

Article

Learning-Based Near-Infrared Band Simulation with Applications on Large-Scale Landcover Classification

Xiangtian Yuan , Jiaojiao Tian *  and Peter Reinartz 

German Aerospace Center (DLR), Münchner Str. 20, 82234 Weßling, Germany; xiangtian.yuan@dlr.de (X.Y.); peter.reinartz@dlr.de (P.R.)

* Correspondence: Jiaojiao.tian@dlr.de

Abstract: Multispectral sensors are important instruments for Earth observation. In remote sensing applications, the near-infrared (NIR) band, together with the visible spectrum (RGB), provide abundant information about ground objects. However, the NIR band is typically not available on low-cost camera systems, which presents challenges for the vegetation extraction. To this end, this paper presents a conditional generative adversarial network (cGAN) method to simulate the NIR band from RGB bands of Sentinel-2 multispectral data. We adapt a robust loss function and a structural similarity index loss (SSIM) in addition to the GAN loss to improve the model performance. With 45,529 multi-seasonal test images across the globe, the simulated NIR band had a mean absolute error of 0.02378 and an SSIM of 89.98%. A rule-based landcover classification using the simulated normalized difference vegetation index (NDVI) achieved a Jaccard score of 89.50%. The evaluation metrics demonstrated the versatility of the learning-based paradigm in remote sensing applications. Our simulation approach is flexible and can be easily adapted to other spectral bands.

Keywords: multispectral; remote sensing; NIR; RGB; cGAN; Sentinel-2; SEN12MS; robust loss; SSIM



Citation: Yuan, X.; Tian, J.; Reinartz, P. Learning-Based Near-Infrared Band Simulation with Applications on Large-Scale Landcover Classification. *Sensors* **2023**, *23*, 4179. <https://doi.org/10.3390/s23094179>

Academic Editors: Angel D. Sappa and Riad Hammoud

Received: 20 March 2023

Revised: 12 April 2023

Accepted: 18 April 2023

Published: 22 April 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

Multispectral remote sensing is one of the important means of Earth observation. It has been extensively employed to explore the physical, chemical and biological properties of the Earth's surface. In addition to visible spectra that the human visual system can perceive, multispectral sensors capture signals from additional spectral ranges. Since ground objects respond differently to light of certain wavelengths, the wider spectral range allows additional information to be extracted from ground objects.

Due to limitations of budget, technology, intended application and various other reasons, not every sensor is capable of capturing a wide range of wavelengths across the electromagnetic spectrum. Moreover, differences in the wavelength characteristics of different sensors can make it challenging to use data from multiple sensors simultaneously, necessitating the process of harmonization [1]. Some pixels might be corrupted during the data down-link from satellites, which can hinder further analysis [2]. In the case of unmanned aerial vehicle (UAV) remote sensing, low-cost, off-the-shelf cameras typically capture only visible light in the red, green, and blue wavelengths, which limits their potential for downstream applications that require near-infrared spectra, such as vegetation monitoring. As a result, researchers have modified commercial cameras to capture the NIR bands, but registration of each band is often required [3–6].

Among the common spectral bands, the near-infrared (NIR) bands have been used extensively in Earth observation tasks. In combination with visible spectra, NIR bands contain additional features of ground objects, especially vegetation. For example, indices using NIR bands have been developed for tasks such as land cover classification. These indices include the normalized difference vegetation index (NDVI) and the normalized difference water index (NDWI), which have been shown to be effective in highlighting

vegetation and open water in remotely sensed imagery [7]. In addition to identifying vegetation and water, *NIR* spectroscopy can also help detect materials such as plastics [8,9], certain minerals [10], and tree health problems [11,12]. In addition, *NIR*-derived indices have also been used in tasks such as atmospheric correction [13]. In data-hungry machine learning or deep learning methods for land cover classification, *NIR* bands are able to improve coarse ground truth or correct mislabelling for some classes that are sometimes challenging for the human eye to interpret. Therefore, the generation of additional spectral bands from known bands has potential practical applications in Earth observation but has not yet been extensively explored. However, the underlying problem is that there are no precise physical models that map a spectral response from another wavelength. Signals that the sensor receives depend on many factors, including atmospheric effects, weather condition, land cover type, and terrain, etc. Ignoring these effects, we want to test whether a simple end-to-end model is sufficient to generate additional bands from known bands on a large scale, without the myriad of input parameters of complicated models.

The generation of artificial *NIR* bands using only the visible spectrum can be considered as a nonlinear mapping problem. Neural networks have been shown to be effective in nonlinear mapping [14]. It could also be viewed as a generative problem, which can be addressed by neural networks and especially GANs. Unlike computer vision tasks, which usually have some level of abstraction from the input, our task is to ensure that the generated *NIR* band is also consistent in structure and spatial distribution. To this end, additional loss functions, such as L1 or L2, are added to the GAN loss to ensure that the output is close to the ground truth [15]. However, such losses are prone to outliers. Several robust loss functions are able to handle outliers by being less sensitive to large errors. A single robust loss function proposed by Barron [16] integrates several common robust loss functions that are controlled by a single continuous-valued parameter that can also be tuned when training neural networks.

In our preliminary work, we performed experiments on a dataset acquired in summer [17]. It was shown that a conditional GAN model was able to generate a realistic *NIR* band from the visible spectrum. The generated *NIR* band retained the original image's texture and correctly displayed the spectral characteristics of certain land cover. However, in this proof of concept work, only a relative small dataset was used for the evaluation to eliminate the potential influence of the seasons on the results. In this work, we will present more details about the method to generate artificial spectral bands from the visible bands using cGAN. The method is tested with the complete SEN12MS dataset with four full seasons. In addition to the pixel-wise evaluation, by comparing with the ground truth, a rule-based land cover classification is carried out to assess the quality of the artificially generated *NIR* band in a downstream remote sensing application.

The contribution of our work is twofold: based on our previous proof of concept work, we integrated a structural similarity loss to further improve the performance of the method; we evaluated the method with the complete four seasons of the SEN12MS dataset, validating the effectiveness and revealing the limitations of this method.

The paper is structured as follows: In Section 2, we review relevant literature; in Section 3, we introduce the model in detail and the dataset used in this study; in Section 4, we describe the experimental settings; in Section 5, the results are presented and analyzed, followed by the conclusion in Section 6.

2. Related Work

Traditional techniques for solving remote sensing problems have been continuously challenged by machine learning methods over the last decades. Machine learning methods have been extended to remote sensing images, which have some peculiarities compared to natural images. Among machine learning methods, GANs have gained increasing attention due to their versatility and performance. For example, GANs have been applied to tasks such as image super-resolution and pan-sharpening. Sajjadi et al. [18] proposed a GAN-based network to super-resolve natural images using automated texture synthesis combined

with a perceptual loss that produced realistic texture, instead of just approximating the ground truth pixel during training. Jiang et al. [19] applied a GAN-based model with an edge enhancement subnetwork for super-resolution, which effectively removed noise in the reconstructed image and enhanced the contour of ground objects. This model takes advantage of the adversarial learning strategy, which is insensitive to noise. GAN-based models have also been applied in remote sensing image pre-processing, such as image dehazing and cloud removal. Enomoto et al. [20] adopted NIR bands with an RGB image in a GAN-based framework to remove clouds from satellite images. Grohnfeldt et al. [21] fused Sentinel-1 SAR data and Sentinel-2 optical multi-spectral imagery in a GAN framework to produce cloud-free optical images, which has shown more robust performance due to the properties of the synthetic aperture radar (SAR) sensor. GAN-based models have been applied to treat many problems as generative, such as monocular height estimation [22,23], DSM refinement [24], PAN-sharpening [25], image super-resolution [26], and change detection [27].

One way of approaching the spectral band simulation problem is to treat it as a classification problem. Within this framework, the mapping process could be viewed as involving the injection of spectral signatures into corresponding classes of ground objects. In recent years, neural networks have been widely used for land cover classification and have become a popular classification paradigm in the remote sensing community. The problem could also be viewed as a generative one as well, and, thus, can be tackled by neural-network- and GAN-based methods. Therefore, it is theoretically possible to simulate spectral reflectance from other bands using neural-network-based methods. Specifically, hyperspectral or multispectral image reconstruction, in which responses at multiple electromagnetic wavelengths need to be simulated, is a hot research top in spectral simulation using deep learning methods. The paper by Fu et al. [14] proposed a network structure for hyperspectral image reconstruction from RGB bands. The network consists of a spectral sub-network, which performs the spectral nonlinear mapping, and a spatial sub-network, which models the spatial correlation. The network uses a traditional loss function (mean square error) to force the generated bands to be numerically similar to the real ones. Deng et al. [28] proposed a neural-network-based method (M2H-Net) to reconstruct hyperspectral image from arbitrary number of input bands within spectral range of 300 to 2500 nm. The method was verified by UAV and satellite data captured at different locations in China. Zhao et al. [29] used a model trained by a hyperspectral benchmark dataset WHU-Hi-Honghu HSI [30] to convert true RGB to natural RGB, which was subsequently used with its multispectral pair to train an HSCNN-R network [31] for reconstruction. The model was trained with a multi-temporal and multispectral dataset of a maize field and successfully tested on the imagery of a rice field. Like many other remote sensing issues, the hyperspectral reconstruction work can be tackled by GAN-based methods as well. Alvarez-Gila et al. [32] used a conditional GAN to reconstruct a hyperspectral image from an RGB image. The method was trained and tested on 201 natural images of 1392×1300 . Liu and Zhao [33] proposed a scale attention pyramid UNet (SAPUNet) that adopted a dilated convolution for feature extraction and an attention mechanism for feature selection. SAPW-Net was proposed in the same work, with an additional branch for boundary supervision. The work achieved improved results on the interdisciplinary Computational Vision Lab at Ben Gurion University (ICVL) dataset [34].

