

Article

Deep Learning-Based Virtual Optical Image Generation and Its Application to Early Crop Mapping

No-Wook Park ^{1,*}, Min-Gyu Park ², Geun-Ho Kwak ³ and Sungwook Hong ⁴¹ Department of Geoinformatic Engineering, Inha University, Incheon 22212, Republic of Korea² Munji R&D Complex, Satrec Initiative, Daejeon 34051, Republic of Korea³ Geoinformatic Engineering Research Institute, Inha University, Incheon 22212, Republic of Korea⁴ Department of Environment, Energy, and Geoinformatics, Sejong University, Seoul 05006, Republic of Korea

* Correspondence: nwpark@inha.ac.kr

Abstract: This paper investigates the potential of cloud-free virtual optical imagery generated using synthetic-aperture radar (SAR) images and conditional generative adversarial networks (CGANs) for early crop mapping, which requires cloud-free optical imagery at the optimal date for classification. A two-stage CGAN approach, including representation and generation stages, is presented to generate virtual Sentinel-2 spectral bands using all available information from Sentinel-1 SAR and Sentinel-2 optical images. The dual-polarization-based radar vegetation index and all available multi-spectral bands of Sentinel-2 imagery are particularly considered for feature extraction in the representation stage. A crop classification experiment using Sentinel-1 and -2 images in Illinois, USA, demonstrated that the use of all available scattering and spectral features achieved the best prediction performance for all spectral bands, including visible, near-infrared, red-edge, and shortwave infrared bands, compared with the cases that only used dual-polarization backscattering coefficients and partial input spectral bands. Early crop mapping with an image time series, including the virtual Sentinel-2 image, yielded satisfactory classification accuracy comparable to the case of using an actual time-series image set, regardless of the different combinations of spectral bands. Therefore, the generation of virtual optical images using the proposed model can be effectively applied to early crop mapping when the availability of cloud-free optical images is limited.

Keywords: crop classification; virtual image; deep learning; generative adversarial networks

Citation: Park, N.-W.; Park, M.-G.; Kwak, G.-H.; Hong, S. Deep Learning-Based Virtual Optical Image Generation and Its Application to Early Crop Mapping. *Appl. Sci.* **2023**, *13*, 1766. <https://doi.org/10.3390/app13031766>

Academic Editor: Nathan J Moore

Received: 23 December 2022

Revised: 27 January 2023

Accepted: 28 January 2023

Published: 30 January 2023



Copyright: © 2023 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

Timely crop condition monitoring and crop yield forecasting are essential for grain supply and demand control, agricultural commodity price forecasting, and agricultural policy establishment [1–3]. Since any risk in crop production greatly affects the international grain market, the resulting volatility in international grain price significantly impacts the domestic market of crop-importing countries [4,5]. Therefore, it is crucial that crop-importing countries monitor crop growth conditions in major foreign crop-producing regions in the context of food security. In particular, the early identification of crop types before the end of crop seasons is a critical component in the timely forecasting of crop yield and production [6,7].

Remote sensing imagery can be an effective information source for crop classification in foreign countries because it can provide multi-temporal thematic information on the spectral and scattering responses of different crop types in inaccessible regions at various spatial scales [8–12]. Despite the great potential of crop classification using remote sensing imagery, crop type maps derived from remote sensing imagery inevitably include errors that significantly affect further crop yield prediction modeling; thus, it is critical to generate reliable crop type maps [13].

There are practical issues to be resolved regarding the accurate identification of crop types using remote sensing imagery, particularly for early crop mapping. Multi-temporal

images are often used as inputs for crop classification to fully account for the phenological status of crops [8–16]. However, an incomplete image time series is often available for early crop mapping to identify crop types at the early crop growth stage, which may achieve poor classification accuracy when compared to using a complete image time series [17]. Cloud contamination, common in optical imagery, is a major obstacle to obtaining cloud-free optical images at the dates optimal for early crop mapping. Synthetic-aperture radar (SAR) imagery, which can be acquired regardless of the weather conditions, can be used for crop mapping as an alternative to optical imagery [18–21]. However, it may be challenging to achieve high classification accuracy comparable to optical imagery when specific crop types have similar physical structures but different spectral responses. Furthermore, the difficulty in the visual interpretation of SAR imagery compared with optical imagery is an obstacle to collecting sufficient training samples for supervised classification in inaccessible regions, such as in foreign countries. The joint use of SAR and optical imagery has often been employed to improve classification performance by synergizing complementarity between SAR and optical imagery [22–26]. However, the availability of cloud-free optical imagery at the critical crop growth stage for early crop mapping is still limited in the multi-sensor fusion approach.

Regarding early-season crop mapping using remote sensing imagery, most previous studies have focused on the selection of multi-temporal image combinations [27,28], the comparison of various classification methods [29], and the collection of reliable training samples [13,30]. However, the difficulty in collecting cloud-free optical imagery at the optimal date for early crop mapping has not been fully explored. From a data availability perspective, the generation of cloud-free virtual optical images can be a feasible solution to overcome the missing problems in optical image acquisition for early crop mapping, which is the primary focus of this study. To this end, SAR imagery can be used as an important source of information for virtual optical image generation. Since both optical and SAR images have different imaging mechanisms and physical meanings, it is necessary to apply advanced methods for such a heterogeneous image translation [31].

Due to the great potential of deep learning (DL) for remote sensing image processing, many studies have attempted to generate virtual images using DL-based image-to-image translation. Convolutional neural networks (CNNs), which have been widely applied to image processing [32,33], lack a coordination mechanism appropriate for multi-sensor image fusion [31] and manually design effective loss functions [34]. Meanwhile, inspired by Goodfellow et al. [35], conditional generative adversarial networks (CGANs) [34,36] have been employed for virtual image generation in the remote sensing community because of their adversarial learning procedures suitable for image-to-image translation. The virtual reflectance of visible bands from other bands and satellites was generated using a CGAN-based prediction [37–39]. Regarding the optical image generation from SAR imagery, many studies have also conducted SAR-to-optical image translations based on CGANs, which have proven effective for virtual optical image generation [40–44]. Moreover, variants of CGAN, including cycle-consistent adversarial networks (CycleGANs) [45–47] and DualGANs [48], have also been proposed for image-to-image translation.

The performance of the SAR-to-optical image translation depends on the effective extraction of the key features that account for the quantitative relationships between optical and SAR images. Thus, it is critical to quantify the different reflectance and scattering characteristics of various land-cover types. Both scattering and spectral features that are useful to discriminate different crop types should be used as inputs for CGANs in the context of crop classification. Despite the potential of red-edge and shortwave infrared (SWIR) spectral bands for crop classification [49–51], the generation of those spectral bands has not yet been conducted using CGANs. For example, the spatial resolution of Sentinel-2 red-edge and SWIR bands is 20 m, which is different from the visible and near-infrared (NIR) bands with a spatial resolution of 10 m. Thus, conventional CGAN models cannot be applied to all spectral bands of Sentinel-2 without scale conversion. Most previous studies focused on the generation of visible and NIR (VNIR) band images and only used

backscattering coefficients as input variables [40,43,44,46–48]. Furthermore, the benefit of using an additional feature derived from dual-polarization SAR data (e.g., Sentinel-1 images), such as the radar vegetation index [52,53], has not been thoroughly investigated.

From a practical perspective, the virtual optical image should be used as the input for specific tasks such as land-cover classification or change detection. However, most studies have focused solely on virtual optical image generation with improved accuracy. They have not analyzed the impacts of the virtual image on specific applications when optical imagery is unavailable at the desired time. Very few studies have been conducted where the generated virtual optical image is directly applied to classification tasks, except for Bermudez et al. [42], where crop classification and wildfire detection were conducted. Despite the great importance of early crop mapping, to the best of our knowledge, the value of virtual optical imagery generated at the optimal date has not yet been evaluated.

The objectives of this study are to generate cloud-free virtual optical images using SAR images and CGANs and explore the effects of using the generated virtual optical image for early crop mapping. More specifically, this study demonstrates (1) how well the CGAN model predicts the reflectance of spectral bands from SAR imagery using additional information and (2) how predicted virtual imagery could be helpful for early crop mapping. The main contributions of this study are summarized in the following points:

- (1) Specific feature representation and extraction procedures tailored to image translation in croplands are considered in this study. Various features extracted from Sentinel-1 SAR and Sentinel-2 multi-spectral optical images are utilized for virtual Sentinel-2 image generation.
- (2) Unlike most previous studies that focused on virtual optical image generation itself, the benefits of virtual optical imagery for a real application, such as early crop mapping, are demonstrated.

The rest of this paper is organized as follows. The study area and data sets are described in Section 2. Section 3 presents the methodological framework and experimental settings employed in this study. Experiment results of virtual image generation and early crop mapping are presented in Section 4, followed by a discussion in Section 5 and conclusions in Section 6.

2. Materials

2.1. Study Area

A crop classification experiment for both virtual Sentinel-2 optical image generation and crop classification was conducted in the subarea in Illinois within the US Corn Belt. Illinois was selected because Korea mainly imports corn and soybean from the US [54] and Illinois is one of the core producing states of corn and soybean in the US [5,55].

In this study, the subarea in the southwest agricultural statistics district, where diverse crops are grown [56], was first selected as a pilot study area considering the diversity of land-cover types. Both the quantitative evaluations of the virtual Sentinel-2 imagery generated by the CGANs and the effect analysis of the generated virtual Sentinel-2 imagery on early crop mapping require multi-temporal cloud-free images. Hence, we finally selected the small pilot study area where cloud-free Sentinel-2 images in 2019 were available (Figure 1).

From the cropland data layer (CDL) provided by the National Agricultural Statistics Service (NASS) of the United States Department of Agriculture (USDA) [57,58], the crops in the study area consist of five types, including corn, soybean, wheat, alfalfa, and double cropping of winter wheat and soybean (Figure 1). The non-crop class includes developed areas, deciduous forests, grassland, and water. Corn and soybean, as major crops in the study area, occupy approximately 71.6% of the study area.

