8568	0.8707
other hay	0.5110	0.7252	0.7556	0.7700	0.7766	0.7623	0.8912	0.9036
AA	0.7455	0.7942	0.8215	0.8202	0.8838	0.8851	0.8925	0.9249
OA	0.8027	0.7894	0.8412	0.8378	0.8983	0.9110	0.9130	0.9389
Kappa	0.7337	0.7133	0.7960	0.7926	0.8640	0.8823	0.8850	0.9191
F1_score	0.7970	0.7760	0.8396	0.8359	0.8966	0.9103	0.9121	0.9390

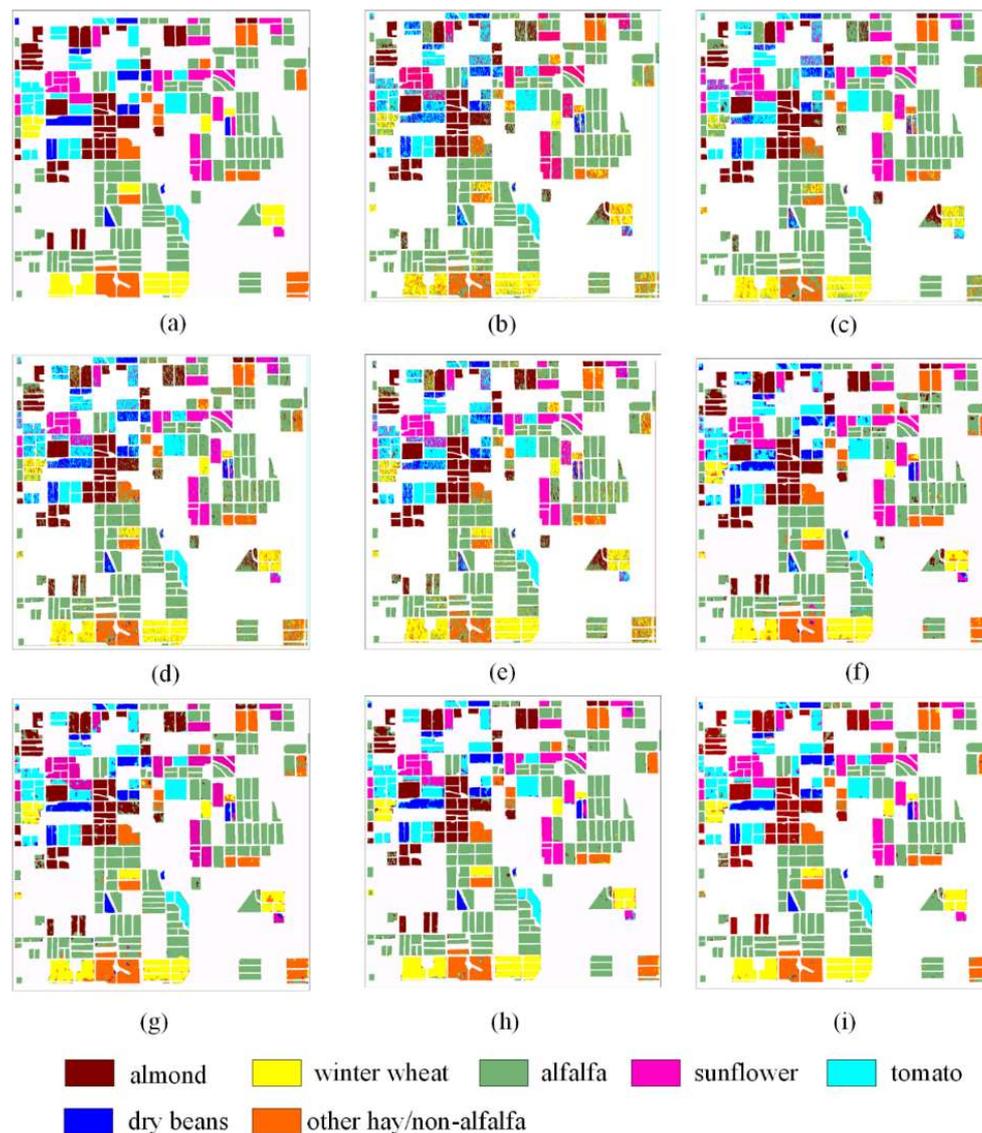


Figure 8. Maps of study area 2: (a) Ground truth map, (b) SVM, (c) RF, (d) Conv1D, (e) LSTM, (f) Net A, (g) Net B, (h) Net C, and (i) DSCRNN.

4.2. Influence of Different Input Data

In this section, experiments are implemented with different input data to verify the improvement when using the covariance matrix of Sentinel-1 images instead of backscattering (VV and VH). The common methods (RF, Net A, and Net B) with the input of amplitude are abbreviated as RF-v1, Net A-v1, and Net B-v1, while the models using covariance matrix input are noted as RF-v2, Net A-v2, and Net B-v2. Experiments are carried out on the Study area 1.

The results are reported in Table 6. In this study area 1, the RF with the covariance matrix as input always shows better performances, which confirms our hypothesis that the phase information indeed provides more useful information for the crop classification task. Clearly, the performance of Net B-v2 is much better than that of Net B-v1 which demonstrates that the depthwise separable convolution is helpful for extracting the underlying correlations of the phase information. However, Net A-v1 and Net A-v2 have similar classification accuracy. This means that conventional convolution has limited ability to extract information from the SAR phase.

Table 6. Results of study area 1 (v1 is amplitude data and v2 is covariance matrix data).

	AA	OA	Kappa	F1_Score
RF-v1	0.9051	0.8652	0.7875	0.8592
Net A-v1	0.8818	0.8953	0.8580	0.8968
Net B-v1	0.8821	0.9004	0.8640	0.9009
RF-v2	0.8914	0.8843	0.8256	0.8795
Net A-v2	0.8688	0.9092	0.8757	0.9086
Net B-v2	0.9268	0.9477	0.9284	0.9477

5. Discussion

5.1. Phase Information Importance