Due to the prohibitively high cost of hyperspectral imagery, the amount of open-source hyperspectral datasets is limited, and the size of the available datasets is relatively limited. According to the review by [35], the available open-source hyperspectral datasets are only of small to medium size [35]. However, multispectral datasets have better availability, and many satellite missions have global coverage, such as the Sentinel mission, enabling large-scale experiments and analysis.

The integration of additional NIR bands in cameras has practical applications in various remote sensing research fields, including vegetation monitoring. However, the widespread adoption of NIR-capable cameras is limited by cost and technical constraints [3]. To over-

come this, numerous researchers have modified commercial RGB cameras to enable the capture of additional near-infrared band radiation for vegetation and soil monitoring by unmanned aerial vehicles (UAVs). For instance, Rabatel et al. [3] removed the NIR blocking filter and added an external long-wavelength pass filter to a single commercial camera (Canon 500D). The optimal external filter was determined by BSOP (band simulation by orthogonal projection), which relies on known sensitivity curves of the camera. Other studies employed two cameras to capture RGB and NIR images separately [4,5], which requires accurate pixel alignment. Brown and Ssstrunk [6] created a 477-image RGB-NIR dataset captured by a modified single-lens reflex (SLR) camera. The NIR and RGB bands were registered using a feature-based alignment algorithm [36] with robust estimation of a similarity motion model. The joint entropy analysis suggested that NIR contains significantly different information from the visible bands.

Different from the methods that involve hardware modification, learning-based methods focusing on the image data using generative methods have also been studied. Our previous work [17] was one of the first to apply GANs to simulate NIR bands from RGB bands. Subsequently, several studies have been published to explore the potential of simulated NIR bands. Koshelev et al. [37] synthesized an NIR band to boost the performance of hogweed crops segmentation. Sa et al. [38] proposed a fruit detection dataset with RGB and NIR bands. Aslahishahri et al. [39] curated a dataset with blue, green, red, NIR, and red edge bands covering canola, lentil, dry bean, and wheat breeding fields. These works provide valuable datasets that can be used for specific remote sensing tasks.

3. Method and Data

In this section, details of our method will be introduced in detail. An overview of the method is shown in Figure 1.

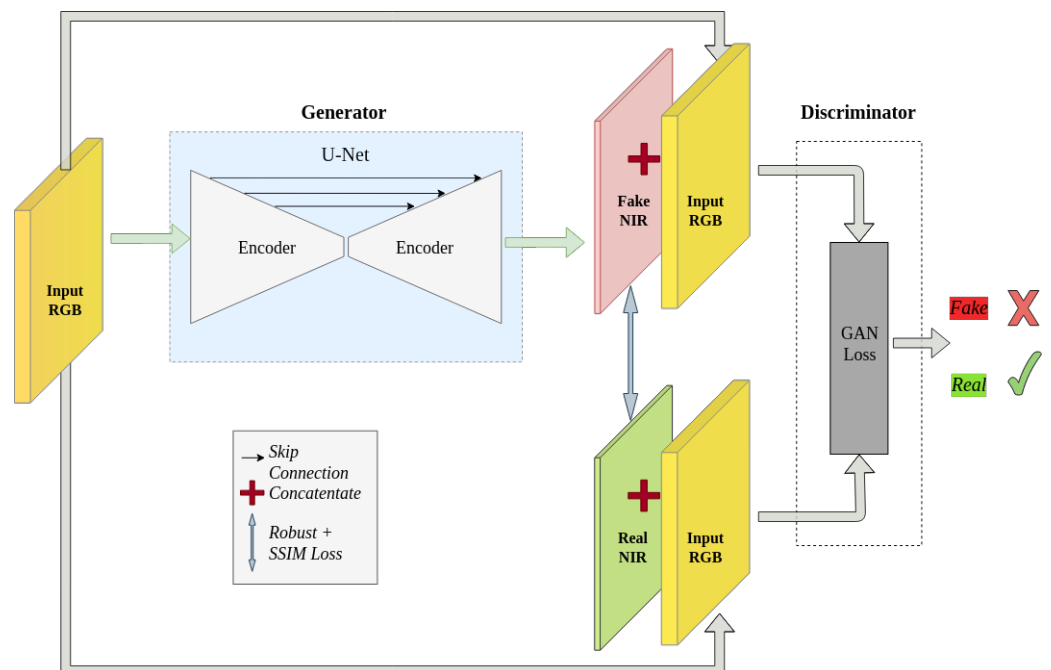


Figure 1. The workflow of the method. The GAN loss is also used for back propagation in the generator, which is not shown in the visualization. The weights of the generator and the discriminator are updated alternately.

3.1. Conditional GAN

A GAN consists of two components: a generator and a discriminator. The generator's aim is to produce fake results, while the discriminator tries to differentiate between the generator's output and the real distribution [40]. These two components work in a

competitive manner inspired by game theory, where any improvement in one leads to a corresponding deterioration in the other. In a typical setup for vision tasks, GANs take in random noise vectors from a particular distribution and produce realistic-looking images as outputs. In contrast to traditional GANs, conditional GANs (cGANs) [41] require specific input data. In the context of spectral band simulation, the network must generate the *NIR* band that is related to the RGB bands. In addition, the discriminator is also conditioned on the input RGB bands. By attaining equilibrium between the generator and the discriminator, the cGAN can produce a realistic-looking NIR band that corresponds to the input RGB bands.

3.1.1. Generator

The task of generating a realistic NIR band from RGB bands can be considered as a mapping from the visible to the near-infrared spectrum. As the input and output represent the same ground objects, they should possess matching structures, similar textures, and encode the same semantics. Many GAN generators use encoder-decoder structures that downsample the input and then gradually upsample it to the output, which leads to the loss of low-level information from earlier stages and consequently degraded details. Hence, several CNN models use long skip connections to retain fine-grained information from earlier layers that are lost during down-sampling, and, thereby, perform better for tasks that require finer spatial information. Therefore, the encoder-decoder network with a skip connection, famously known as the U-Net structure [42], is well-suited for this task. The U-Net structure comprises eight blocks in both the encoder and decoder, with each block containing convolution, batch normalization, and a leaky rectified linear unit (LeakyReLU) with a slope of 0.2 in the encoder. The decoder blocks comprise transposed convolution, batch normalization, and ReLU layers. The convolution has a filter size of 4 and a stride of 2 in both the encoder and decoder. In contrast to some conditional GANs that use Gaussian noise as input to the generator to avoid deterministic results matching the Dirac delta function [15], our task does not require much randomness, so we do not use this approach. Instead, we use dropout in the generator during both training and testing phases, which reduces stochasticity but is suitable for our task. The first three blocks in the decoder have dropout layers with a rate of 50%.

3.1.2. Discriminator

The discriminator should be neither too strong nor too weak with respect to the generator. If the discriminator is too weak, the generator can easily fool it, possibly leading to an unrealistic NIR band. On the other hand, if the discriminator is too powerful, the generator might not pass its scrutiny. There are various options available for the discriminator, depending on the intended purpose. One choice is the Markovian discriminator, also known as PatchGAN [15], as shown in Figure 2. This discriminator classifies whether a patch of size $N \times N$ in the input image is real or fake and averages all the patches in the image. It consists of several blocks of 2D convolution, batch normalization, and LeakyReLU layers, with a stride of 2 for all convolutions, except for the last two convolutions. The PatchGAN can be considered a form of texture loss, as it focuses on the local structure of the image.

Another option is the pixel discriminator, which operates on the pixel level and does not consider texture information. The kernel size and stride in the pixel discriminator are both 1, keeping the feature map size constant throughout the network. Figure 3 illustrates the architecture of the pixel discriminator. Both discriminators have a binary cross-entropy layer as their last layer, which classifies the generated image as true or false, and the results are averaged over the entire image. When training the cGAN, the weights of the discriminator are updated first, and the weights of the generator are updated subsequently.

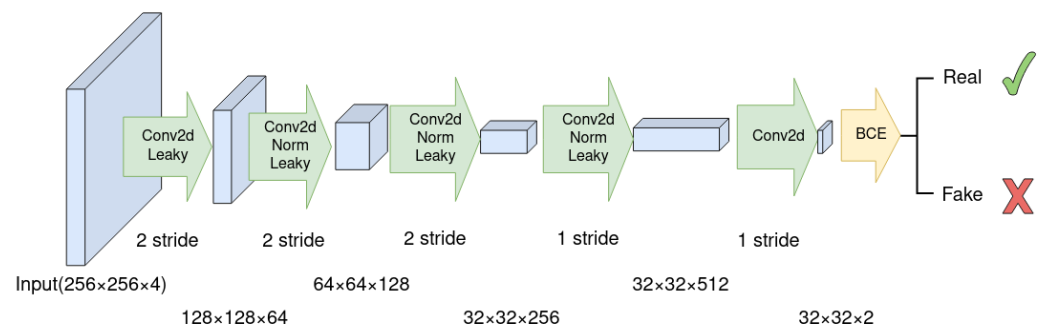


Figure 2. The PatchGAN discriminator, as shown in the diagram, is made up of several blocks comprising of 2D convolutions, a LeakyReLU activation function, and a binary cross-entropy function at the end. The first three convolutions have a filter size of 4×4 and a stride of 2. The last two convolution layers have a stride of 1 to retain the spatial resolution. The output of the discriminator after binary cross-entropy (BCE) is obtained by averaging the outputs of all the patches.