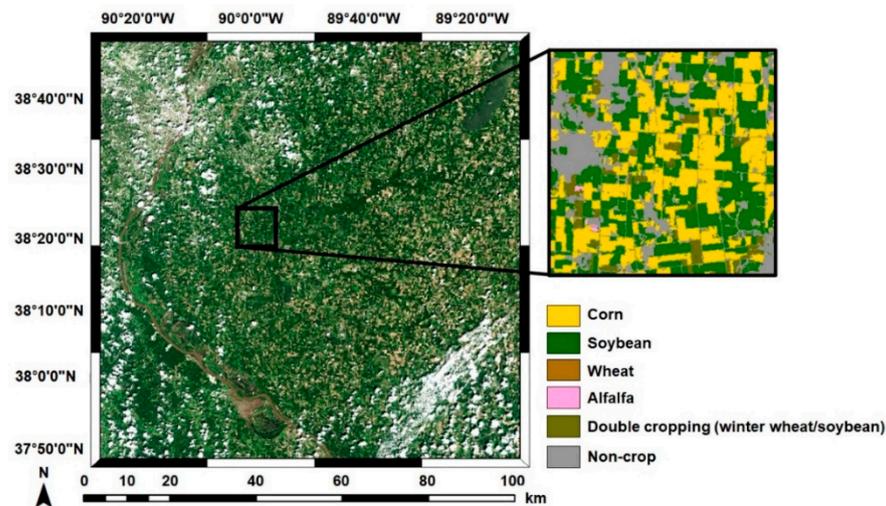


Figure 1. Location of the study area. The black box denotes the subarea where virtual image generation and crop classification are conducted. The background in the left image is the Sentinel-2 image on 30 July 2019, and the cropland data layer in 2019 in the study area is enlarged.

2.2. Data

Multi-temporal Sentinel-2 images taken in 2019 that covered the study area were considered for the crop classification task. Using a twin-satellite constellation, the Sentinel-2A and -2B images can be acquired in the same area every two or three days. The Sentinel-2 level-2A Bottom-of-Atmosphere (BOA) products covering the study area were downloaded from the Copernicus Open Access Hub [59]. The reflectance values of nine spectral bands in Table 1 were considered inputs for the experiment by considering their effectiveness for crop classification. The spatial resolutions of the nine multi-spectral bands of the Sentinel-2 images vary from 10 m to 20 m. To match the spatial resolution of all the spectral bands, three red-edge bands and two SWIR bands with a spatial resolution of 20 m were resampled to 10 m using bilinear resampling.

Table 1. Specification of Sentinel-2 multi-spectral bands (RE: red-edge; NIR: near-infrared; SWIR: shortwave infrared).

Spectral Band		Central Wavelength (nm)	Spatial Resolution (m)
B1	Blue	492.4	10
B2	Green	559.8	
B3	Red	664.6	
B4	RE1	704.1	20
B5	RE2	740.5	
B6	RE3	782.8	
B7	NIR	832.8	10
B8	SWIR2	1613.7	20
B9	SWIR3	2202.4	

The Sentinel-2 images acquired from June to October were first collected by considering the growing stages of the corn and soybean in the study area. Before the selection of the Sentinel-2 images to be used for the experiment, the normalized difference vegetation index (NDVI) values for corn and soybean parcels were first calculated using multi-temporal Sentinel-2 images and the CDL data in the study area. In cases where the Sentinel-2 images covering the study area contained cloud-contaminated areas such as clouds and shadows, the NDVI values were only computed using cloud-free subareas. The temporal variations in the average NDVI values for corn and soybean were then analyzed to identify the optimal date for early crop mapping.

Figure 2 shows the temporal profiles of the average NDVI values for corn and soybean. The corn in the study area begins to grow from the end of June. After reaching the mature stage by the end of August, harvesting begins in mid-September. On the other hand, the growing stage of soybean begins later than corn (mid-July). Soybean grows continuously until the end of August and is harvested from mid-September, similar to corn. Based on the growth cycles of corn and soybean identified in Figure 2, the optimal date for early crop mapping is either the end of July or early August. As fewer images are preferred for early crop mapping, the Sentinel-2 images acquired from June to July should be used for crop classification.

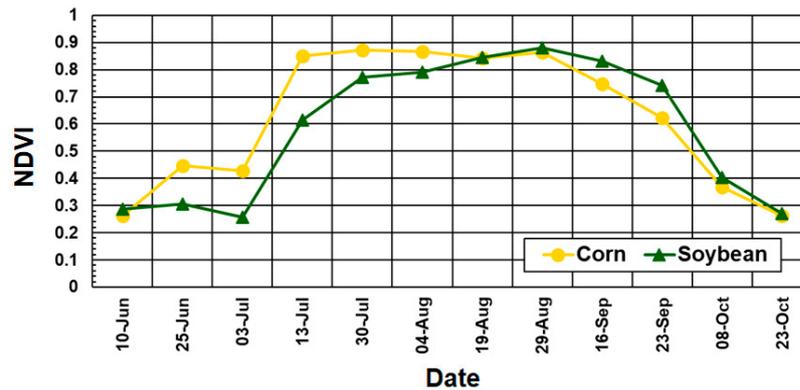


Figure 2. Temporal profiles of NDVI values for corn and soybean in 2019.

A total of 24 Sentinel-2 images were acquired from June to July, corresponding to the growing seasons of corn and soybean. However, as shown in Table 2, only seven cloud-free Sentinel-2 images were available in the study area, emphasizing the necessity of virtual cloud-free optical image generation for crop classification. Considering the temporal interval between the cloud-free Sentinel-2 images, four images, including two June images and two July images, were considered for crop classification (Figure 3). The 30 July Sentinel-2 image was selected as a target image for the virtual image generation because this image was acquired at the optimal date for early crop mapping. That is, the 30 July Sentinel-2 image was assumed to be unavailable or heavily contaminated by cloud. This image was considered the unavailable actual image to evaluate the predictive performance of virtual Sentinel-2 image generation using CGANs. The multi-temporal Sentinel-2 images and the virtual Sentinel-2 image were also used as inputs for crop classification.

Table 2. List of the availability of cloud-free Sentinel-2 images in the study area from June to July 2019 (×: cloud-free image unavailable; ○: cloud-free image available; ●: cloud-free image used for the experiment; *: target image for virtual image generation).

Date	Availability	Date	Availability
3 June	×	3 July	×
5 June	○	5 July	×
8 June	×	8 July	×
10 June	●	10 July	×
13 June	×	13 July	●
15 June	×	15 July	×
18 June	×	18 July	×
20 June	×	20 July	×
23 June	×	23 July	○
25 June	●	25 July	×
28 June	○	28 July	×
30 June	×	30 July	●*



Figure 3. Sentinel-2 true color composite images in the study area.

C-band (5.4 GHz) Sentinel-1 SAR images with dual-polarization modes (VV and VH) were also collected for the virtual Sentinel-2 image generation. The Sentinel-1A image in an ascending mode on 29 July was finally selected to be input into the CGAN model in consideration of the target date established for virtual image generation (i.e., 30 July). The level-1 Ground Range Detected (GRD) high-resolution product with a pixel size of 10 m in the interferometric wide (IW) swath mode was downloaded from the Copernicus Open Access Hub [59]. Several preprocessing procedures, including the update of the orbit state vectors, thermal noise removal, border noise removal, radiometric calibration, speckle filtering, range Doppler terrain correction, and projection, were employed using the Sentinel Application Platform (SNAP) software [60]. After preprocessing, the calibrated backscattering coefficients in the dB unit for both VV and VH polarizations were finally prepared. As the geometric errors between the Sentinel-1 and -2 images affect the predictive performance of virtual image generation, the geometric correction of the Sentinel-1 image was implemented using ground control points manually extracted from the Sentinel-2 image.

The CDL in 2019 in Figure 1 was used as the ground truth data for the extraction of crop parcels and crop classification. The original 30 m CDL data were resampled to 10 m using nearest neighbor resampling to match the spatial resolution of the Sentinel images. All of the land-cover types in the study area were reclassified into six classes: corn, soybean, wheat, alfalfa, double cropping of winter wheat and soybean, and non-crop. As crop classification is the main objective, water, built-up, and forest classes were merged as the non-crop class. The CDL data were used to extract training and test samples for crop classification.

3. Methods

The entire data processing procedures applied in this study consist of two analysis steps, including virtual Sentinel-2 image generation using CGAN and crop classification using random forest (RF). The virtual image generation step includes (1) the extraction of paired training patches and (2) feature extraction and prediction using the CGAN (Figure 4). The effectiveness of the virtual Sentinel-2 image generation is evaluated through crop classification.

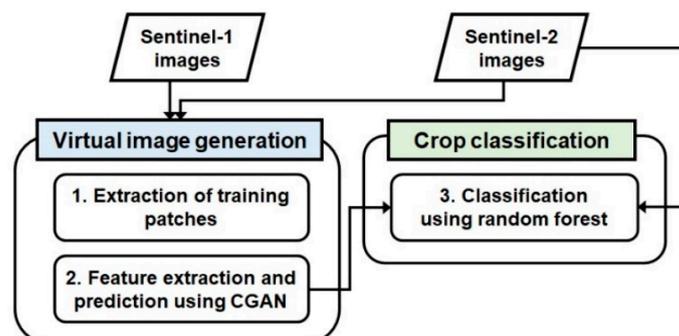


Figure 4. Schematic diagram of the processing flow applied in this study.

3.1. Extraction of Paired Training Patches

The procedure for extracting the SAR-optical training patches is illustrated in Figure 5. Prior to patch extraction, all of the images with different data value ranges were normalized to have a value between 0 and 1 to be used as inputs for the CGAN. In the first processing step, the training patches were extracted from the paired Sentinel-1 (29 July) and Sentinel-2 (30 July) images to train the generation model. As the Sentinel-2 image on 30 July was assumed to be unavailable in the pilot study area, the training patches were extracted from the cloud-free regions outside of the pilot study area within the full scenes of Sentinel-1 and -2 images. After the clouds and shadows were masked out by visual inspection, the Sentinel-1 and -2 training patches were manually extracted from the cloud-free regions. The size of the training patches was experimentally set to 512 by 512 pixels.