In this section, we discuss the contribution of the covariance matrix for crop classification. For most of the previous researches on Sentinel-1 images crop mapping, it is common to only consider amplitude information extraction and neglect the unique phase information of SAR. However, the complex-valued polarization scattering matrix can extract useful phase information, and thus generate accurate descriptions of crop type.

The classification results in Section 4 demonstrated that the RF with the covariance matrix as input has greater overall classification accuracies than backscatter features. It should be noted that the classification accuracy of the conventional convolution on the two data sets is similar (0.8953 and 0.9092, respectively). One possible reason may be that there is a significant difference between the phase information and the amplitude information, so the phase information may not be fully utilized when using conventional amplitude image classification methods (e.g., conventional CNN).

Further, to validate the importance of feature representations [48], the RF classifier with all available features (90 features) is utilized to investigate the importance of the input features for crop classification. For visual comparison, we add up the contribution score of features to represent the importance of each feature and list them through the temporal axis in Figure 9. It is clear from Figure 8 that the features derived from the covariance matrix are generally more important compared to the backscatter ones. This suggests that neglected phase information can be used to identify different types of scatterers. Also, for the images collected in January, the covariance matrix is the most important feature representation with the importance of 0.0653 and 0.0935, respectively. In particular, for the entire time series data dataset, the images collected from January to March, the covariance matrix has the most important impact on crop classification. This indicates that the largest separability amongst crop types in study area 1 occurred during this period. In contrast, for the May to June time series imagery, the phase information has limited influence on crop identification. It is important to note that the more time series images are collected, the less important the phase information may be.