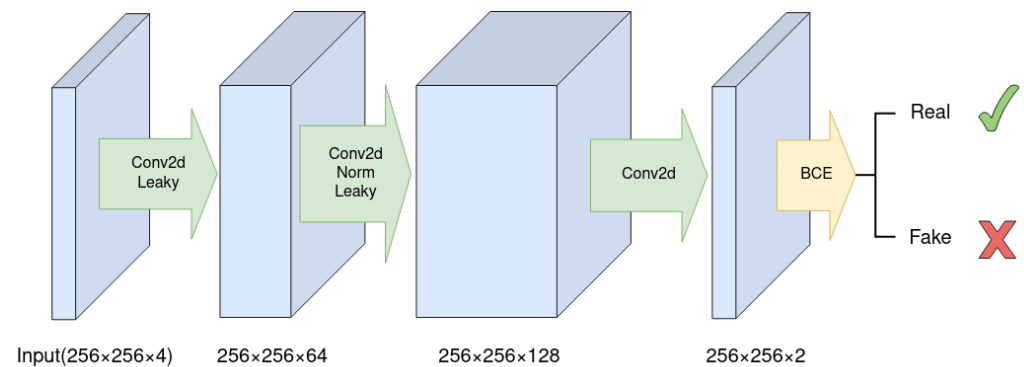


Figure 3. The PixelGAN discriminator. Each 1×1 convolution has a stride of 1. The prediction is only on pixel level. The output is averaged of all the pixels. The PixelGAN discriminator does not encode texture information.

3.2. Loss Function

The objective in our experiment consists of three terms: conditional GAN loss, robust loss and SSIM loss. The conditional GAN loss is used by both the generator and discriminator for back propagation, while the robust loss and SSIM are only considered by the generators.

3.2.1. GAN Loss

In unconditional GANs, the generator G maps a random noise z that follows a certain probability distribution to the output y . The discriminator tries to predict the generator's output as fake. The loss function of the unconditional GANs can be written as:

$$L_{GAN}(G, D) = \mathbb{E}_y[\text{Log}D(y)] + \mathbb{E}_z[\text{log}(1 - D(G(z)))] \quad (1)$$

Conditional GANs learn mapping from both random noise z and input image x : $G : \{x, z\} \rightarrow y$. In our case the noise is realized by dropout. The discriminator D is trained adversarially against the generator G to distinguish between the real image and the generated image. The objective function of conditional GANs can be expressed as:

$$L_{cGAN}(G, D) = \mathbb{E}_{x,y}[\text{Log}D(x, y)] + \mathbb{E}_{x,z}[\text{log}(1 - D(G(x, z)))] \quad (2)$$

In detail, the generator tries to maximize $\log(D(G(x, z)))$ instead of minimizing the objective $\mathbb{E}_{x, z}[\log(1 - D(G(x, z)))]$, as suggested in the original GAN paper [40].

3.2.2. Robust Loss

Combining GAN objectives with traditional losses, such as L1 and L2, has been proven beneficial [43]. In cGAN, the generator should also produce output that is close to the ground truth numerically. Traditional loss functions, such as L1 and L2, are susceptible to noise in the data. They are heavily skewed by outliers that might result from many factors, such as sensor errors, data transmission errors or pre-processing errors. Robustness is a crucial property that is desired in machine learning models. There are several robust loss functions that demonstrate reduced sensitivity to large errors, such as Cauchy/Lorentzian [44], Geman-McClure [45], Welsch [46], Charbonnier [47] and generalized Charbonnier [48]. When the loss is large, some of these loss functions can have a saturated or decreased gradient. To address this issue, a robust loss function has been proposed by Barron [16], which includes several common robust loss functions as a subset. This loss function has the capability to adjust its robustness continuously during training through a parameter α . The general form of the robust loss function can be expressed as:

$$f(x, \alpha, c) = \frac{|\alpha - 2|}{\alpha} \left(\left(\frac{(x/c)^2}{|\alpha - 2|} + 1 \right)^{\alpha/2} - 1 \right) \quad (3)$$

This loss function is a versatile generalization of several existing loss functions mentioned above. In Equation (3), $c > 0$ represents the scale parameter that determines the size of the quadratic bowl near $x = 0$. Additionally, the robust loss function can be used to construct a general probability distribution, where the negative log-likelihood (NLL) of the probability density is a shifted version of the robust loss function. The distribution is written as:

$$p(x|\mu, \alpha, c) = \frac{1}{cZ(\alpha)} \exp(-f(x - \mu, \alpha, c)) \quad (4)$$

$Z(\alpha)$ in Equation (4) is a partition function defined as:

$$Z(\alpha) = \int_{-\infty}^{\infty} \exp(-f(x - \mu, \alpha, c)) \quad (5)$$

To simplify the partition function, the logarithmic of $Z(\alpha)$ is approximated using the cubic Hermit spline. The monotonicity of the general loss with respect to α will result in simply minimizing α to reduce the cost of outliers. Therefore, the final robust loss function is the NLL of the general distribution (Equation (4)) instead of the general loss function (Equation (3)). The values for c and μ are 1 and 0, respectively. For more details, readers are referred to the paper from Barron [16]. The objective for cGAN with robust loss function can be expressed as:

$$L_{cGAN}(G, D) + \lambda L_{Robust}(G, \alpha) \quad (6)$$

$$L_{Robust}(G, \alpha) = NLL(p(x|0, \alpha, 1)) \quad (7)$$

3.2.3. Structural Loss

In addition to the traditional loss and the robust loss, we integrate an additional structural loss function during training. The structural similarity index (SSIM) is a widely used metric for evaluating the perceived quality of generated images [49]. The SSIM index is a comprehensive metric for measuring image similarity within a defined window as it takes into account luminance (l), contrast (c), and structure (s). It is defined as the weighted combination of the three comparative measures:

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma \quad (8)$$

In Equation (8), the l , c , and s are defined as follow:

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (9)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (10)$$

$$s(x, y) = \frac{2\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (11)$$

In Equation (8), α, β, γ define the relative importance of the three components. In Equations (9)–(11), μ_x and μ_y are the intensity values of the two windows for comparison; σ_x and σ_y are the standard deviations; C_1, C_2 , and C_3 are constants to avoid issue caused by zero denominator instability, and are related to the dynamic range of pixel values. A Gaussian sliding window with size of 11 and $\sigma = 1.5$ is applied to the image. If $C_3 = C_2/2$, and α, β and γ are all equal to 1, Equation (8) simplifies to:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (12)$$

The SSIM is integrated in the final loss function defined as:

$$L = L_{cGAN}(G, D) + \lambda L_{Robust}(G, \alpha) + (1 - SSIM) \quad (13)$$

3.3. Dataset

To validate the method, we used the SEN12MS dataset, which is based on the Sentinel-1 and Sentinel-2 datasets [50]. The Sentinel-2 data from the SEN12MS dataset are level 1-C top of atmosphere (TOA) reflectance products. The TOA reflectance is the digital number (DN) divided by 10,000. This data format saves more memory compared with using the floating point format. It should be noted that this dataset is not atmospherically corrected. A cloud filter was applied during the data selection process to make sure the effect of clouds was minimized. The imagery has in total 13 bands with the spatial resolution ranging from 10 m to 60 m. In our experiment, we selected the red (R), green (G), blue (B) and near-infrared (NIR) bands with 10 m resolution for the experiment. Details of the band information can be found in Table 1. The dataset encompasses diverse areas, including desert, field, forests, urban areas and water bodies. The images in SEN12MS are distributed globally with varying longitudes and latitudes, and is categorised by seasons, as illustrated in Figure 4. The dataset is categorized into the four meteorological seasons of the northern hemisphere: winter (1 December 2016 to 28 February 2017), spring (1 March 2017 to 30 May 2017), summer (1 June 2017 to 31 August 2017), and fall (1 September 2017 to 30 November 2017).

Table 1. Data description from Sentinel-2 user handbook [51]. The wavelength is the center wavelength.

Bands	Wavelength (nm)		Bandwidth (nm)	Resolution (m)
	S2A	S2B		
Red	664.6	664.9	31	10
Green	559.8	559.0	36	10
Blue	492.4	492.1	66	10
NIR	832.8	832.9	106	10

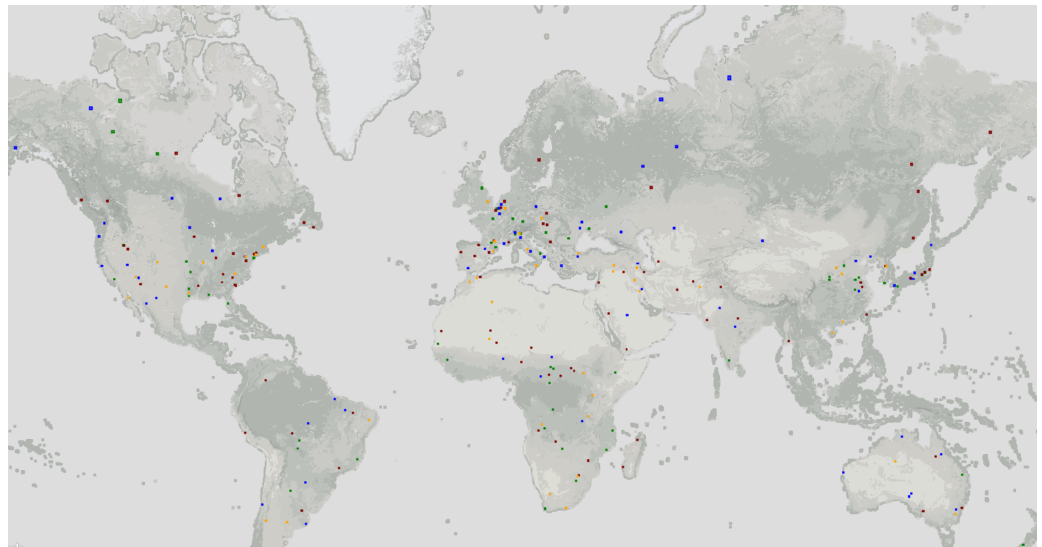


Figure 4. Visualization of the region of interest (ROIs) in the dataset. The visualization is realized in Google Earth Engine [52]. The colors represent different seasons: ■ Spring ■ Summer ■ Autumn ■ Winter.