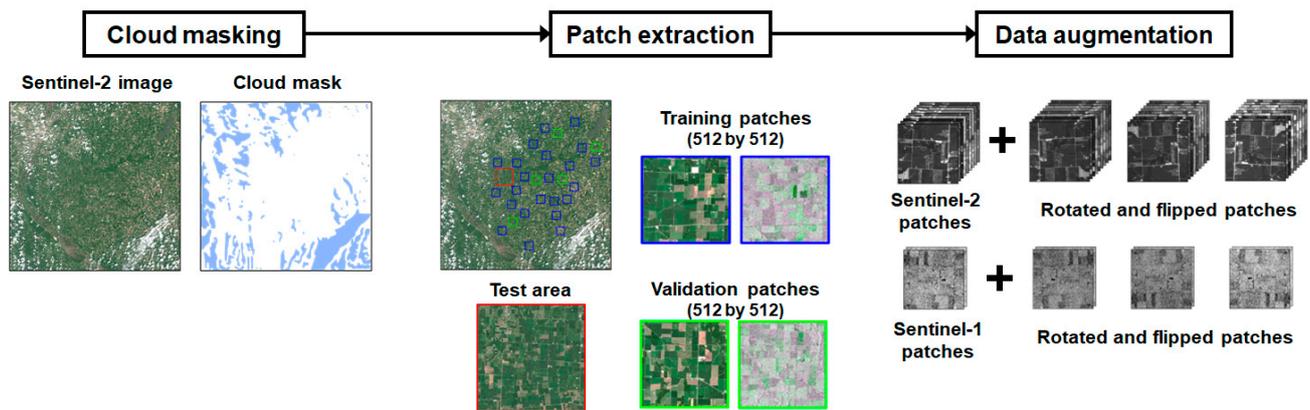


Figure 5. Workflow for the preparation of paired training patches employed in this study. Not all training and validation patches are shown in this figure.

As the generated virtual Sentinel-2 image will be used for crop classification, the training patches should include crop areas and be extracted from the regions with similar spatial characteristics to the pilot study area. To this end, regions containing more than 70% of crops within the patch window were extracted using the CDL and then selected as training patches. To avoid over-fitting and find the optimal hyperparameters of the generation model, the extracted 1120 training patches were divided into two disjoint groups, including training (1000) and validation (120) patches for the model training and the selection of optimal parameters of the generation model, respectively. Furthermore, data augmentation (DA), which artificially enlarges the number of training data [42,61], was applied to improve the generalization ability of the trained model by increasing the diversity of the training patches. Four DA schemes, including rotations and flips in the vertical and horizontal directions, were applied to expand the number of training patches intentionally. Finally, 4000 pieces of training data were considered for the model training.

3.2. Virtual Image Generation

3.2.1. Conditional Generative Adversarial Network (CGAN)

As a base model of a CGAN, a generative adversarial network (GAN) is a CNN-based architecture that consists of a generator and a discriminator [35], where the former generates virtual images after learning from the input images, and the latter distinguishes the virtual images from real images. The virtual output images can be generated through such an adversarial learning procedure.

A CGAN was proposed to overcome the limitations of the GAN, including the generation of unconditional or uncontrolled random images [34,36]. As implied from its name, the CGAN is an extended model of GANs, where both the generator and discriminator are conditioned on additional data [62]. More specifically, the generator combines the prior input noise (\mathbf{z}) and an additional condition variable (\mathbf{y}) in joint hidden representation, and

input data (x) and y are utilized as inputs for the discriminator [36]. The objective function of CGAN is defined as

$$L_{CGAN}(G, D) = E_{x, y}[\log D(x, y)] + E_{z, y}[\log(1 - D(G(z, y), y))], \tag{1}$$

where G and D are the generator and the discriminator, respectively. $E[\cdot]$ is an expectation operator. G tries to minimize the above objective function, whereas D tries to maximize it through the two-player min-max game [34].

The L1 loss function for training the generator is also introduced to alleviate blurring as

$$L_{L1} = E_{x, y, z}[\|x - G(z, y)\|_1]. \tag{2}$$

Finally, both the objective function in Equation (1) and the L1 loss function in Equation (2) are simultaneously optimized [29] as

$$\min_G \max_D L_{CGAN}(G, D) + \lambda L_{L1}, \tag{3}$$

where λ is the hyperparameter that controls the weights of the two functions.

3.2.2. Two-Stage CGAN Model

In this study, a two-stage CGAN model (hereafter referred to as CGAN_{2S}), including representation and generation stages, is presented to generate the virtual Sentinel-2 image (Figure 6).

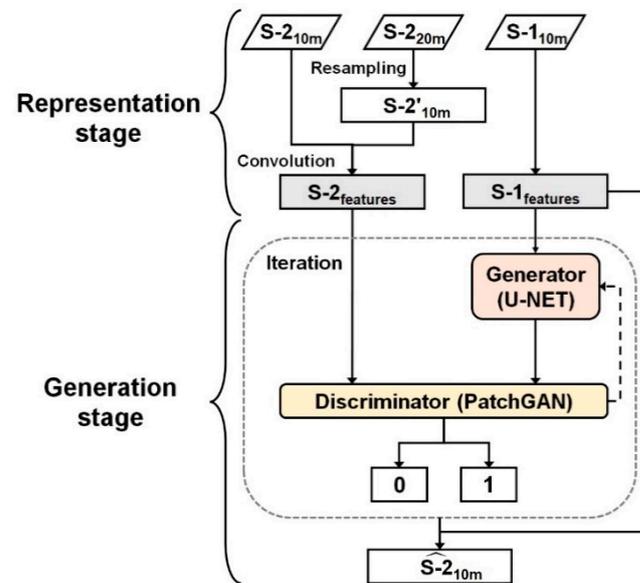


Figure 6. Two-stage CGAN model applied in this study (S-2—Sentinel-2 optical image; S-1—Sentinel-1 SAR image).

A representation stage for feature extraction is particularly designed to utilize all available information from Sentinel-1 and -2 images (Figure 7). The features extracted in the representation stage are fed into a CGAN-based generation stage to create the virtual Sentinel-2 image.

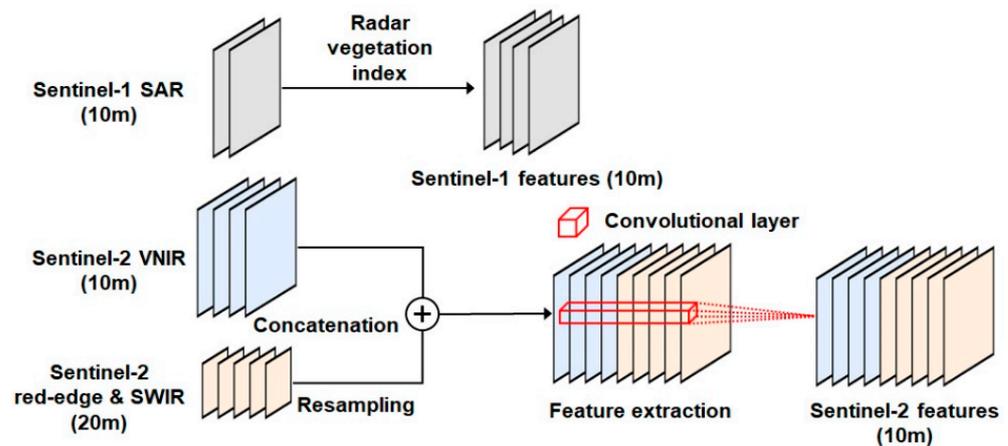


Figure 7. Feature extraction procedures from Sentinel-1 and -2 images in the representation stage.

In the representation stage, various inputs, which are helpful in representing scattering characteristics in croplands, were first considered from the Sentinel-1 images. The dual-polarization-based radar vegetation index (RVI) was considered an additional SAR-based feature. The RVI was considered an indicator of the vegetation growth dynamics [52,63,64] and is defined as

$$RVI = 4\sigma_{VH}^0 / (\sigma_{VV}^0 + \sigma_{VH}^0), \tag{4}$$

where σ_{VH}^0 and σ_{VV}^0 are the backscattering coefficients for VH and VV polarizations, respectively.

Figure 8 shows the dual-polarization backscattering coefficients and RVI of the Sentinel-1 imagery on 29 July 2019. Three inputs, including two backscattering coefficients for VH and VV polarizations, and RVI, were utilized as SAR-based features.

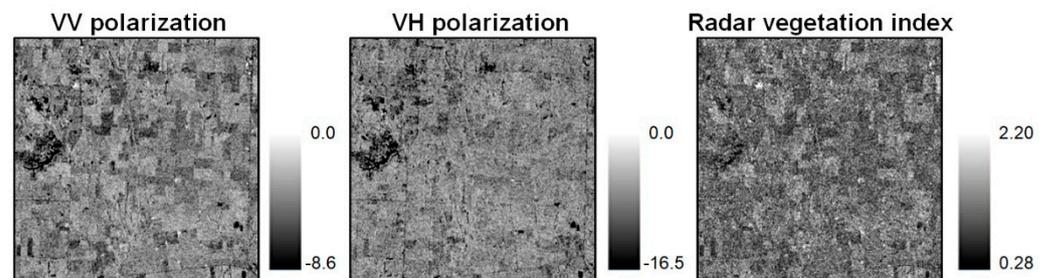


Figure 8. Sentinel-1 SAR dual-polarization images and radar vegetation index on 29 July 2019.

In addition to the above SAR-based features, multi-spectral bands in Table 1 were also used as inputs for feature extraction. The core part of Sentinel-2-based feature extraction is to extract optical features from all of the spectral information via a convolution operation. After concatenating five resampled red-edge and SWIR band images with VNIR band images, the convolution operation was applied to generate Sentinel-2-based features (see Figure 7). Consequently, the extracted features provide useful information to the discriminator of the CGAN model compared to that of each original spectral band.

The Pix2Pix architecture [34] shown in Figure 9 was adopted in this study because it can adequately reproduce spectral information [37–39,47]. The Pix2Pix model requires a paired set of SAR-based and optical-based features. The convolution layer in the representation stage is regarded as the input layer of the discriminator in the Pix2Pix model.

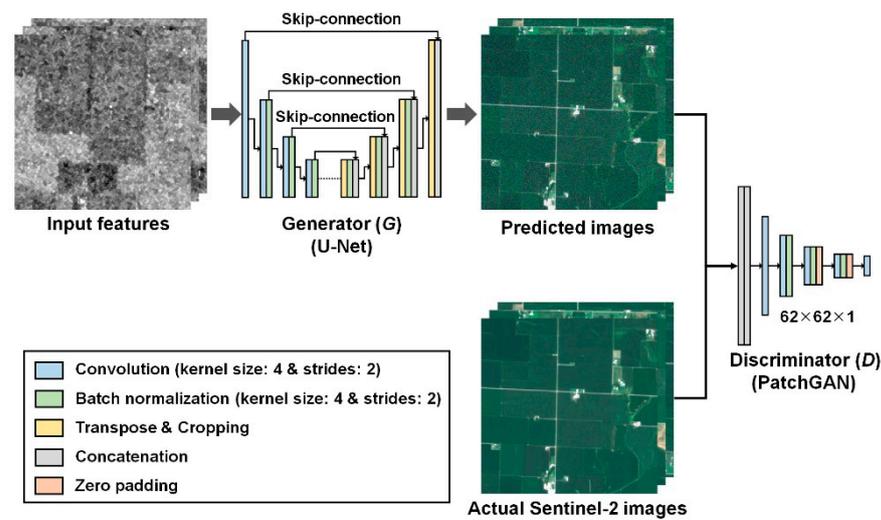


Figure 9. Pix2Pix architecture employed in the generation stage.