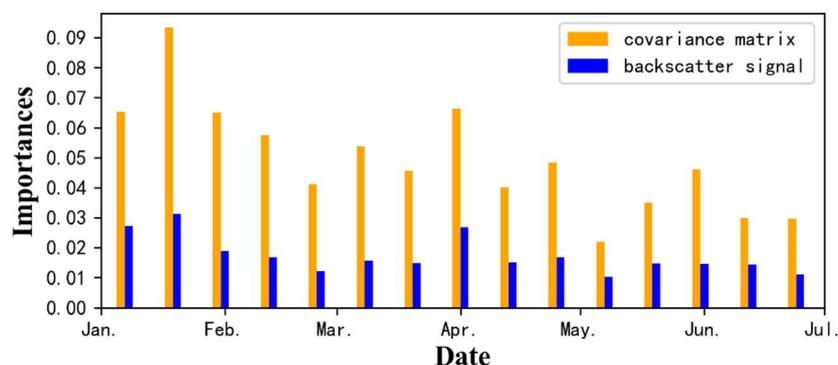


Figure 9. Importance validation with random forest (RF) classifier. The sum of the importance of all variables represented by each bar. For example, the first orange bar represents the sum of the importance of the four variables in the Sentinel-1 covariance matrix on January 5.

5.2. Pros and Cons

In this work, we demonstrated that combining phase information and amplitude information from the Sentinel-1 time-series data can be used to classify crops in complex agricultural areas. Most of the time, for the Sentinel-1 covariance matrix images, both the classical method and the model proposed in this study obtain good classification accuracy. It is worth noting that the proposed DSCRNN has the highest overall classification accuracy in two study areas—AA, OA, Kappa, and F1-score of DSCRNN are above 0.91 for both study areas. Moreover, we explored the contributions of depthwise separable convolution and Attentive LSTM to the Sentinel-1 images classification to demonstrate the robustness of the DSCRNN. To be specific, in the comparison between Net A and Net B, it can be found that the overall accuracy of Net B is about 5 and 2 percentage points higher than that of Net A for both two study areas. Similarly, from the classification results of Net A and Net C, Attentive LSTM extracts temporal features from Sentinel-1 images are beneficial to the recognition of some complex crops (such as sugar beets). In terms of Net B and Net C, the difference between them is not obvious in study area 1. However, in study area 2, it is clear that Net C is superior to Net B. It seems that differences are related to the complexity of the study area. This indicates that for complex agricultural areas, it is difficult to fully describe the unique structure of crops by scattering characteristics alone, which can result in false recognition. In contrast, the unique growth patterns of crops are much easier to be distinguished than their structures.

In addition, some limitations of the proposed methodology must be stressed. First, the impacts of inaccurate labeled data. The classification method relies heavily on ground-truth maps. This means that the low quality of labeled data will have a negative impact on the performances. Therefore, we visually checked the CDL data and manual labeling the ground truth maps for higher accuracy. Another weakness of DSCRNN is the complexity of the model with a high computational cost.

5.3. Potential Applications

In this section, we summarize some of the studies in this paper and their potential impact on collaborative Sentinel-1 and Sentinel-2 data for crop mapping. The analysis of Sentinel-1 time series data led to recommendations for Sentinel-1 data selection in the synergy use of these two time-series. Our results have shown that the classification results with the Sentinel-1 covariance matrix as the input are better than the backscatter images, suggesting that much more valuable information is provided by the covariance matrix. As such, it is possible to use the covariance matrix data to replace the commonly used backscatter images to coordinate with the Sentinel-2 data. The analysis of DSCRNN classification results led to recommendations for Sentinel-1 branch selection in custom architectures. To be specific, in some collaborative Sentinel-1 and Sentinel-2 studies, the use of well-designed custom networks has been proposed (e.g., the two-branch architecture). With the help of two network branches, different types of Sentinel-2 and Sentinel-1 information can be processed differently to obtain the information on the two branches). DSCRNN has shown strong competitiveness in the experiments of the two study areas. Therefore, DSCRNN as a Sentinel-1 branch in two-branch networks may be of great help to improve crop classification performance.