4. Experimental Section

In our previous work [17], we verified the plausibility of using a cGAN to generate an artificial NIR band, and achieved the best numeric results using the cGAN with a pixel discriminator and robust loss. However, only 300 images in summer were used for evaluation, which was insufficient for deep learning research. In this paper, we tested the method on the full autumn dataset (September 2017 to 30 November 2017) first, and then expanded the experiment to the full dataset (1 December to 30 November 2017).

4.1. One Season Experiment

To test the plausibility of the method, we first experimented with satellite images acquired in autumn to minimize the effect of land cover seasonal variance. Some land cover types, such as vegetation and water bodies, display different properties in different seasons. We selected all images acquired in autumn in the northern hemisphere for the training and testing. The images in the southern hemisphere were not selected because the season was the opposite. Among the training data, 10 percent of images were used for validation. The patch size was 256×256 . The data fed into the network were normalized by the band-specific mean and standard deviation calculated in advance. The model was trained for 50 epochs with learning rate decay from the 25th epoch until it reached 0 in the last epoch. The λ in Equation (13) was set to 10 empirically. The initial learning rate was set to 0.002. The model was trained on a TITAN Xp GPU with 12 GB memory. In this experiment, we replaced all the batch normalization with instance normalization since the former produced unsatisfactory results, especially in the rule-based classification. We did not observe this issue in our previous work [17], possibly due to the small amount of test data that did not have large intra-variance.

4.2. Full Season Experiment

After experimenting with data from a single season, we moved further to using the full season data to test our method. Within the full season data, we could clearly see the distinct characteristics of the same land cover type in different seasons. We followed the same training protocol as in the one season experiment, except for training and testing with the full season data. The number of training images was 135,133 in total, among which 121,620 images were used for training and 13,513 for validation. The total number of test images was 45,529.

4.3. Evaluation Metrics

Various quantitative evaluation metrics were employed to assess the accuracy of our model. As is typical for GAN applications, our method aims to produce realistic-looking results. However, we did not assess the realism of the generated images through visual perception experiments, which are commonly used in other computer vision applications. In this subsection, we provide a detailed introduction to the evaluation metrics used in our study.

4.3.1. Mean Absolute Error

Mean absolute error is a very common metric. It is intuitive and easy to calculate. It is defined as:

$$MAE(x, y) = \frac{\sum_{i=1}^n |\hat{y}_i - y_i|}{n} \quad (14)$$

In Equation (14), y_i is the true value and \hat{y}_i is the predicted value. We use this metric to evaluate the quality of the generated images. The smaller MAE is, the more similar the generated image is to the original one numerically.

4.3.2. Normalized Root Mean Squared Error

In addition to *MAE*, we also use *NRMSE*. It has many variants. In our evaluation, *NRMSE* is the root mean square error normalized by the difference between the minimum and the maximum of the theoretical NIR reflectance. In our case, the theoretical reflectance ranged from 0 to 1.

$$NRMSE(x, y) = \frac{1}{y_{max} - y_{min}} \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (15)$$

4.3.3. Structural Similarity Index

The structural similarity index (SSIM) is used to evaluate the perceptual quality of simulated images based on the assumption that human visual perception is highly adapted for extracting structural information [49]. The definition is the same as for the SSIM loss function. See Section 3.2.3 for the detailed definition.

4.3.4. Normalized Difference Vegetation Index (NDVI)

Different vegetation types demonstrate dissimilar responses to spectral signals, and phenology studies have shown that vegetation has a distinct response to seasonal variation. Therefore, vegetation indices are useful tools for analysing landcover types, especially vegetation. We evaluated the artificial *NIR* band using *NDVI* to test how well it could represent vegetation. It is calculated as:

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (16)$$

4.3.5. Normalized Difference Water Index (NDWI)

Another useful index based on spectral bands is *NDWI*. It has shown a significant capacity for highlighting water from remote sensing imagery. In the dataset, water bodies are a common landcover type. It can be calculated as:

$$NDWI = \frac{Green - NIR}{Green + NIR} \quad (17)$$

4.3.6. NDVI Based Classification

In addition to approximating the reflectance values, we expected the near-infrared channel to accurately reflect the characteristics of different land cover types. Specifically, the generated *NDVI* should have lower values for water bodies and higher values for vegetation areas. To evaluate this, we implemented a simple rule-based classification scheme that quantizes the *NDVI* and derives four classes: water, barren land, low vegetation, and high vegetation. The class definitions are shown in Equation (18). It is important to note that this classification scheme overgeneralizes all land cover types. However, since our goal was to assess whether the artificial *NIR* can distinguish classes with distinctive spectral characteristics, this classification scheme still provides a meaningful evaluation method. The four classes are derived according to the following rule:

$$\begin{aligned}
 \text{Pixel} &= \text{Water}, \quad \text{if:} \\
 &\quad -1 \leq \text{NDVI} < -0.1 \\
 \text{Pixel} &= \text{Barren}, \quad \text{if:} \\
 &\quad -0.1 \leq \text{NDVI} < 0.1 \\
 \text{Pixel} &= \text{LowVegetation}, \quad \text{if:} \\
 &\quad 0.1 \leq \text{NDVI} < 0.4 \\
 \text{Pixel} &= \text{HighVegetation}, \quad \text{if:} \\
 &\quad 0.4 \leq \text{NDVI} \leq 1.0
 \end{aligned} \tag{18}$$

We evaluated this 4-class classification result using the Jaccard index, also known as the intersection over union (IoU) score, which is widely used in semantic segmentation tasks. It is defined as:

$$J = \frac{|x \cap y|}{|x \cup y|} \tag{19}$$

The Jaccard Index is the area of intersection between prediction (x) and ground truth (y) divided by the area of union of prediction and ground truth. It provides a fair evaluation even when the class distributions are unbalanced.

5. Results and Discussion

5.1. Result

5.1.1. Results of Single Season Experiment

Surprisingly, a serendipitous outcome of the method arose during the analysis of the images with the worst performance. As shown in Figure 5, the test images had blank values that were probably caused by errors in the data down-link or data processing. With the complete RGB bands, the model was able to recover the corrupted *NIR* band. This discovery suggested another use case for the method. Representative image patches are presented in Figures 6–9, which are located in the United States, Russia, Canada, and Pakistan. The images are visualized in pseudo-color and the Jet color map is used for indices visualization. In general, the results appear realistic without distinguishable fake traits. As shown in Table 2, the cGAN-PixelID model with GAN loss, robust loss, and structural loss outperformed the other model variants. In Figures 6–9, the generated *NIR* bands can highlight ground objects that are salient in *NDVI* and *NDWI*. However, some inaccuracies are also observed. For example, in Figure 7f, only a portion of the river is highlighted. Histograms of the real and artificial *NIR* bands show that, while the fake *NIR* band can capture the overall distribution of the data, there are still significant shifts from the real distribution in some intervals.

5.1.2. Results of Multi-Season Experiment

We conducted further experiments using full season data to test the performance of our model under seasonal influence. The results are presented in Table 2 and Figures 10–13,

which demonstrate that the model was able to produce reasonable results despite the varying responses of ground objects to seasonal changes. Notably, using the full season dataset achieved better quantitative results than using only autumn data. However, in some cases, the model failed to accurately capture the characteristics of the ground objects. For example, as shown in Figure 10c, small water bodies were not correctly reflected in the generated *NDWI*. After inspecting the satellite image of this area, we discovered that these water bodies had a similar color to vegetation due to high pollution levels, the presence of algae or aquatic plants, leading to the model's failure to generalize and produce an *NIR* response similar to vegetation. Another error can be observed in Figure 12c,f, where the real and fake *NDWI* contradict each other. Upon visual inspection of the region of interest using Google Earth, we could not identify the object but can only assume that it was a man-made object that strongly reflected the *NIR* band. Similar to the single season experiment, the generated *NIR* band generally approximated the overall distribution of the real data. However, in some intervals, the shift from the real distribution was still noticeable.

Table 2. The results of autumn and full season. G, R, S denote GAN loss, robust loss and structural similarity loss, respectively.

Network	Loss	MAE	MAE	MAE	NMSE	SSIM	IoU
		($\times 10^{-3}$) <i>NIR</i>	($\times 10^{-3}$) <i>NDVI</i>	($\times 10^{-3}$) <i>NDWI</i>	(%) <i>NIR</i>	(%)	(%)
Autumn							
cGAN-PixelD	G + R	22.91	32.01	33.90	2.88	90.88	84.41
	G + R + S	22.92	31.70	33.56	2.88	90.95	84.37
cGAN-PatchD	G + R	25.46	35.24	37.31	3.19	87.90	82.71
	G + R + S	25.02	34.48	36.48	3.14	88.30	83.12
Full Season							
cGAN-PixelD	G + R + S	23.78	28.06	30.40	3.00	89.98	89.50

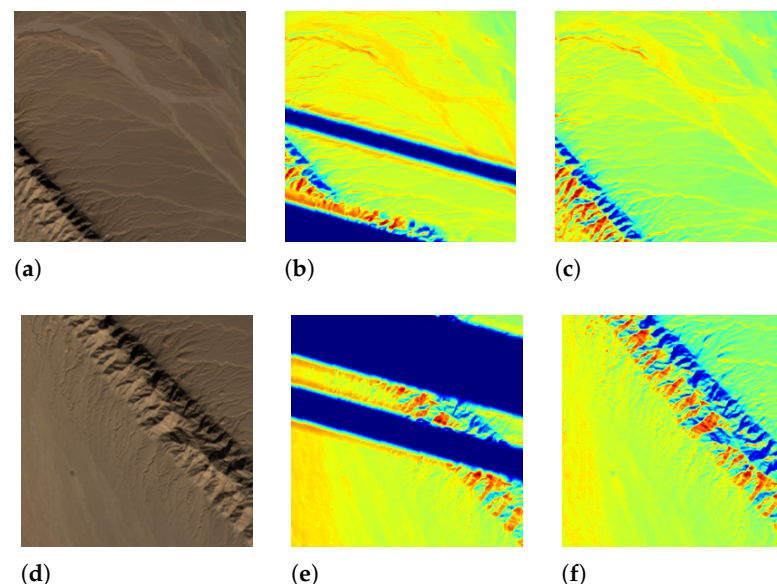


Figure 5. Selected example patches from Sentinel-2 dataset (a,d). Artifacts can be seen in the real *NIR* band (b,e). The corrupted pixels in the *NIR* band are recovered by the model (c,f).