A U-Net architecture with a specific encoder-decoder scheme [65] was utilized as the generator in this study (Figure 9). Virtual images are predicted from the generator through down-sampling for feature extraction and up-sampling for the reconstruction from the extracted features. Skip-connection was applied to mitigate the loss of detailed information during feature extraction. As shown in Figure 9, the down-sampling and up-sampling include several processing layers, such as convolution, batch normalization, and activation functions. Thus, many hyperparameters must be optimized. In this study, the optimal hyperparameters were determined through a trial-and-error approach. The structure and optimal hyperparameters are listed in Table 3. For feature enhancement and dimension reduction, we extracted the features by reducing their dimensions with the predefined kernel size of the patch and strides (4 and 2, respectively). Moreover, dropout was applied during the up-sampling instead of the input noise in Equation (1) [34].

Table 3. List of the structures and hyperparameters of the U-Net-based generator (P: patch size; F: number of features; Conv: convolution; BN: batch normalization; L: leaky ReLU; Deconv: transposed convolution and cropping; D: dropout; R: ReLU).

Down-Sampling		Up-Sampling	
Layers	Output Dimension (P,P,F)	Layers	Output Dimension (P,P,F)
Conv, L	(256,256,64)	DeConv, BN, D, R	(2,2,1024)
Conv, BN, L	(128,128,128)	DeConv, BN, D, R	(4,4,1024)
Conv, BN, L	(64,64,256)	DeConv, BN, D, R	(8,8,1024)
Conv, BN, L	(32,32,512)	DeConv, BN, D, R	(16,16,1024)
Conv, BN, L	(16,16,512)	DeConv, BN, D, R	(32,32,1024)
Conv, BN, L	(8,8,512)	DeConv, BN, D, R	(64,64,512)
Conv, BN, L	(4,4,512)	DeConv, BN, D, R	(128,128,256)
Conv, BN, L	(2,2,512)	DeConv, BN, D, R	(256,256,128)
Conv, R	(1,1,512)	DeConv, R	(512,512,9)

The discriminator is a CNN-based model that discriminates the validity of the generated images from the generator. A PatchGAN, which tries to determine whether the generated images are real or fake in a predefined patch unit [34], not in an entire image, was adopted as the discriminator in this study. As the neighboring pixels in the remote sensing images are more similar than distant pixels, PatchGAN was applied to determine whether they are real or fake with an appropriate patch size that maintains the autocorrelation between the pixels.

The discriminator includes convolution, batch normalization, and activation function layers to extract local features from the input image. The Sentinel-2-based features extracted through the convolution operation were utilized as inputs for the discriminator. Both the kernel size and strides of the convolution layer were set to 1 to maintain the size of the input patch. A sigmoid activation function was also applied to distinguish between real (1) and fake (0) (Table 4). Similar to the generator, the features were extracted by reducing the dimension with the kernel size and strides set to 4 and 2, respectively.

Table 4. List of the structures and hyperparameters of the PatchGAN-based discriminator (P: patch size; F: number of features; Conv: convolution; BN: batch normalization; L: leaky; S: sigmoid; Z: zero padding; *: convolution operation in the representation stage).

Layers	Output Dimension (P,P,F)
* Conv, L	(512,512,9)
Conv, L	(256,256,32)
Conv, BN, L	(128,128,64)
Conv, BN, L, Z	(66,66,128)
Conv, BN, L, Z	(63,63,256)
Conv, S	(62,62,1)

The virtual Sentinel-2 image was finally created using the Sentinel-1 image acquired on the prediction date as the input for the trained CGAN model.

3.3. Crop Classification Using Random Forest

In this study, the validity of the virtual image generation was assessed in relation to crop classification using the generated Sentinel-2 image as one of the inputs. Of the various machine learning-based classifiers, RF was employed as a classifier because of its superior classification performance in remote sensing image classification [66–68]. It should be noted that the main objective of crop classification is to analyze the impact of the virtual image on the classification performance, not the selection of the best classifier. As an ensemble classifier, RF maximizes the diversity of decision trees through the bootstrapping of training samples and the random subsampling of input features to improve classification performance [69]. The optimal values of the two hyperparameters, including the number of trees to be grown in the forest and the number of variables for the node partitioning, were determined using a grid search.

The training samples for the RF-based supervised classification were extracted from the CDL in 2019. In early crop mapping in foreign countries, it is often difficult to collect sufficient training samples through visual inspection. To mimic this early crop mapping case, a very small number of pixels were extracted as training samples. From the preliminary test, 1056 training samples (approximately 0.1% of the total pixels in the study area) were finally utilized as training samples for RF-based classification. The number of training samples for each class was determined by considering the relative proportions of classes in the study area (Table 5). As a result, the number of training samples for minor crops in the study area, including wheat and alfalfa, is very small. Furthermore, to mitigate a biased sampling problem of major classes in supervised classification [13], we set the number of training samples for the double cropping class equal to corn and soybean. All of the pixels except the training samples were considered test samples for the accuracy evaluation. As very small training samples were extracted, the relative proportion of training samples to test samples was still approximately 0.1%, which is suitable for thoroughly evaluating the classification results for early crop mapping.

Table 5. Numbers of training and test samples for five crop types. The percentage value in parentheses of the training samples indicates the relative proportion for each class.

Class	Training Samples	Test Samples
Corn	314 (29.7%)	352,473
Soybean	314 (29.7%)	397,161
Wheat	41 (3.9%)	274
Double cropping	314 (29.7%)	51,511
Alfalfa	73 (7.0%)	3407
Total	1056	1,047,520

3.4. Experimental Setup and Evaluation

3.4.1. Virtual Image Generation

The prediction performance of the CGAN-based virtual image generation was evaluated in relation to the different input combinations. Since the virtual Sentinel-2 imagery is predicted using the SAR-based features as inputs after training the CGAN model, the effect of using different SAR-based features was investigated by comparing the case using RVI with the case of only using the dual-polarization backscattering coefficients. Moreover, due to the fact that the CGAN_{2S} presented in this study utilizes the features extracted from all spectral bands of the Sentinel-2 imagery (i.e., VNIR, red-edge, and SWIR bands), two CGAN models using different input spectral bands were additionally considered for comparison purposes. The first CGAN model only utilized four 10 m VNIR bands as inputs for generating virtual VNIR bands (hereafter referred to as CGAN_{VNIR}). As for the second CGAN model (hereafter referred to as CGAN_{RESW}), five resampled red-edge/SWIR band images were utilized to generate virtual red-edge and SWIR bands. Thus, the prediction performance of CGAN_{2S} was compared with CGAN_{VNIR} and CGAN_{RESW} to generate virtual VNIR and red-edge/SWIR bands, respectively (Table 6).

Table 6. Comparison cases for virtual Sentinel-2 image generation.

Cases	Input Features
CGAN _{2S}	Sentinel-1 SAR features and all Sentinel-2 spectral bands
CGAN _{VNIR}	Sentinel-1 SAR features and Sentinel-2 VNIR bands
CGAN _{RESW}	Sentinel-1 SAR features and Sentinel-2 red-edge/SWIR bands

The prediction accuracy of the generated virtual images was quantitatively evaluated using an actual Sentinel-2 image on 30 July. Two accuracy statistics, including the root-mean-square error (RMSE) and the correlation coefficient (CC), were used as the quantitative measures of prediction accuracy. The structural similarity (SSIM) and the peak signal-to-noise ratio (PSNR) [70] were also used to measure the spatial similarity between the actual and the virtual images. The larger CC, PSNR, and SSIM values and the smaller RMSE value indicate that the virtual image is more similar to the actual image.

3.4.2. Crop Classification

An incremental classification approach, frequently applied to classification using multi-temporal images [17,71], was employed for crop classification since early crop mapping was the primary purpose of crop classification. The five incremental classification cases were considered in this study, as shown in Table 7. By starting the classification using only the 10 July Sentinel-2 image (case 1), the new Sentinel-2 image was added as an input along with all of the previously acquired Sentinel-2 images. As the optimal date for early crop mapping was at the end of July (see Figure 2), the 30 July Sentinel-2 image was considered the last image to be used for the incremental classification (case 4). The effect of using the virtual Sentinel-2 image on early crop mapping (case 5) was evaluated through comparisons with cases 3 and 4. In particular, the quality of the virtual Sentinel-2 image generation can be indirectly assessed by comparing case 5 with case 4.

Table 7. Incremental classification cases considered for early crop mapping.

Case	Input Images
1	10 June
2	10 June + 25 June
3	10 June + 25 June + 13 July
4	10 June + 25 June + 13 July + 30 July (actual Sentinel-2 image)
5	10 June + 25 June + 13 July + 30 July (virtual Sentinel-2 image)

The crop classification accuracy was quantitatively evaluated using the reference samples shown in Table 5. After preparing a confusion matrix, the overall accuracy (OA) was calculated and used as a quantitative accuracy measure for five incremental classification cases.

3.5. Implementation

Virtual image generation was implemented using the Pix2Pix-based code [34], TensorFlow [72], Keras [73], and NumPy [74] libraries on Python 3.6.7. The Scikit-learn library [75] was utilized for RF-based classification. In addition, ENVI software version 5.6 (L3Harris Technologies, Broomfield, CO, USA) and Python programming were used for data preprocessing, visualization, and quantitative accuracy evaluation. All the procedures were run on the CentOS 7.0 operating system with an Intel XEON E5–2630 v4 @ 2.2 GHz CPU and two NVIDIA GTX1080ti GPUs with 11 GB of memory.

4. Results

4.1. Virtual Image Generation Results

4.1.1. Prediction Performance for SAR-Based Feature Combinations

When comparing the prediction performance of the different SAR-based features of CGAN_{2S}, the inclusion of the RVI to the dual-polarization backscattering coefficients exhibited the best prediction accuracy for all of the spectral bands (Table 8). The maximum improvement in the RMSE of the case using all of the SAR features over the case without RVI was 8.97% for the green band. When comparing the relative improvement in the different spectral bands, the relative improvement in RMSE was substantial for the VNIR bands compared to the red-edge and SWIR bands. The structural similarity and PSNR were also improved by adding the RVI as additional feature information. The RVI improved the SSIM by 3.2%p and 2.4%p for the NIR and red-edge3 bands, respectively.