6. Conclusions

In this study, we proposed a combined strategy for crop classification using Sentinel-1 time series data. Different from previous studies, the Sentinel-1 time series stack was replaced by the complex-valued covariance matrix instead of the commonly used backscatter signals. In this way, the original information of Sentinel-1 images could be effectively retained. Moreover, we proposed the DSCRNN architecture to characterization of crop type from multiple perspectives (spatial characteristics, phase correlation, and temporal information). The architecture utilized depthwise separable convolution to better formulate the potential correlation of the phase and spatial information of Sentinel-1 images. On this basis, we further introduced the Attentive LSTM into the network to extract

the temporal relationship from the feature sequences. Compared to previous studies, the proposed method provided an accurate crop mapping result even with complex crop areas. For future works, we will focus on the combination of Sentinel-1 and Sentinel-2 data to boost crop mapping accuracy.

Author Contributions: Y.Q. and W.Z. developed the main idea that led to this paper. Y.Q. and W.Z. provided Sentinel-1 processing, classifications and their descriptions. Z.Y. and J.C. helped with the experiments and results analysis. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the National Key Research and Development Program of China (Grant No. 2018YFC1508903), and the Fundamental Research Funds for the Central Universities (Grant No. 2018NTST01).

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

References

- Godfray, H.C.J.; Beddington, J.R.; Crute, I.R.; Haddad, L.; Lawrence, D.; Muir, J.F.; Pretty, J.; Robinson, S.; Thomas, S.M.; Toulmin, C. Food security: The challenge of feeding 9 billion people. *Science* **2010**, *327*, 812–818. [[CrossRef](#)] [[PubMed](#)]
- Johnson, J.A.; Runge, C.F.; Senauer, B.; Foley, J.; Polasky, S. Global agriculture and carbon trade-offs. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 12342–12347. [[CrossRef](#)] [[PubMed](#)]
- Thenkabail, P.S.; Knox, J.W.; Ozdogan, M.; Gumma, M.K.; Congalton, R.G.; Wu, Z.; Milesi, C.; Finkral, A.; Marshall, M.; Mariotto, I. Assessing future risks to agricultural productivity, water resources and food security: How can remote sensing help? *Photogramm. Eng. Remote Sens.* **2012**, *78*, 773–782.
- Jiao, X.; Kovacs, J.M.; Shang, J.; McNairn, H.; Walters, D.; Ma, B.; Geng, X. Object-oriented crop mapping and monitoring using multi-temporal polarimetric RADARSAT-2 data. *ISPRS J. Photogramm. Remote Sens.* **2014**, *96*, 38–46. [[CrossRef](#)]
- Zhong, L.; Hu, L.; Zhou, H. Deep learning based multi-temporal crop classification. *Remote Sens. Environ.* **2019**, *221*, 430–443. [[CrossRef](#)]
- 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](#)]
- Griffiths, P.; Nendel, C.; Hostert, P. Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. *Remote Sens. Environ.* **2019**, *220*, 135–151. [[CrossRef](#)]
- Sonobe, R.; Tani, H.; Wang, X.; Kobayashi, N.; Shimamura, H. Random forest classification of crop type using multi-temporal TerraSAR-X dual-polarimetric data. *Remote Sens. Lett.* **2014**, *5*, 157–164. [[CrossRef](#)]
- Skriver, H. Crop classification by multitemporal C-and L-band single-and dual-polarization and fully polarimetric SAR. *IEEE Trans. Geosci. Remote Sens.* **2011**, *50*, 2138–2149. [[CrossRef](#)]
- Ullmann, T.; Schmitt, A.; Roth, A.; Duffe, J.; Dech, S.; Hubberten, H.; Baumhauer, R. Land cover characterization and classification of arctic tundra environments by means of polarized synthetic aperture X- and C-band radar (PolSAR) and landsat 8 multispectral imagery — richards island, Canada. *Remote Sens.* **2014**, *6*, 8565–8593. [[CrossRef](#)]
- Zhang, X.; Dierking, W.; Zhang, J.; Meng, J. A polarimetric decomposition method for ice in the Bohai Sea using C-band PolSAR data. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2014**, *8*, 47–66. [[CrossRef](#)]
- Inglada, J.; Vincent, A.; Arias, M.; Marais-Sicre, C. Improved early crop type identification by joint use of high temporal resolution SAR and optical image time series. *Remote Sens.* **2016**, *8*, 362. [[CrossRef](#)]
- Navarro, A.; Rolim, J.; Miguel, I.; Catalão, J.; Silva, J.; Painho, M.; Vekerdy, Z. Crop monitoring based on SPOT-5 Take-5 and sentinel-1A data for the estimation of crop water requirements. *Remote Sens.* **2016**, *8*, 525. [[CrossRef](#)]
- Ndikumana, E.; Minh, D.H.T.; Baghdadi, N.; Courault, D.; Hossard, L. Deep recurrent neural network for agricultural classification using multitemporal SAR Sentinel-1 for camargue, France. *Remote Sens.* **2018**, *10*, 1217. [[CrossRef](#)]
- Zhang, Z.; Wang, H.; Xu, F.; Jin, Y. Complex-valued convolutional neural network and its application in polarimetric SAR image classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 7177–7188. [[CrossRef](#)]
- Cloude, S.R.; Pottier, E. An entropy based classification scheme for land applications of polarimetric SAR. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 68–78. [[CrossRef](#)]