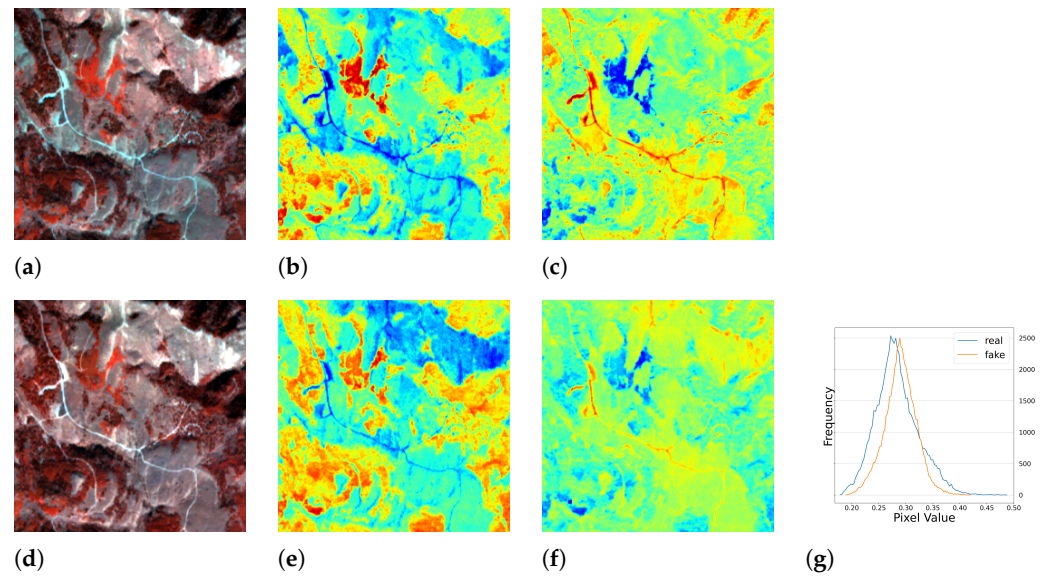


Figure 6. The single season result of cGAN-PixelID with robust loss. The ROI is located in Idaho, USA. (a) pseudo color image with the real NIR band; (b) NDVI from the real NIR band; (c) NDWI from the real NIR band; (d) pseudo color image with the fake NIR band; (e) NDVI from the fake NIR band; (f) NDWI from the fake NIR band; (g) histogram plot of the real and fake NIR bands.

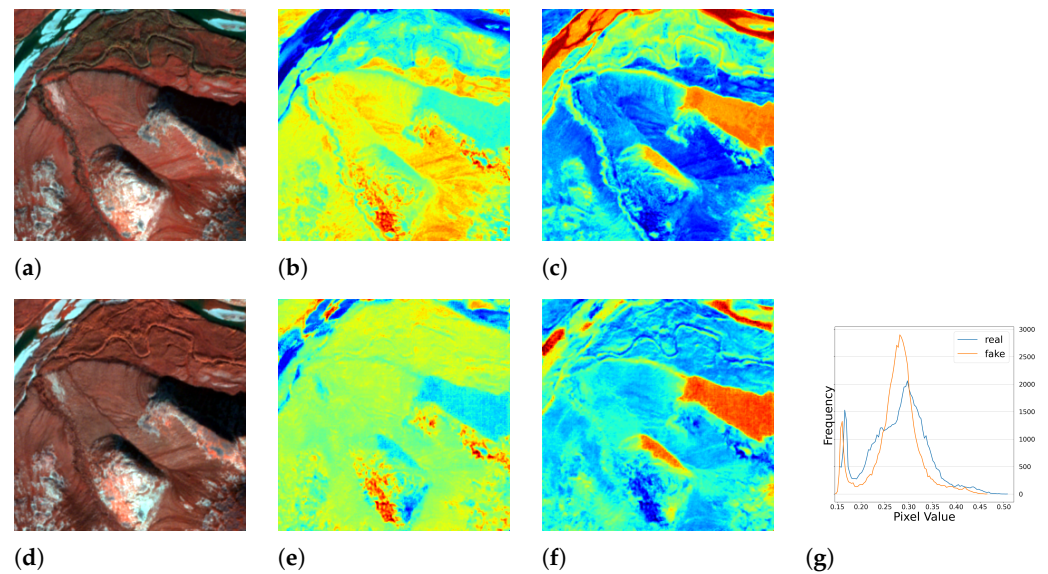


Figure 7. The single season result of cGAN-PixelID with robust loss. The ROI is located in the northeast of Russia. (a) pseudo color image with the real NIR band; (b) NDVI from the real NIR band; (c) NDWI from the real NIR band; (d) pseudo color image with the fake NIR band; (e) NDVI from the fake NIR band; (f) NDWI from the fake NIR band; (g) histogram plot of the real and fake NIR bands.

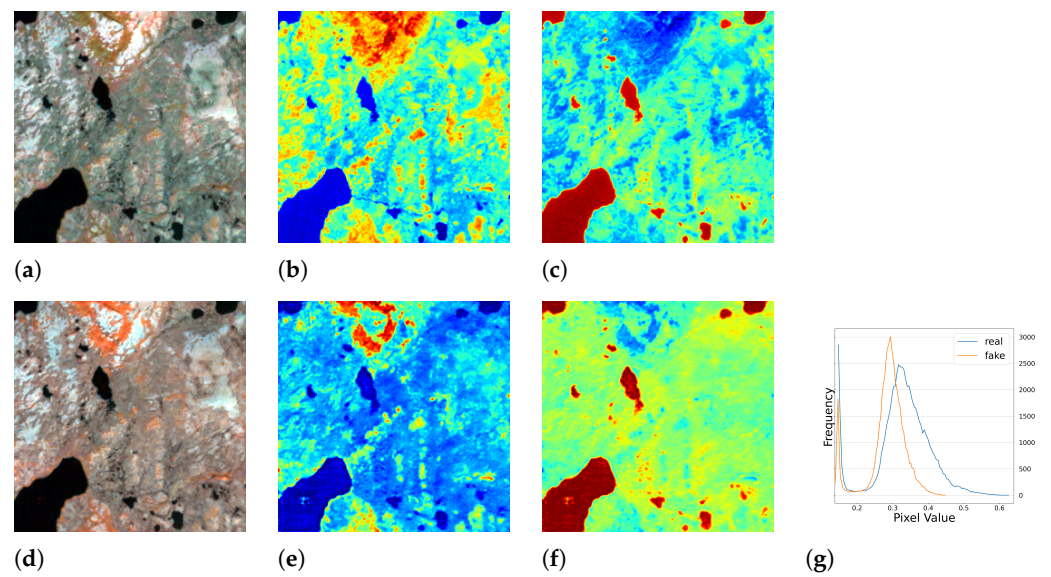


Figure 8. The single season result of cGAN-PixelD with robust loss. The ROI is located in Newfoundland and Labrador, Canada. (a) pseudo color image with the real NIR band; (b) NDVI from the real NIR band; (c) NDWI from the real NIR band; (d) pseudo color image with the fake NIR band; (e) NDVI from the fake NIR band; (f) NDWI from the fake NIR band; (g) histogram plot of the real and fake NIR bands.

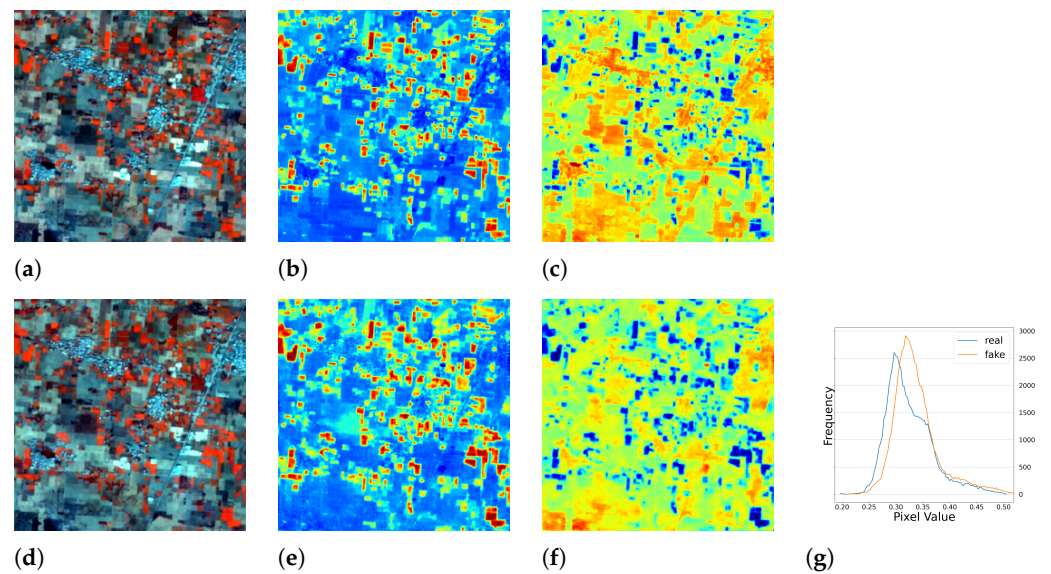


Figure 9. The single season result of cGAN-PixelD with robust loss. The ROI is located in Sialkot, Pakistan. (a) pseudo color image with the real NIR band; (b) NDVI from the real NIR band; (c) NDWI from the real NIR band; (d) pseudo color image with the fake NIR band; (e) NDVI from the fake NIR band; (f) NDWI from the fake NIR band; (g) histogram plot of the real and fake NIR bands.