Table 8. Accuracy statistics for different SAR feature combinations (RE: red-edge; NIR: near infrared; SWIR: shortwave infrared; RMSE: root mean square error; CC: correlation coefficient; SSIM: structural similarity; PSNR: peak signal-to-noise ratio).

Band	CGAN _{2S} without RVI				CGAN _{2S} with RVI			
	RMSE	CC	SSIM	PSNR	RMSE	CC	SSIM	PSNR
Blue	0.021	0.557	0.934	33.410	0.020	0.614	0.945	34.192
Green	0.022	0.615	0.932	33.021	0.020	0.673	0.944	33.848
Red	0.028	0.661	0.894	31.190	0.025	0.708	0.913	31.987
NIR	0.070	0.766	0.688	23.152	0.064	0.799	0.720	23.846
RE1	0.024	0.778	0.920	32.583	0.020	0.815	0.931	33.197
RE2	0.045	0.789	0.811	26.947	0.043	0.820	0.828	27.435
RE3	0.061	0.798	0.732	24.332	0.057	0.833	0.756	24.816
SWIR2	0.029	0.844	0.906	30.862	0.027	0.864	0.921	31.376
SWIR3	0.024	0.811	0.921	32.350	0.023	0.832	0.932	32.856

Figure 10 shows the prediction results with the actual Sentinel-2 imagery zoomed in on the subarea for a visual comparison. The local details in the actual imagery were not perfectly depicted in the two prediction results. However, the prediction using RVI

as additional feature information could reproduce the overall patterns better than the prediction using only backscattering coefficients. In particular, the parcel boundaries were well depicted in the prediction using RVI, as confirmed by the higher SSIM value in Table 8. These comparison results confirm the potential of using the RVI for the extraction of valuable information for vegetation discrimination.

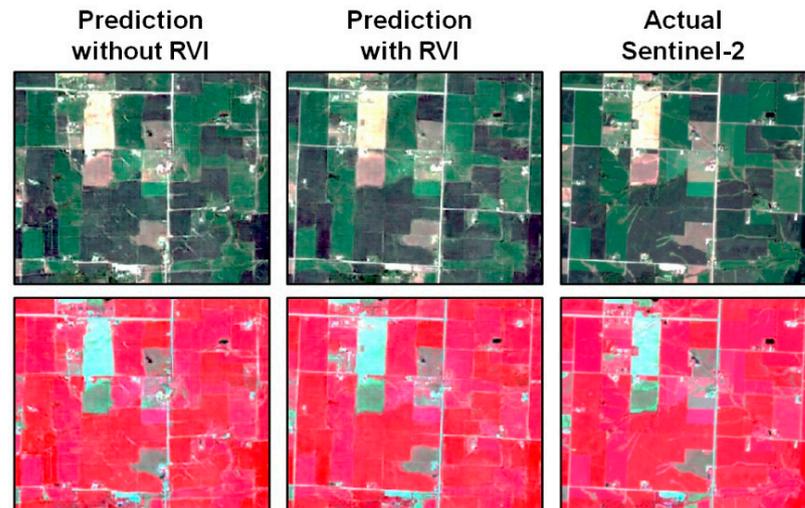


Figure 10. Comparison of prediction results for different SAR-based features with the actual Sentinel-2 imagery. The first and second rows represent the true and false color composite (NIR-red-green as RGB) images, respectively.

4.1.2. Prediction Performance for Sentinel-2-Based Feature Combinations

Table 9 summarizes the comparison results of the three CGAN models with different Sentinel-2 band combinations, as listed in Table 6. Based on the comparison results of the SAR-based feature combinations, three SAR-based features, including two backscattering coefficients and RVI, were used for the three CGAN models.

Table 9. Accuracy statistics of three CGAN models with different Sentinel-2 feature combinations (RE: red-edge; NIR: near infrared; SWIR: shortwave infrared; RMSE: root mean square error; CC: correlation coefficient; SSIM: structural similarity; PSNR: peak signal-to-noise ratio).

Band	CGAN _{VNIR}				CGAN _{RESW}				CGAN _{2S}			
	RMSE	CC	SSIM	PSNR	RMSE	CC	SSIM	PSNR	RMSE	CC	SSIM	PSNR
Blue	0.032	0.600	0.853	30.027					0.020	0.614	0.945	34.192
Green	0.036	0.649	0.865	28.874					0.020	0.673	0.944	33.848
Red	0.050	0.691	0.745	26.087					0.025	0.708	0.913	31.987
NIR	0.141	0.771	0.569	17.040					0.064	0.799	0.720	23.846
RE1					0.027	0.687	0.904	31.379	0.022	0.815	0.931	33.197
RE2					0.055	0.662	0.767	25.167	0.043	0.820	0.828	27.435
RE3					0.076	0.675	0.664	22.351	0.057	0.833	0.756	24.816
SWIR2					0.036	0.751	0.882	28.782	0.027	0.864	0.921	31.376
SWIR3					0.028	0.730	0.900	30.928	0.023	0.832	0.932	32.856

When comparing the conventional CGAN models that utilized inputs with a single spatial resolution (CGAN_{VNIR} and CGAN_{RESW}), the prediction using CGAN_{2S} showed superior predictive performance across all the spectral bands. In particular, the effect of a multi-band combination was much more significant for the prediction of the VNIR bands. The use of red-edge and SWIR bands as additional features could improve the RMSE from 37.5% to 54.61%. The maximum improvement in RMSE was achieved in relation to the NIR band. Although the resampled red-edge and SWIR bands were utilized as inputs,

the features from the highly correlated red-edge2 and red-edge3 bands could improve the prediction accuracy of the NIR band. The SSIM values for the red and NIR bands were also improved by 16.8%p and 15.1%p, respectively. The accuracy improvement was relatively minor when comparing the predictive performance of CGAN_{2S} with CGAN_{RESW}. However, the accuracy values of the red-edge and SWIR bands predicted by CGAN_{2S} were still much greater than that of CGAN_{RESW}. The use of the VNIR bands with finer spatial resolutions could improve the predictive accuracy of the red-edge and SWIR bands. These comparison results demonstrate the necessity of a multi-band combination to predict optical imagery.

The visual comparison results in two subareas are shown in Figure 11. The NIR and red-edge3 bands were selected for the visual comparison because their sensitivity to vegetation vitality is helpful for crop classification. For the NIR band prediction, CGAN_{VNIR} and CGAN_{2S} generated similar results in this subarea, although their RMSE values over the study area were significantly different. In contrast, CGAN_{RESWIR} generated much-degraded prediction results for the red-edge3 band. In particular, spectral distortion in the prediction results was severe in the left subarea consisting of lakes, developed areas, and grassland. CGAN_{RESWIR} utilized resampled input images to consider the spatial resolution difference between Sentinel-1 and -2 images. Hence, it has a limitation to properly learn the complex spatial patterns from the pair images. Furthermore, the Sentinel-1 SAR imagery with a spatial resolution of 10 m and speckle noise filtering may be responsible for the failure to depict the spatial patterns of small objects. On the other hand, CGAN_{2S} could overcome such limitations by considering the fine-scale VNIR bands as additional features, which yielded much better prediction results for both subareas with different land-cover types.

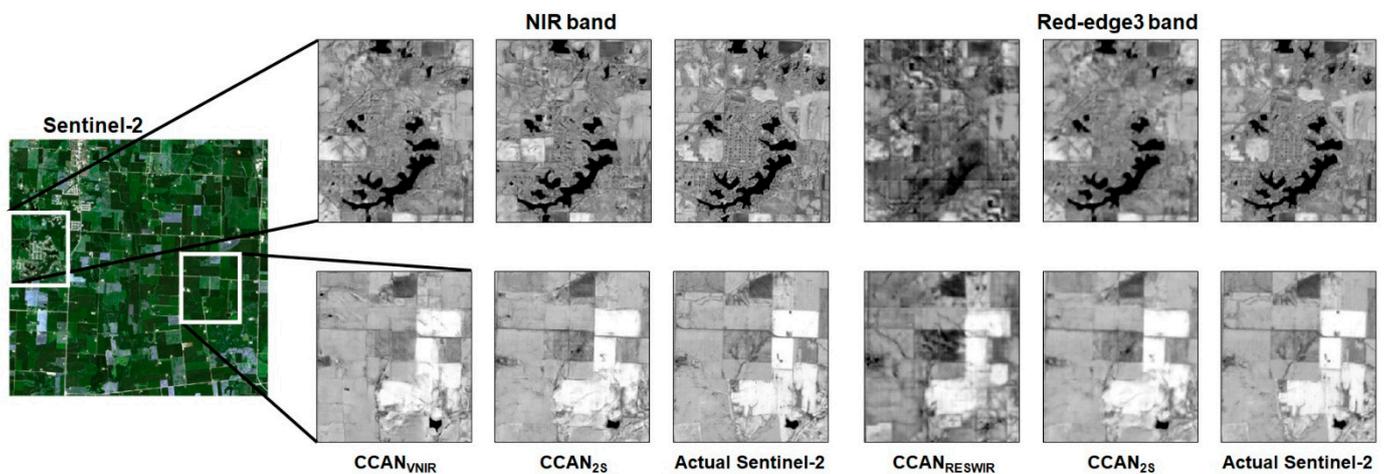


Figure 11. Prediction results of NIR and red-edge3 bands for different CGAN models with the actual Sentinel-2 imagery. The two subareas marked with white boxes in the Sentinel-2 imagery are enlarged.

As the actual Sentinel-2 imagery is available, the errors per pixel for the NIR and red-edge3 bands were calculated and further analyzed using the CDL (Figure 12). For both spectral bands, the magnitude of errors of CGAN_{2S} was relatively smaller than those of CGAN_{VNIR} or CGAN_{RESWIR}. As expected, the superiority of CGAN_{2S} over CGAN_{RESWIR} was substantial in relation to the red-edge3 band. The overestimation (red color) and underestimation (purple color) of spectral reflectance were significantly reduced for CGAN_{2S}. However, despite the significant reduction in errors when using CGAN_{2S}, there still exist relatively significant errors in the western, northwestern, and northeastern parts of the study area. The land-cover types in the western part include developed areas. Thus, the land-cover distributions in this area are relatively more complex than the other crop parcel areas, which resulted in scattered error distributions in the NIR band prediction. The clustered errors in the northwestern and northeastern parts indicate the parcel-unit errors,

mainly due to the learning process in a patch unit for all of the CGAN models. The different spatial distributions and data ranges of the SAR-based features within each patch unit may be responsible for the failure to restore the actual spectral reflectance values. These effects were relatively severe for the urban areas, including complex land-cover types and objects of different sizes. Since more than 70% of the pixels in the training patch are composed of crops, large errors may occur in the developed areas and roads. The errors at road and parcel boundaries were also more significant than those in the interior of the crop parcels, mainly due to the use of the features generated by a convolution operation in the generator. Despite these intrinsic errors, CGAN_{2S} showed the best prediction performance.