17. Freeman, A.; Durden, S.L. A three-component scattering model for polarimetric SAR data. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 963–973. [[CrossRef](#)]
18. Chen, Y.; He, X.; Xu, J.; Zhang, R.; Lu, Y. Scattering feature set optimization and polarimetric SAR classification using object-oriented RF-SFS algorithm in coastal wetlands. *Remote Sens.* **2020**, *12*, 407. [[CrossRef](#)]
19. Loosvelt, L.; Peters, J.; Skriver, H.; Lievens, H.; van Coillie, F.; de Baets, B.; Verhoest, N. Random Forests as a tool for estimating uncertainty at pixel-level in SAR image classification. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *19*, 173–184. [[CrossRef](#)]
20. She, X.; Yang, J.; Zhang, W. The boosting algorithm with application to polarimetric SAR image classification. In Proceedings of the 2007 1st Asian and Pacific Conference on Synthetic Aperture Radar, Huangshan, China, 5–9 November 2007; pp. 779–783.
21. Shang, R.; He, J.; Wang, J.; Xu, K.; Jiao, L.; Stolkin, R. Dense connection and depthwise separable convolution based CNN for polarimetric SAR image classification. *Knowl. Based Syst.* **2020**. [[CrossRef](#)]
22. Wang, P.; Zhang, H.; Patel, V.M. SAR image despeckling using a convolutional neural network. *IEEE Signal Process. Lett.* **2017**, *24*, 1763–1767. [[CrossRef](#)]
23. Chen, S.; Wang, H.; Xu, F.; Jin, Y.-Q. Target classification using the deep convolutional networks for SAR images. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 4806–4817. [[CrossRef](#)]
24. Zhou, Y.; Wang, H.; Xu, F.; Jin, Y.-Q. Polarimetric SAR image classification using deep convolutional neural networks. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 1935–1939. [[CrossRef](#)]
25. Li, L.; Ma, L.; Jiao, L.; Liu, F.; Sun, Q.; Zhao, J. Complex contourlet-CNN for polarimetric SAR image classification. *Pattern Recognit.* **2019**, *194*, 107110. [[CrossRef](#)]
26. Chollet, F. Xception: Deep learning with depth wise separable convolutions. In Proceedings of the Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1800–1807.
27. 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.* **2019**, *57*, 924–935. [[CrossRef](#)]
28. Shi, X.; Chen, Z.; Wang, H.; Yeung, D.; Wong, W.; Woo, W. Convolutional LSTM Network: A machine learning approach for precipitation nowcasting. In Proceedings of the Neural Information Processing Systems, Montreal, Quebec, Canada, 8–13 December 2014; pp. 802–810.
29. Rubwurm, 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 Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1496–1504.
30. Pathak, T.B.; Maskey, M.L.; Dahlberg, J.A.; Kearns, F.; Bali, K.M.; Zaccaria, D. Climate change trends and impacts on California agriculture: A detailed review. *Agronomy* **2018**, *8*, 25. [[CrossRef](#)]
31. Belgiu, M.; Csillik, O. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sens. Environ.* **2017**, *204*. [[CrossRef](#)]
32. Zhong, L.; Gong, P.; Biging, G.S. Phenology-based crop classification algorithm and its implications on agricultural water use assessments in California’s Central Valley. *Photogramm. Eng. Remote Sens.* **2012**, *78*, 799–813. [[CrossRef](#)]
33. Dyer, A.R.; Rice, K.J. Effects of competition on resource availability and growth of a California bunchgrass. *Ecology* **1999**, *80*, 2697–2710. [[CrossRef](#)]
34. Li, M.; Bijker, W. Vegetable classification in Indonesia using Dynamic Time Warping of Sentinel-1A dual polarization SAR time series. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *78*, 268–280. [[CrossRef](#)]
35. Boryan, C.G.; Yang, Z.; Mueller, R.; Craig, M. Monitoring US agriculture: The US department of agriculture, national agricultural statistics service, cropland data layer program. *Geocarto Int.* **2011**, *26*, 341–358. [[CrossRef](#)]
36. Li, H.; Zhang, C.; Zhang, S.; Atkinson, P.M. Full year crop monitoring and separability assessment with fully-polarimetric L-band UAVSAR: A case study in the Sacramento Valley, California. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *74*, 45–56. [[CrossRef](#)]
37. Liu, F.; Jiao, L.; Hou, B.; Yang, S. POL-SAR image classification based on Wishart DBN and local spatial information. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 3292–3308. [[CrossRef](#)]
38. Howard, A.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861. Available online: <https://arxiv.org/abs/1704.04861> (accessed on 17 April 2017).