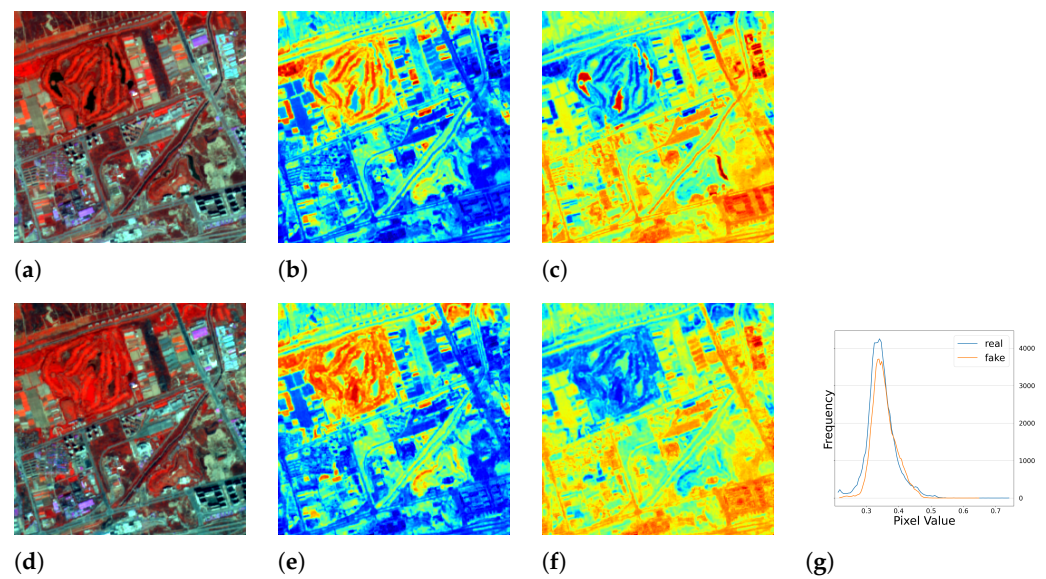


Figure 10. The full season result of cGAN-PixelD with robust loss, spring. The ROI is located in Xi'an, China. (a) pseudo color image with the real NIR band; (b) NDVI from the real NIR band; (c) NDWI from the real NIR band; (d) pseudo color image with the fake NIR band; (e) NDVI from the fake NIR band; (f) NDWI from the fake NIR band; (g) histogram plot of the real and fake NIR bands.

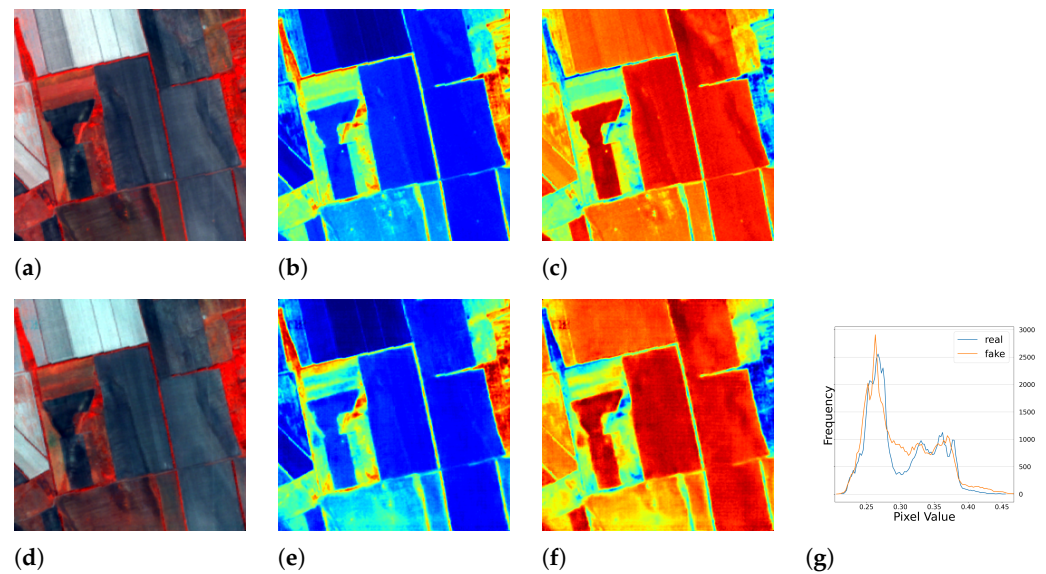


Figure 11. The full season result of cGAN-PixelD with robust loss, summer. The ROI is located in Novosavys'ts'ke, Ukraine, near the border with Moldova. (a) pseudo color image with the real NIR band; (b) NDVI from the real NIR band; (c) NDWI from the real NIR band; (d) pseudo color image with the fake NIR band; (e) NDVI from the fake NIR band; (f) NDWI from the fake NIR band; (g) histogram plot of the real and fake NIR bands.

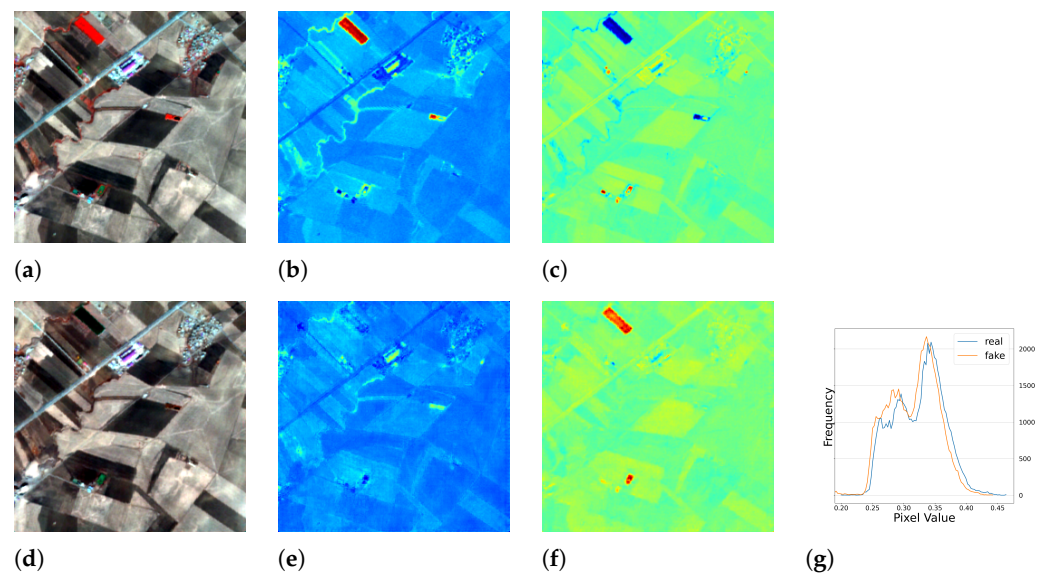


Figure 12. The full season result of cGAN-PixelD with robust loss, autumn. The ROI is located in Qararnaz, Iraq. (a) pseudo color image with the real NIR band; (b) NDVI from the real NIR band; (c) NDWI from the real NIR band; (d) pseudo color image with the fake NIR band; (e) NDVI from the fake NIR band; (f) NDWI from the fake NIR band; (g) histogram plot of the real and fake NIR bands.

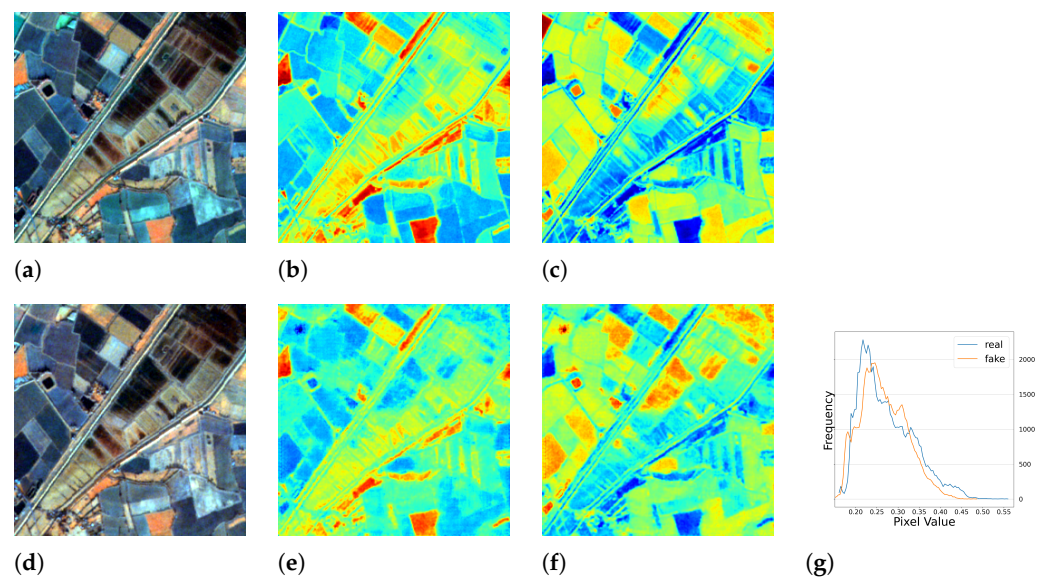


Figure 13. The full season result of cGAN-PixelD with robust loss, winter. The ROI is located in East Cambridgeshire, the United Kingdom. (a) pseudo color image with the real NIR band; (b) NDVI from the real NIR band; (c) NDWI from the real NIR band; (d) pseudo color image with the fake NIR band; (e) NDVI from the fake NIR band; (f) NDWI from the fake NIR band; (g) histogram plot of the real and fake NIR bands.