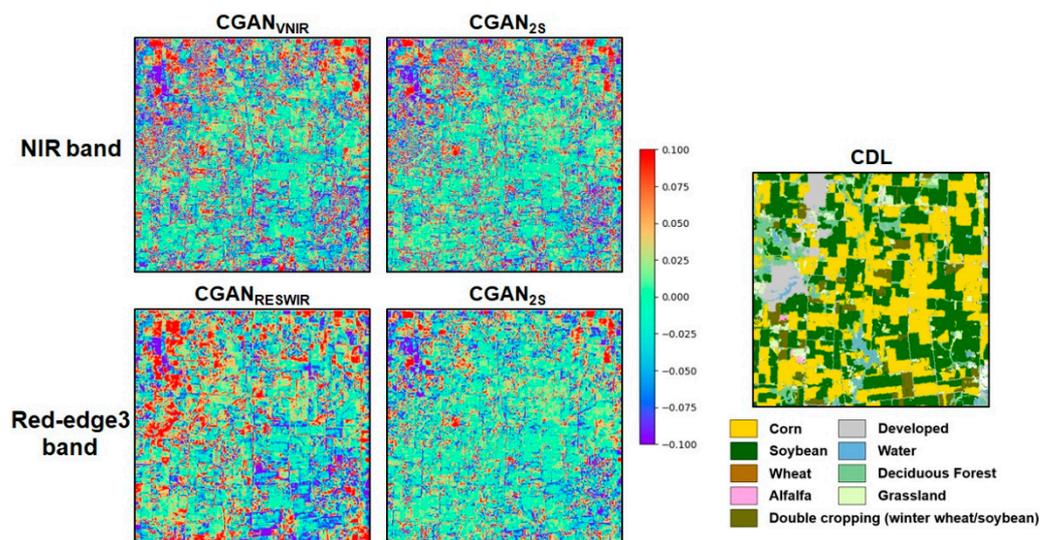


Figure 12. Error distributions over the study area of NIR and red-edge3 bands for different CGAN models with the CDL.

4.2. Crop Classification Results

The virtual spectral bands predicted using CGAN_{2S} were used as inputs for the incremental classification (case 5 in Table 7). The classification accuracy values of the incremental classification results for different Sentinel-2 input spectral bands are shown in Figure 13. As more input Sentinel-2 images were used for classification, the classification accuracy increased regardless of the CGAN model type. When comparing the effects of using different input spectral bands, the poorest OA was achieved for the case using VNIR bands for all cases. Except for cases 1 and 2, the difference in OA between using red-edge and SWIR bands and all of the spectral bands was insignificant for cases 3 to 5.

Cases 1 to 2, where the Sentinel-2 images acquired before July were used as inputs for classification, achieved poor classification accuracy values of less than 80%. The OA of case 3 using the Sentinel-2 image on 13 July was larger than that of case 2; however, it was still lower than cases 4 and 5. The inclusion of the Sentinel-2 imagery on 30 July for cases 4 and 5 could significantly increase the classification accuracy, achieving an OA value of more than 90%. The best OA was achieved for case 4, which used the actual Sentinel-2 imagery on 30 July. Case 5, which used the virtual Sentinel-2 imagery predicted by CGAN, yielded an OA comparable to case 4. The difference in OA between cases 4 and 5 was approximately 1.72%p on average for the three different inputs. When comparing case 3, the inclusion of all of the spectral bands of the virtual Sentinel-2 imagery on 30 July increased the OA by 5.03%p. Despite the errors in the virtual spectral bands predicted by CGAN_{2S}, the promising accuracy for case 5, which is higher than case 3 and comparable to case 4, indicates the feasibility of the virtual optical image generation for early crop mapping.

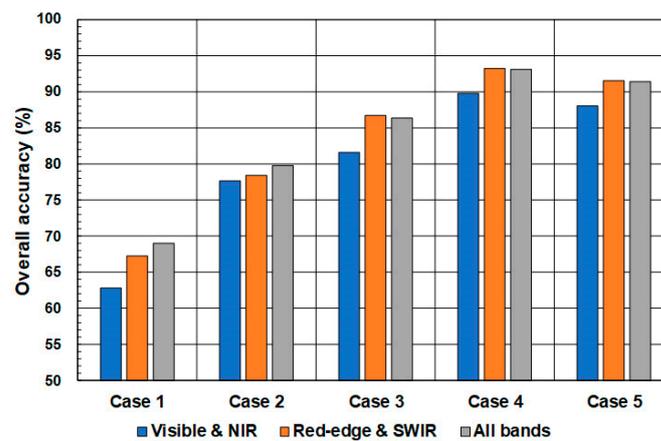


Figure 13. Overall accuracy values of five different incremental classification results for different input combinations. The case numbers can be referred to in Table 7.

To further analyze the classification result, the classification map of case 5 using all of the spectral bands was compared with the CDL, which was regarded as the ground truth, and a classification error map was generated (Figure 14). As only crop parcels were considered for classification, non-crop areas were masked in the error map. Notably, the parts with large error values in Figure 12 were correctly classified, which implies that the errors in spectral reflectance did not significantly affect the classification performance. From the error map in Figure 14, most misclassification occurred along the parcel boundaries, except for some crop parcels. The resampling of the original 30 m CDL to 10 m may not fully represent the actual land-cover type near the parcel boundary, which resulted in misclassification along the parcel boundaries. The actual crop type in the misclassified crop parcels in Figure 14 (black color) was mostly soybean and was misclassified as alfalfa. However, as alfalfa is a minor crop type in the study area, this misclassification did not significantly affect the overall accuracy. From both the quantitative and the qualitative evaluation results, it can be concluded that the cloud-free virtual imagery generated at the optimal date for early crop mapping could be a practical input for crop classification and generate reliable results.

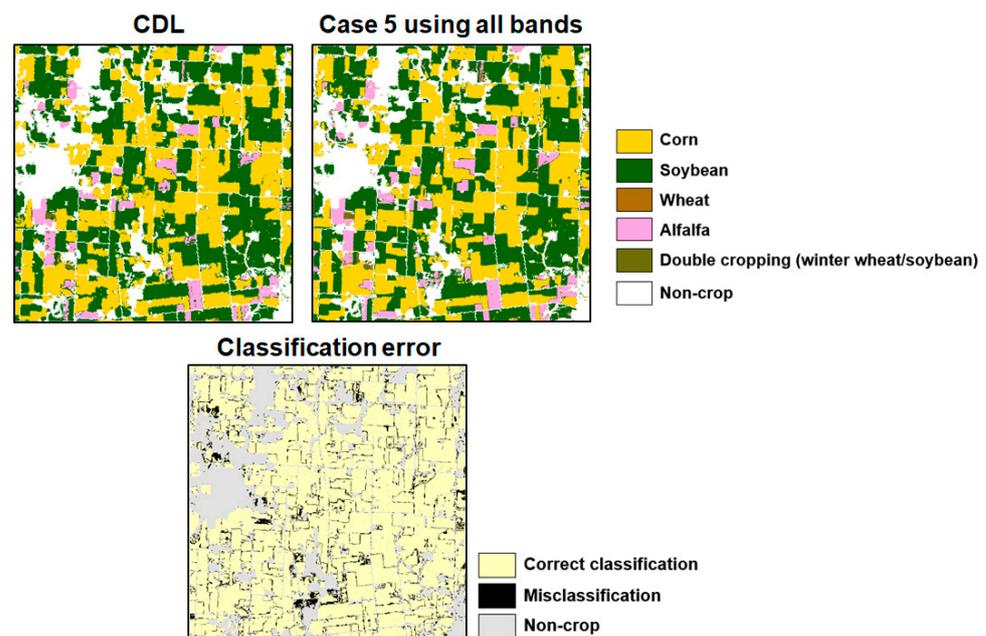


Figure 14. Comparison of the classification result of case 5 using all bands with CDL and the classification error map.

5. Discussion

Concerning virtual image generation, this study utilized all of the available spectral bands as inputs in the representation stage for Sentinel-2-based feature extraction. From the experiment results, the use of VNIR bands could improve the prediction accuracy for the prediction of the red-edge and SWIR bands, and vice versa. The feature extraction from highly correlated additional spectral bands could improve prediction accuracy. In this study, after resampling the red-edge and SWIR bands with a spatial resolution of 20 m to 10 m, the resampled bands were fed into the representation stage, which extracted features, including all the spectral information, via a convolution operation. When multi-spectral bands have different spatial resolutions (i.e., Sentinel-2), a simple resampling method, such as nearest neighbor and bilinear, may not properly account for the differences in the spatial resolution. Although the improved prediction accuracy could be obtained from CGAN_{2S}, the CGAN_{2S} prediction still included the blurred spectral reflectance of the red-edge and SWIR bands. To generate fine-scale band images that maintain the original spectral information, it is necessary to apply advanced spatial downscaling and super-resolution methods, including area-to-point regression kriging [76] and deep neural networks [77,78]. In future work, it will be worthwhile to analyze the effects of spatial resolution enhancement on the prediction of reflectance.

Previous studies on the virtual optical image generation using SAR imagery and CGANs reported that loss of texture information of surface objects and distortion in reflectance were still observed in the generated optical imagery [40,79]. A significant increase in prediction performance was achieved in the CGAN_{2S} prediction in this study. However, relatively large errors at parcel boundaries and distortion in reflectance were also observed in the CGAN_{2S} prediction, as shown in Figure 12. Higher prediction errors did not directly result in misclassification in this study. Since virtual spectral bands are directly used as inputs for crop classification, the generation of virtual imagery with fewer errors is a prerequisite for reliable early crop mapping. In this study, the comparison between CGAN_{2S} and other CGAN-based models was not performed because the main focus was on the evaluation of prediction performance with respect to the different input feature combinations, which was adopted in previous CGAN-based and application-oriented studies [42,53]. Despite the promising early crop mapping results, it is still challenging to create reliable virtual optical imagery due to the difference in imaging mechanism between SAR and optical images. From a methodological point of view, reliable virtual optical images can be generated through the use of previous time pairs of SAR and optical images as additional information [40], the modification of model structures, and loss functions [79–81]. Thus, further methodological developments and comparisons with other advanced CGAN-based models should be performed to improve the prediction performance of virtual image generation.