39. Kamal, K.C.; Yin, Z.; Wu, M.; Wu, Z. Depthwise separable convolution architectures for plant disease classification. *Comput. Electron. Agric.* **2019**, *165*, 104948.
40. Zhang, T.; Zhang, X.; Shi, J.; Wei, S. Depthwise separable convolution neural network for high-speed SAR ship detection. *Remote Sens.* **2019**, *11*, 2483. [[CrossRef](#)]
41. Zhang, L.; Dong, H.; Zou, B. Efficiently utilizing complex-valued PolSAR image data via a multi-task deep learning framework. *ISPRS J. Photogramm. Remote Sens.* **2019**, *157*, 59–72. [[CrossRef](#)]
42. Gers, F.A.; Schmidhuber, J.; Cummins, F. Learning to forget: Continual prediction with LSTM. *Neural Comput.* **2000**, *12*, 2451–2471. [[CrossRef](#)] [[PubMed](#)]
43. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)]
44. Chen, S.; Tao, C. PolSAR image classification using polarimetric-feature-driven deep convolutional neural network. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 627–631. [[CrossRef](#)]
45. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:Learning/1412.6980. Available online: <https://arxiv.org/abs/1412.6980> (accessed on 22 December 2014).
46. 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.
47. Liu, X.; Jiao, L.; Tang, X.; Sun, Q.; Zhang, D. Polarimetric convolutional network for PolSAR image classification. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 3040–3054. [[CrossRef](#)]
48. Strobl, C.; Boulesteix, A.; Kneib, T.; Augustin, T.; Zeileis, A. Conditional variable importance for random forests. *BMC Bioinform.* **2008**, *9*, 307. [[CrossRef](#)] [[PubMed](#)]



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).