5.1.3. Results of Different Satellite Experiment

We also conducted tests to evaluate the generalization ability of our model trained with the Sentinel-2 dataset on the Landsat 8 dataset. As we expected, the model failed to generalize well to the Landsat 8 dataset. This difference could be attributed to various factors, such as differences in illumination, atmospheric or sensor conditions. For instance, Sentinel-2 and Landsat 8 OLI have slightly different bandpasses, and different measuring times can result in different solar angles. Additionally, the ground sampling distance of

Landsat 8 is 30 m for corresponding bands, which can cause mixed pixels to have a different spectral response than pure ones.

5.2. Reflections on Deficiency of the Method

From the previous section, we can see that the proposed method can work reasonably well for the same sensor. The numeric evaluation demonstrated the effectiveness of the method and its potential to some degree. The model is an end-to-end trainable black box, which is difficult to interpret. However, instead of reiterating the effectiveness of the method, we want to focus more on the limitations. We hereby address some aspects that need special consideration in future work.

5.2.1. Discussion on Atmospheric Correction

In remote sensing datasets, ground objects often exhibit vastly different responses to the sensor when atmospheric effects are present, adding to the complexity of the problem. In our experiment, we utilized the non-atmospheric corrected dataset for several reasons. Firstly, the primary objective of this research was to test the feasibility of utilizing GANs to simulate the *NIR* signal received by the sensor, irrespective of whether it was surface reflectance or radiance. Therefore, atmospheric correction was not required in this problem setting. Secondly, the *NIR* band we aimed to synthesize had a wavelength of approximately 800 nm, and its response to atmospheric effects was lower than that of the RGB bands. Even if the atmospheric effects were substantial in some scenes, the signal received by the sensor for the same ground object would not differ significantly. Furthermore, we employed a dataset that excluded images with severe haze or cloud effects, minimizing the influence of the atmosphere on the signal received by the sensor. Lastly, atmospheric correction techniques are not always 100 percent accurate and require various parameters, thereby complicating the problem. In some remote sensing applications, atmospheric correction is not essential. For instance, it has been argued that atmospheric correction does not improve land cover classification accuracy [53]. In multi-temporal change detection, atmospheric correction is typically unnecessary. When the training and testing data originate from different locations and times, a simple dark subtraction method is usually sufficient [54]. However, we still believe that using atmospherically corrected data could enhance the model's performance. Atmospheric correction removes the effects of atmosphere on the signal received by the sensor. Atmospheric effects are caused by water drops, aerosols, and gas molecules, which could scatter, refract or absorb the sunlight. When such effects are removed, the problem of simulating spectral bands simply boils down to injecting spectral features into each land cover type, despite the difficulty in classifying numerous ground objects accurately. We expect less error with atmospherically corrected data. However, this statement remains to be substantiated with further investigation. The biggest constraint is the fact that not every band can be inferred from other bands, as observed in the case of water. For example, short-wave near-infrared (SWIR) bands can pass through haze and fire smoke while visible bands can not. In this scenario, the SWIR band cannot be generated from the RGB bands accurately.

5.2.2. Discussion on Transferring to a Different Dataset

We tested the model trained on Sentinel-2 data on the Landsat-8 dataset, but it did not perform well due to the distinct characteristics of the two types of imagery. This problem cannot be solved without integrating professional knowledge in sensor and signal processing. However, a harmonization workflow has been proposed to enable the simultaneous use of Landsat and Sentinel-2 for land cover monitoring [1]. This workflow involves radiometric and geometric adjustments to provide a consistent surface reflectance record for time-series applications. First, the data is cloud-masked and atmospherically corrected using a radiative transfer algorithm. Then, the data is normalized to a common nadir view geometry using bi-directional reflectance distribution function (BRDF) estimation. Finally, the spectral bandpass adjustment is applied to Landsat-8 Operational Land Imager (OLI),

and both Sentinel-2 and Landsat-8 datasets are gridded to a common resolution, projection, and spatial extent. This workflow sheds light on using a model trained on different data for *NIR* band generation, and it could be integrated into the design of sensors that do not capture *NIR* bands. In the future, it would be worthwhile to investigate the performance of this model in the context of drone imagery. If applicable, this approach could greatly expand the applications of old data.

6. Conclusions

In this paper, we propose an innovative application of conditional generative adversarial network in multispectral remote sensing. We investigate the plausibility of simulating Sentinel-2 *NIR* band TOA reflectance from RGB bands using a cGAN. Furthermore, we replace the traditional loss function with a robust loss function and a structural similarity loss function to further improve the results. We also extend the scale of the experiment data and the scope of analysis compared with our pilot study, which was intended to examine the plausibility of this application.

We tested the proposed framework on SEN12MS multispectral datasets and achieved reasonable results, demonstrating the potential of applying this framework in various tasks, such as supplementing low-cost sensors with additional bands, recovering corrupted band information, and deriving supplemental indices for landcover classification. Although the method demonstrated reasonable results, we believe that additional knowledge is essential for more accurate results. Specifically, knowledge in the fields of signal processing, sensor calibration, and atmospheric modeling could contribute to a more robust model. In the meantime, the limitations of the method should be highlighted. In space-borne imaging, atmospheric effects play a crucial role in characterizing ground objects. These effects can result in distinct representations of ground objects by the same spectral bands, leading to wrong outputs. More importantly, the black box model might be able to imitate the human visual system, but whether it can accurately model physical phenomenon remains to be answered. Nevertheless, this framework could have greater potential in airborne datasets where the atmospheric effect is not as significant as on a spaceborne sensor. In addition, the possibility of combining ground object measurement and imagery captured by a spaceborne sensor can help to calibrate and fine-tune the model for higher accuracy. It could be applied to supplement airborne datasets for better characterisation of ground objects. In the future, we will seek to test the method on a multispectral airborne dataset to derive a plausible method to supplement a visible camera with additional artificial spectral bands to expand the usage of history data. We will also further investigate the performance of the generated *NIR* band in large-scale vegetation monitoring and supervised landcover classification.

Author Contributions: Conceptualization, J.T. and X.Y.; methodology, X.Y. and J.T.; validation, X.Y. and J.T.; formal analysis, X.Y. and J.T.; investigation, X.Y.; resources, P.R.; data curation, X.Y.; writing—original draft preparation, X.Y.; writing—review and editing, J.T. and P.R.; visualization, X.Y.; supervision, J.T. and P.R.; project administration, X.Y., J.T. and P.R.; funding acquisition, P.R. All authors have read and agreed to the published version of the manuscript.

Funding: We gratefully acknowledge the funding from the German Academic Exchange Service (DAAD), within the scholarship programme DLR/DAAD Research Fellowships—Doctoral Studies in Germany, under the personal reference number of 91704516.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The SEN12MS dataset can be downloaded at <https://github.com/schmitt-muc/SEN12MS>, accessed on 10 April 2023.

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

Abbreviations

The following abbreviations are used in this manuscript:

SLR	Single-lens reflex
GAN	Generative adversarial network
cGAN	Conditional generative adversarial network
DSM	Digital surface model
TOA	Top of atmosphere
NIR	Near-infrared
RGB	Red, green and blue
NDVI	Normalized difference vegetation index
NDWI	Normalized difference water index