In this study, the subarea in Illinois state was selected as the experimental pilot study area. Early crop mapping results are usually used as inputs for crop yield forecasting and crop acreage estimation on a regional scale (e.g., state unit) [2,3,5]. Several aspects of virtual image generation should be further considered when extending the study area. When the virtual optical imagery should be generated over a large region, several images must be stitched to generate the large-scale optical imagery. Stitching several virtual images acquired along different paths inevitably includes uneven seam lines of reflectance between the scene boundaries. Hence, the generation of large-scale seamless virtual optical imagery requires seam elimination, such as weighted averaging [47]. Furthermore, different growth environments and stages of the same crop types within a large study area [56] require a sophisticated design for collecting training pairs of SAR and optical images. Such specific considerations for an extension of the study area should be reflected in future work so that the findings of this study could be more practical for early crop mapping in foreign countries.

6. Conclusions

Cloud contamination in optical imagery is a critical obstacle to constructing a complete optical image time series for crop classification, particularly for early crop mapping. In this study, the benefits of using virtual cloud-free optical imagery using CGANs were investigated in the context of early crop mapping. This study presented a two-stage CGAN approach, including feature representation and extraction procedures tailored to SAR-to-optical image translation in croplands. The quality of the generated virtual optical imagery was assessed via early crop mapping using the virtual optical imagery as inputs for crop classification. The major findings from a crop classification experiment using Sentinel-1 and -2 images are summarized as follows:

- (1) The use of various available inputs for feature extraction in CGANs, including the radar vegetation index and all of the correlated spectral bands, achieved a significant improvement in prediction performance compared to conventional CGAN models, which only used the backscattering coefficients from Sentinel-1 SAR imagery and the prediction target bands of Sentinel-2 imagery as inputs.
- (2) The use of the predicted virtual spectral bands taken at the optimal date as inputs for early crop mapping led to a classification accuracy comparable to the classification case using actual cloud-free optical imagery.

Based on the above findings, with further development, virtual optical image generation from the presented CGAN model could be beneficial for reconstructing missing information in the cloud-contaminated optical imagery and for thematic mapping tasks using an optical image time series.

Author Contributions: Conceptualization, N.-W.P. and M.-G.P.; methodology, M.-G.P., N.-W.P. and G.-H.K.; formal analysis, M.-G.P.; data curation, M.-G.P., N.-W.P., G.-H.K. and S.H.; writing—original draft preparation, N.-W.P. and M.-G.P.; writing—review and editing, G.-H.K. and S.H.; supervision, N.-W.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. NRF-2022R1F1A1069221).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The Sentinel-1 and -2 images are made publicly available by European Space Agency via the Copernicus Open Access Hub at <https://scihub.copernicus.eu> (accessed on 10 January 2022). The Cropland Data Layer data are also publicly available by the United States Department of Agriculture via National Agricultural Statistics Service at <https://nassgeodata.gmu.edu/CropScape> (accessed on 10 January 2022).

Acknowledgments: The authors thank the three anonymous reviewers for providing constructive comments that greatly improved the presentation of the manuscript.

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

References

1. Doraiswamy, P.C.; Sinclair, T.R.; Hollinger, S.; Akhmedov, B.; Stern, A.; Prueger, J. Application of MODIS derived parameters for regional crop yield assessment. *Remote Sens. Environ.* **2005**, *97*, 192–202. [[CrossRef](#)]
2. Lee, K.-D.; Na, S.-I.; Hong, S.-Y.; Park, C.-W.; So, K.-H.; Park, J.-M. Estimating corn and soybean yield using MODIS NDVI and meteorological data in Illinois and Iowa, USA. *Korean J. Remote Sens.* **2017**, *33*, 741–750, (In Korean with English Abstract).
3. Ban, H.-Y.; Kim, K.S.; Park, N.-W.; Lee, B.-W. Using MODIS data to predict regional corn yields. *Remote Sens.* **2017**, *9*, 16. [[CrossRef](#)]
4. Seo, B.; Lee, J.; Lee, K.-D.; Hong, S.; Kang, S. Improving remotely-sensed crop monitoring by NDVI-based crop phenology estimators for corn and soybeans in Iowa and Illinois, USA. *Field Crops Res.* **2019**, *238*, 113–128. [[CrossRef](#)]
5. Pazhanivelan, S.; Geethalakshmi, V.; Tamilmounika, R.; Sudarmanian, N.S.; Kaliaperumal, R.; Ramalingam, K.; Sivamurugan, A.P.; Mrunalini, K.; Yadav, M.K.; Quicho, E.D. Spatial rice yield estimation using multiple linear regression analysis, semi-physical approach and assimilating SAR satellite derived products with DSSAT crop simulation model. *Agronomy* **2022**, *12*, 2008. [[CrossRef](#)]

6. Atzberger, C. Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. *Remote Sens. Environ.* **2013**, *5*, 949–981. [[CrossRef](#)]
7. Weiss, M.; Jacob, F.; Duveiller, G. Remote sensing for agricultural applications: A meta-review. *Remote Sens. Environ.* **2020**, *236*, 111402. [[CrossRef](#)]
8. Rahman, M.S.; Di, L.; Yu, E.; Zhang, C.; Mohiuddin, H. In-season major crop-type identification for US cropland from Landsat images using crop-rotation pattern and progressive data classification. *Agriculture* **2019**, *9*, 17. [[CrossRef](#)]
9. Wardlow, B.D.; Egbert, S.L. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains. *Remote Sens. Environ.* **2008**, *112*, 1096–1116. [[CrossRef](#)]
10. Immitzer, M.; Vuolo, F.; Atzberger, C. First experience with Sentinel-2 data for crop and tree species classifications in central Europe. *Remote Sens.* **2016**, *8*, 166. [[CrossRef](#)]
11. 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.
12. Lee, D.-H.; Kim, H.-J.; Park, J.-H. UAV, a Farm Map, and machine learning technology convergence classification method of a corn cultivation area. *Agronomy* **2021**, *11*, 1554. [[CrossRef](#)]
13. Kim, Y.; Park, N.-W.; Lee, K.-D. Self-learning based land-cover classification using sequential class patterns from past land-cover maps. *Remote Sens.* **2017**, *9*, 921. [[CrossRef](#)]
14. Marais Sicre, C.; Inglada, J.; Fieuzal, R.; Baup, F.; Valero, S.; Cros, J.; Demarez, V. Early detection of summer crops using high spatial resolution optical image time series. *Remote Sens.* **2016**, *8*, 591. [[CrossRef](#)]
15. Waldhoff, G.; Lussem, U.; Bareth, G. Multi-data approach for remote sensing-based regional crop rotation mapping: A case study for the Rur catchment, Germany. *Int. J. Appl. Earth Obs.* **2017**, *61*, 55–69. [[CrossRef](#)]
16. Simón Sánchez, A.-M.; González-Piqueras, J.; de la Ossa, L.; Calera, A. Convolutional neural networks for agricultural land use classification from Sentinel-2 image time series. *Remote Sens.* **2022**, *14*, 5373. [[CrossRef](#)]
17. Kwak, G.-H.; Park, C.-w.; Lee, K.-d.; Na, S.-i.; Ahn, H.-y.; Park, N.-W. Potential of hybrid CNN-RF model for early crop mapping with limited input data. *Remote Sens.* **2021**, *13*, 1629. [[CrossRef](#)]
18. Skriver, H.; Mattia, F.; Satalino, G.; Balenzano, A.; Pauwels, V.R.; Verhoest, N.E.; Davidson, M. Crop classification using short-revisit multitemporal SAR data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2011**, *4*, 423–431. [[CrossRef](#)]
19. Villa, P.; Stroppiana, D.; Fontanelli, G.; Azar, R.; Brivio, P.A. In-season mapping of crop type with optical and X-band SAR data: A classification tree approach using synoptic seasonal features. *Remote Sens.* **2015**, *7*, 12859–12886. [[CrossRef](#)]
20. Khabbazan, S.; Vermunt, P.; Steele-Dunne, S.; Ratering Arntz, L.; Marinetti, C.; van der Valk, D.; Iannini, L.; Molijn, R.; Westerdijk, K.; van der Sande, C. Crop monitoring using Sentinel-1 data: A case study from The Netherlands. *Remote Sens.* **2019**, *11*, 1887. [[CrossRef](#)]
21. Guo, J.; Li, H.; Ning, J.; Han, W.; Zhang, W.; Zhou, Z.S. Feature dimension reduction using stacked sparse auto-encoders for crop classification with multi-temporal, quad-pol SAR data. *Remote Sens.* **2020**, *12*, 321. [[CrossRef](#)]
22. 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](#)]
23. 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](#)]
24. Sonobe, R.; Yamaya, Y.; Tani, H.; Wang, X.; Kobayashi, N.; Mochizuki, K.I. Assessing the suitability of data from Sentinel-1A and 2A for crop classification. *GISci. Remote Sens.* **2017**, *54*, 918–938. [[CrossRef](#)]
25. Zhao, W.; Qu, Y.; Chen, J.; Yuan, Z. Deeply synergistic optical and SAR time series for crop dynamic monitoring. *Remote Sens. Environ.* **2020**, *247*, 111952. [[CrossRef](#)]
26. Guo, L.; Zhao, S.; Gao, J.; Zhang, H.; Zou, Y.; Xiao, X. A novel workflow for crop type mapping with a time series of synthetic aperture radar and optical images in the Google Earth Engine. *Remote Sens.* **2022**, *14*, 5458. [[CrossRef](#)]
27. Ren, T.; Liu, Z.; Zhang, L.; Liu, D.; Xi, X.; Kang, Y.; Zhao, Y.; Zhang, C.; Li, S.; Zhang, X. Early identification of seed maize and common maize production fields using Sentinel-2 images. *Remote Sens.* **2020**, *12*, 2140. [[CrossRef](#)]
28. Yi, Z.; Jia, L.; Chen, Q.; Jiang, M.; Zhou, D.; Zeng, Y. Early-season crop identification in the Shiyang River Basin using a deep learning algorithm and time-series Sentinel-2 data. *Remote Sens.* **2022**, *14*, 5625. [[CrossRef](#)]
29. Lin, C.; Zhong, L.; Song, X.-P.; Dong, J.; Lobell, D.B.; Jin, Z. Early- and in-season crop type mapping without current-year ground truth: Generating labels from historical information via a topology-based approach. *Remote Sens. Environ.* **2022**, *274*, 112994. [[CrossRef](#)]
30. Yan, Y.; Ryu, Y. Exploring Google Street View with deep learning for crop type mapping. *ISPRS J. Photogramm. Remote Sens.* **2021**, *146*, 278–296. [[CrossRef](#)]
31. Liu, P.; Li, J.; Wang, L.; He, G. Remote sensing data fusion with generative adversarial networks: State-of-the-art methods and future research directions. *IEEE Geosci. Remote Sens. Mag.* **2022**, *10*, 295–328. [[CrossRef](#)]
32. Zhang, R.; Isola, P.; Efros, A.A. Colorful image colorization. *arXiv* **2016**, arXiv:1603.08511.
33. Pathak, D.; Krahenbuhl, P.; Donahue, J.; Darrell, T.; Efros, A.A. Context encoders: Feature learning by inpainting. *arXiv* **2016**, arXiv:1604.07379.
34. Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. *arXiv* **2016**, arXiv:1611.07004.

35. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *arXiv* **2014**, arXiv:1406.2661. [[CrossRef](#)]
36. Mirza, M.; Osindero, S. Conditional generative adversarial nets. *arXiv* **2014**, arXiv:1411.1784.
37. Kim, K.; Kim, J.-H.; Moon, Y.-J.; Park, E.; Shin, G.; Kim, T.; Kim, Y.; Hong, S. Nighttime reflectance generation in the visible band of satellites. *Remote Sens.* **2019**, *11*, 2087. [[CrossRef](#)]
38. Park, J.-E.; Kim, G.; Hong, S. Green band generation for advanced baseline imager sensor using Pix2Pix with advanced baseline imager and advanced Himawari imager observations. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 6415–6423. [[CrossRef](#)]
39. Han, K.-H.; Jang, J.-C.; Ryu, S.; Sohn, E.-H.; Hong, S. Hypothetical visible bands of advanced meteorological imager onboard the Geostationary Korea Multi-Purpose Satellite-2A using data-to-data translation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 8378–8388. [[CrossRef](#)]
40. He, W.; Yokoya, N. Multi-temporal sentinel-1 and-2 data fusion for optical image simulation. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 389. [[CrossRef](#)]
41. Bermudez, J.D.; Happ, P.N.; Oliveira, D.A.B.; Feitosa, R.Q. SAR to optical image synthesis for cloud removal with generative adversarial networks. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* **2018**, *IV-1*, 5–11. [[CrossRef](#)]
42. Bermudez, J.D.; Happ, P.N.; Feitosa, R.Q.; Oliveira, D.A.B. Synthesis of multispectral optical images from SAR/optical multitemporal data using conditional generative adversarial networks. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 1220–1224. [[CrossRef](#)]
43. Li, Y.; Fu, R.; Meng, X.; Jin, W.; Shao, F. A SAR-to-optical image translation method based on conditional generation adversarial network (cGAN). *IEEE Access* **2020**, *8*, 60338–60343. [[CrossRef](#)]
44. Zhang, J.; Zhou, J.; Lu, X. Feature-guided SAR-to-optical image translation. *IEEE Access* **2020**, *8*, 70925–70937. [[CrossRef](#)]
45. Zhu, J.-Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV 2017), Venice, Italy, 22–29 October 2017; pp. 2242–2251.
46. Fuentes Reyes, M.; Auer, S.; Merkle, N.; Henry, C.; Schmitt, M. SAR-to-optical image translation based on conditional generative adversarial networks—Optimization, opportunities and limits. *Remote Sens.* **2019**, *11*, 2067. [[CrossRef](#)]
47. Wang, L.; Xu, X.; Yu, Y.; Yang, R.; Gui, R.; Xu, Z.; Pu, F. SAR-to-optical image translation using supervised cycle-consistent adversarial networks. *IEEE Access* **2019**, *7*, 129136–129149. [[CrossRef](#)]
48. Yi, Z.; Zhang, H.; Tan, P.; Gong, M. DualGAN: Unsupervised dual learning for image-to-image translation. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV 2017), Venice, Italy, 22–29 October 2017; pp. 2868–2876.
49. 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](#)]
50. Feng, S.; Zhao, J.; Liu, T.; Zhang, H.; Zhang, Z.; Guo, X. Crop type identification and mapping using machine learning algorithms and sentinel-2 time series data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 3295–3306. [[CrossRef](#)]
51. Yi, Z.; Jia, L.; Chen, Q. Crop classification using multi-temporal Sentinel-2 data in the Shiyang River Basin of China. *Remote Sens.* **2020**, *12*, 4052. [[CrossRef](#)]
52. Nasirzadehdzaji, R.; Sanli, F.B.; Abdikan, S.; Cakir, Z.; Sekertekin, A.; Ustuner, M. Sensitivity analysis of multi-temporal Sentinel-1 SAR parameters to crop height and canopy coverage. *Appl. Sci.* **2019**, *9*, 655. [[CrossRef](#)]
53. Zhang, Q.; Liu, X.; Liu, M.; Zou, X.; Zhu, L.; Ruan, X. Comparative analysis of edge information and polarization on SAR-to-optical translation based on conditional generative adversarial networks. *Remote Sens.* **2021**, *13*, 128. [[CrossRef](#)]
54. USDA Foreign Agricultural Service. Available online: <https://fas.usda.gov/commodities> (accessed on 25 January 2023).
55. Green, T.R.; Kipka, H.; David, O.; McMaster, G.S. Where is the USA Corn Belt, and how is it changing? *Sci. Total Environ.* **2018**, *618*, 1613–1618. [[CrossRef](#)] [[PubMed](#)]
56. Diao, C. Remote sensing phenological monitoring framework to characterize corn and soybean physiological growing stages. *Remote Sens. Environ.* **2020**, *248*, 111960. [[CrossRef](#)]
57. Han, W.; Yang, Z.; Di, L.; Mueller, R. CropScape: A Web service based application for exploring and disseminating US conterminous geospatial cropland data products for decision support. *Comput. Electron. Agric.* **2012**, *84*, 111–123. [[CrossRef](#)]
58. CropScape—Cropland Data Layer. Available online: <https://nassgeodata.gmu.edu/CropScape> (accessed on 10 January 2022).
59. ESA, Copernicus Open Access Hub. Available online: <https://scihub.copernicus.eu> (accessed on 10 January 2022).
60. ESA, SNAP. Available online: <https://step.esa.int/main/toolboxes/snap> (accessed on 10 January 2022).
61. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Adv. Neural Inform. Process. Syst.* **2012**, *25*, 1097–1105. [[CrossRef](#)]
62. Shao, Z.; Lu, Z.; Ran, M.; Fang, L.; Zhou, J.; Zhang, Y. Residual encoder-decoder conditional generative adversarial network for pansharpening. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 1573–1577. [[CrossRef](#)]
63. Kim, Y.; Jackson, T.; Bindlish, R.; Lee, H.; Hong, S. Radar vegetation index for estimating the vegetation water content of rice and soybean. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 564–568.
64. Mandal, D.; Kumar, V.; Ratha, D.; Dey, S.; Bhattacharya, A.; Lopez-Sanchez, J.M.; Rao, Y.S. Dual polarimetric radar vegetation index for crop growth monitoring using sentinel-1 SAR data. *Remote Sens. Environ.* **2020**, *247*, 111954. [[CrossRef](#)]

65. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Interventions, Munich, Germany, 5–9 October 2015; pp. 234–241.
66. Zhang, C.; Zhang, H.; Zhang, L. Spatial domain bridge transfer: An automated paddy rice mapping method with no training data required and decreased image inputs for the large cloudy area. *Comput. Electron. Agric.* **2021**, *181*, 105978. [[CrossRef](#)]
67. Peng, X.; He, G.; She, W.; Zhang, X.; Wang, G.; Yin, R.; Long, T. A comparison of random forest algorithms-based forest extraction with GF-1 WFV, Landsat 8 and Sentinel-2 images. *Remote Sens.* **2022**, *14*, 5296. [[CrossRef](#)]
68. Abida, K.; Barbouchi, M.; Boudabbous, K.; Toukabri, W.; Saad, K.; Bousnina, H.; Chahe, T.S. Sentinel-2 data for land use mapping: Comparing different supervised classifications in semi-arid areas. *Agriculture* **2022**, *12*, 1429. [[CrossRef](#)]
69. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
70. Hore, A.; Ziou, D. Image quality metrics: PSNR vs. SSIM. In Proceedings of the 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010; pp. 2366–2369.
71. Zhao, H.; Chen, Z.; Jiang, H.; Jing, W.; Sun, L.; Feng, M. Evaluation of three deep learning models for early crop classification using Sentinel-1A imagery time series—A case study in Zhanjiang, China. *Remote Sens.* **2019**, *11*, 2673. [[CrossRef](#)]
72. TensorFlow. Available online: <https://tensorflow.org> (accessed on 15 August 2022).
73. Keras Documentation. Available online: <https://keras.io> (accessed on 15 August 2022).
74. NumPy. Available online: <https://numpy.org> (accessed on 15 August 2021).
75. Scikit-Learn: Machine Learning in Python. Available online: <https://scikit-learn.org> (accessed on 15 August 2022).
76. Wang, Q.; Shi, W.; Li, Z.; Atkinson, P.M. Fusion of Sentinel-2 images. *Remote Sens. Environ.* **2016**, *187*, 241–252. [[CrossRef](#)]
77. Lanaras, C.; Bioucas-Dias, J.; Galliani, S.; Baltsavias, E. Super-resolution of Sentinel-2 images: Learning a globally applicable deep neural network. *ISPRS J. Photogramm. Remote Sens.* **2018**, *146*, 305–319. [[CrossRef](#)]
78. Palsson, F.; Sveinsson, J.R.; Ulfarsson, M.O. Sentinel-2 image fusion using a deep residual network. *Remote Sens.* **2018**, *10*, 1290. [[CrossRef](#)]
79. Yang, X.; Zhao, J.; Wei, Z.; Wang, N.; Gao, X. SAR-to-optical image translation based on improved CGAN. *Pattern Recognit.* **2022**, *121*, 108208. [[CrossRef](#)]
80. Turnes, J.N.; Castro, J.D.B.; Torres, D.L.; Vega, P.J.S.; Feitosa, R.Q.; Happ, P.N. Atrous cGAN for SAR to optical image translation. *IEEE Geosci. Remote Sens. Lett.* **2020**, *19*, 1–5. [[CrossRef](#)]
81. Shi, H.; Zhang, B.; Wang, Y.; Cui, Z.; Chen, L. SAR-to-optical image translating through generate-validate adversarial networks. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [[CrossRef](#)]

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