References

- Claverie, M.; Ju, J.; Masek, J.G.; Dungan, J.L.; Vermote, E.F.; Roger, J.C.; Skakun, S.V.; Justice, C. The Harmonized Landsat and Sentinel-2 surface reflectance dataset. *Remote Sens. Environ.* **2018**, *219*, 145–161. [\[CrossRef\]](#)
- S. Clerc and MPC Team. S2 MPC L1C Data Quality Report. 2022. p. 31. Available online: https://sentinel.esa.int/documents/247904/685211/Sentinel-2_L1C_Data_Quality_Report (accessed on 9 April 2023).
- Rabatel, G.; Gorretta, N.; Labbé, S. Getting simultaneous red and near-infrared band data from a single digital camera for plant monitoring applications: Theoretical and practical study. *Biosyst. Eng.* **2014**, *117*, 2–14. [\[CrossRef\]](#)
- Berra, E.F.; Gaulton, R.; Barr, S. Commercial off-the-shelf digital cameras on unmanned aerial vehicles for multitemporal monitoring of vegetation reflectance and NDVI. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 4878–4886. [\[CrossRef\]](#)
- Dare, P. Small format digital sensors for aerial imaging applications. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2008**, *XXXVII Pt B1*, 533–538
- Brown, M.; Süssstrunk, S. Multi-spectral SIFT for scene category recognition. In Proceedings of the CVPR 2011, Colorado Springs, CO, USA, 20–25 June 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 177–184.
- McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [\[CrossRef\]](#)
- Topouzelis, K.; Papakonstantinou, A.; Garaba, S.P. Detection of floating plastics from satellite and unmanned aerial systems (Plastic Litter Project 2018). *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *79*, 175–183. [\[CrossRef\]](#)
- Topouzelis, K.; Papageorgiou, D.; Karagaitanakis, A.; Papakonstantinou, A.; Arias Ballesteros, M. Remote sensing of sea surface artificial floating plastic targets with Sentinel-2 and unmanned aerial systems (plastic litter project 2019). *Remote Sens.* **2020**, *12*, 2013. [\[CrossRef\]](#)
- Jain, R.; Sharma, R.U. Airborne hyperspectral data for mineral mapping in Southeastern Rajasthan, India. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *81*, 137–145. [\[CrossRef\]](#)
- Sankaran, S.; Ehsani, R. Visible-near-infrared spectroscopy based citrus greening detection: Evaluation of spectral feature extraction techniques. *Crop Prot.* **2011**, *30*, 1508–1513. [\[CrossRef\]](#)
- Sankaran, S.; Maja, J.M.; Buchanon, S.; Ehsani, R. Huanglongbing (citrus greening) detection using visible, near infrared and thermal imaging techniques. *Sensors* **2013**, *13*, 2117–2130. [\[CrossRef\]](#)
- Kaufman, Y.J.; Sendra, C. Algorithm for automatic atmospheric corrections to visible and near-IR satellite imagery. *Int. J. Remote Sens.* **1988**, *9*, 1357–1381. [\[CrossRef\]](#)
- Fu, Y.; Zhang, T.; Zheng, Y.; Zhang, D.; Huang, H. Joint camera spectral sensitivity selection and hyperspectral image recovery. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 788–804.
- Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-Image Translation with Conditional Adversarial Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2017, Honolulu, HI, USA, 21–26 July 2017; pp. 1125–1134. [\[CrossRef\]](#)
- Barron, J.T. A general and adaptive robust loss function. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 4331–4339.
- Yuan, X.; Tian, J.; Reinartz, P. Generating artificial near-infrared spectral band from rgb image using conditional generative adversarial network. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *3*, 279–285. [\[CrossRef\]](#)
- Sajjadi, M.S.; Scholkopf, B.; Hirsch, M. Enhancenet: Single image super-resolution through automated texture synthesis. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 4491–4500.
- Jiang, K.; Wang, Z.; Yi, P.; Wang, G.; Lu, T.; Jiang, J. Edge-enhanced GAN for remote sensing image superresolution. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 5799–5812. [\[CrossRef\]](#)
- Enomoto, K.; Sakurada, K.; Wang, W.; Fukui, H.; Matsuoka, M.; Nakamura, R.; Kawaguchi, N. Filmy cloud removal on satellite imagery with multispectral conditional generative adversarial nets. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 48–56.

21. Grohnfeldt, C.; Schmitt, M.; Zhu, X. A conditional generative adversarial network to fuse sar and multispectral optical data for cloud removal from sentinel-2 images. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1726–1729.
22. Ghamisi, P.; Yokoya, N. Img2dsm: Height simulation from single imagery using conditional generative adversarial net. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 794–798. [[CrossRef](#)]
23. Mou, L.; Zhu, X.X. IM2HEIGHT: Height estimation from single monocular imagery via fully residual convolutional-deconvolutional network. *arXiv* **2018**, arXiv:1802.10249.
24. Bittner, K.; Körner, M.; Fraundorfer, F.; Reinartz, P. Multi-Task cGAN for Simultaneous Spaceborne DSM Refinement and Roof-Type Classification. *Remote Sens.* **2019**, *11*, 1262. [[CrossRef](#)]
25. Liu, X.; Wang, Y.; Liu, Q. PSGAN: A generative adversarial network for remote sensing image pan-sharpening. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 873–877.
26. Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z.; et al. Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4681–4690.
27. Gong, M.; Niu, X.; Zhang, P.; Li, Z. Generative adversarial networks for change detection in multispectral imagery. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 2310–2314. [[CrossRef](#)]
28. Deng, L.; Sun, J.; Chen, Y.; Lu, H.; Duan, F.; Zhu, L.; Fan, T. M2H-Net: A reconstruction method for hyperspectral remotely sensed imagery. *ISPRS J. Photogramm. Remote Sens.* **2021**, *173*, 323–348. [[CrossRef](#)]
29. Zhao, J.; Kumar, A.; Banoth, B.N.; Marathi, B.; Rajalakshmi, P.; Rewald, B.; Ninomiya, S.; Guo, W. Deep-Learning-Based Multispectral Image Reconstruction from Single Natural Color RGB Image—Enhancing UAV-Based Phenotyping. *Remote Sens.* **2022**, *14*, 1272. [[CrossRef](#)]
30. Zhong, Y.; Hu, X.; Luo, C.; Wang, X.; Zhao, J.; Zhang, L. WHU-Hi: UAV-borne hyperspectral with high spatial resolution (H₂) benchmark datasets and classifier for precise crop identification based on deep convolutional neural network with CRF. *Remote Sens. Environ.* **2020**, *250*, 112012. [[CrossRef](#)]
31. Shi, Z.; Chen, C.; Xiong, Z.; Liu, D.; Wu, F. Hscnn+: Advanced cnn-based hyperspectral recovery from rgb images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 939–947.
32. Alvarez-Gila, A.; Van De Weijer, J.; Garrote, E. Adversarial networks for spatial context-aware spectral image reconstruction from rgb. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Venice, Italy, 22–29 October 2017; pp. 480–490.
33. Liu, P.; Zhao, H. Adversarial networks for scale feature-attention spectral image reconstruction from a single RGB. *Sensors* **2020**, *20*, 2426. [[CrossRef](#)]
34. Arad, B.; Ben-Shahar, O. Sparse recovery of hyperspectral signal from natural RGB images. In Proceedings of the Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; Proceedings, Part VII 14. Springer: Berlin/Heidelberg, Germany, 2016; pp. 19–34.
35. Zhang, J.; Su, R.; Fu, Q.; Ren, W.; Heide, F.; Nie, Y. A survey on computational spectral reconstruction methods from RGB to hyperspectral imaging. *Sci. Rep.* **2022**, *12*, 11905. [[CrossRef](#)]
36. Szeliski, R. Image alignment and stitching: A tutorial. *Found. Trends® Comput. Graph. Vis.* **2007**, *2*, 1–104. [[CrossRef](#)]
37. Koshelev, I.; Savinov, M.; Menshchikov, A.; Somov, A. Drone-Aided Detection of Weeds: Transfer Learning for Embedded Image Processing. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *16*, 102–111. [[CrossRef](#)]
38. Sa, I.; Lim, J.Y.; Ahn, H.S.; MacDonald, B. deepNIR: Datasets for generating synthetic NIR images and improved fruit detection system using deep learning techniques. *Sensors* **2022**, *22*, 4721. [[CrossRef](#)] [[PubMed](#)]
39. Aslahishahri, M.; Stanley, K.G.; Duddu, H.; Shirliffe, S.; Vail, S.; Bett, K.; Pozniak, C.; Stavness, I. From RGB to NIR: Predicting of near-infrared reflectance from visible spectrum aerial images of crops. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021; pp. 1312–1322.
40. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; pp. 2672–2680.
41. Mirza, M.; Osindero, S. Conditional generative adversarial nets. *arXiv* **2014**, arXiv:1411.1784.
42. 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 Intervention, Munich, Germany, 5–9 October 2015; Springer: Cham, Switzerland, 2015; pp. 234–241.
43. Pathak, D.; Krahenbuhl, P.; Donahue, J.; Darrell, T.; Efros, A.A. Context encoders: Feature learning by inpainting. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2536–2544.
44. Black, M.J.; Anandan, P. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *Comput. Vis. Image Underst.* **1996**, *63*, 75–104. [[CrossRef](#)]
45. Geman, S.; McClure, D. Bayesian image analysis: An application to single photon emission tomography. In Proceedings of the American Statistical Association, Statistical Computing Section Las Vegas, Nevada, 5–8 August 1985; pp. 12–18.

46. Dennis, J.E., Jr; Welsch, R.E. Techniques for nonlinear least squares and robust regression. *Commun. Stat.-Simul. Comput.* **1978**, *7*, 345–359. [[CrossRef](#)]
47. Charbonnier, P.; Blanc-Feraud, L.; Aubert, G.; Barlaud, M. Two deterministic half-quadratic regularization algorithms for computed imaging. In Proceedings of the of 1st International Conference on Image Processing, Austin, TX, USA, 13–16 November 1994; IEEE: Piscataway, NJ, USA, 1994; Volume 2, pp. 168–172.
48. Sun, D.; Roth, S.; Black, M.J. Secrets of optical flow estimation and their principles. In Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; IEEE: Piscataway, NJ, USA, 2010; pp. 2432–2439.
49. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)] [[PubMed](#)]
50. Schmitt, M.; Hughes, L.H.; Qiu, C.; Zhu, X.X. SEN12MS—A Curated Dataset of Georeferenced Multi-Spectral Sentinel-1/2 Imagery for Deep Learning and Data Fusion. *arXiv* **2019**, arXiv:1906.07789.
51. European Space Agency. SENTINEL-2 User Handbook. 2015. Issue 1, Rev2. Available online: http://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook (accessed on 15 January 2020).
52. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [[CrossRef](#)]
53. Lin, C.; Wu, C.C.; Tsogt, K.; Ouyang, Y.C.; Chang, C.I. Effects of atmospheric correction and pansharpening on LULC classification accuracy using WorldView-2 imagery. *Inf. Process. Agric.* **2015**, *2*, 25–36. [[CrossRef](#)]
54. Song, C.; Woodcock, C.E.; Seto, K.C.; Lenney, M.P.; Macomber, S.A. Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sens. Environ.* **2001**, *75*, 230–244. [[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.