

Cross Domain Early Crop Mapping with Label Spaces Discrepancies using MultiCropGAN

Yiqun WANG¹, Hui HUANG¹, Radu STATE¹

¹ Services and Data Management Research Group (SEDAN), Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, Luxembourg - (yiqun.wang, hui.huang, radu.state)@uni.lu

Keywords: Remote Sensing, Early Crop Mapping, Cross Domain, Domain Adaptation, MultiCropGAN, Label Spaces Discrepancies.

Abstract

Mapping target crops before the harvest season for regions lacking crop-specific ground truth is critical for global food security. Utilizing multispectral remote sensing and domain adaptation methods, prior studies strive to produce precise crop maps in these regions (target domain) with the help of the crop-specific labelled remote sensing data from the source regions (source domain). However, existing approaches assume identical label spaces across those domains, a challenge often unmet in reality, necessitating a more adaptable solution. This paper introduces the Multiple Crop Mapping Generative Adversarial Neural Network (MultiCropGAN) model, comprising a generator, discriminator, and classifier. The generator transforms target domain data into the source domain, employing identity losses to retain the characteristics of the target data. The discriminator aims to distinguish them and shares the structure and weights with the classifier, which locates crops in the target domain using the generator's output. This model's novel capability lies in locating target crops within the target domain, overcoming differences in crop type label spaces between the target and source domains. In experiments, MultiCropGAN is benchmarked against various baseline methods. Notably, when facing differing label spaces, MultiCropGAN significantly outperforms other baseline methods. The Overall Accuracy is improved by about 10%.

1. Introduction

Mapping target crops in the early stages before the harvest season is crucial for a variety of agricultural applications such as agricultural planning, resource allocation, crop insurance, and risk management (Waldner et al., 2015, Singha and Swain, 2016). To locate the target crops, the primary method involves analyzing time-series multispectral images (MSI), while MSI provides detailed spectral information, essential for understanding vegetation's spectral characteristics, influenced by its structural composition, leaf biochemistry, and phenological stages. To generate precise crop cultivation maps from these images, various supervised Deep Learning (DL) methods, including Convolutional Neural Networks (CNN) (Wang et al., 2021a, Hamidi et al., 2021), Temporal Convolutional Neural Network (TempCNN) (Pelletier et al., 2019), and Long Short-Term Memory (LSTM) (He et al., 2019, Crisóstomo de Castro Filho et al., 2020), have been explored. Notably, the findings indicate that DL approaches outperform conventional techniques like Support Vector Machine (SVM) (Mathur and Foody, 2008), Decision Trees (DTs) (Pittman et al., 2010, Tariq et al., 2023), and Random Forest (RF) (Duro et al., 2012, Wang et al., 2023b). These methods leverage publicly available datasets such as the United States Department of Agriculture (USDA)'s Cropland Data Layer (CDL) (Boryan et al., 2011) as the ground truth (GT).

Unfortunately, the collection of GT for crop types is expensive. In instances where GT data is absent, prevailing approaches adopt the "direct transfer strategy". This strategy involves training a classifier using labelled data from different regions (source domain) and then applying this trained model to regions lacking GT (target domain) (Hao et al., 2020, Ge et al., 2021). However, the trained model performs badly due to differences in soil composition, climate conditions, and crop progress, leading to discrepancies between the distributions of source and target data,

commonly known as cross-domain or domain shift issue (Konuri et al., 2020).

To address domain shift, Domain Adversarial Neural Networks (DANN) (Ajakan et al., 2014) and its variants, such as Self-Training with Domain Adversarial Network (STDAN) (Kwak and Park, 2022), Phenology Alignment Network (PAN) (Wang et al., 2021b), and Deep Adaptation Crop Classification Network (DACCN) (Wang et al., 2023a), are employed. These models aim to extract invariant features from both target and source domain data, subsequently using these invariant features for crop mapping classification. Alternatively, Generative Adversarial Neural Networks (GAN) (Creswell et al., 2018), such as the Crop Generative Adversarial Network (CropGAN) (Wang et al., 2024), are utilized. CropGAN transforms time-series MSI data from the target domain to the source domain while preserving local structures. This transformation enables a pre-trained crop mapper classifier, using labelled data from the source domain, to accurately locate the target crop using the transformed target domain data.

However, these methodologies are predicated on the assumption that the crop-type label spaces between the target and source domains are similar. While effective in addressing missing labels and mitigating the adverse effects of domain shifts, these strategies are limited by the disparities in label spaces between the domains. When discrepancies in label spaces occur between these domains, the extractor or generator still endeavors to align the data distributions of both domains. Consequently, some data from one domain must correspond to the other domain data associated with the labels absent in the first domain, leading to misclassification by the crop classifier. To address this label space discrepancies issue, the Multiple Crop Mapping Generative Adversarial Neural Network (MultiCropGAN) is presented in this paper. Our contributions can be summarized as follows:

- Introduce the MultiCropGAN, comprising a generator, a

discriminator, and a classifier, as a solution to mitigate the domain shift issue with label space discrepancies encountered in early crop mapping tasks.

- Propose to incorporate identity losses into the generator's loss function to ensure that the generator refrains from making unnecessary alterations to the data, thereby preserving its essential characteristics.
- Conduct the experiments based on study areas encompassing the USA and Canada. A comparative analysis was undertaken, pitting MultiCropGAN against various baseline methods, including the CropGAN, STDAN, DACCN, TempCNN, and RF. MultiCropGAN demonstrates the highest classification metrics when handling divergent label spaces in the target and source domains.

2. Related Work

Within a confined area, specific crop types often exhibit minimal variability in the growth period, offering reliable priors for accurate sample inference. However, outside these labelled regions or domains, significant phenological disparities arise within the same crop type due to variations in environmental conditions. This presents a considerable challenge for cross-domain classification. Existing approaches address this challenge through two distinct perspectives:

From a feature perspective, previous works introduced DANN variant methods to map samples from diverse regions into a shared feature subspace, thereby reducing differences in deep features. For instance, (Kwak and Park, 2022) presents the STDAN, a novel unsupervised domain adaptation framework for crop type classification. Moreover, the DACCN (Wang et al., 2021b) and the PAN (Wang et al., 2023a), extended the loss function using the Maximum Mean Discrepancy (MMD) and the Multiple Kernel variant of Maximum Mean Discrepancy (MK-MMD), achieving improved accuracy compared to CNN and LSTM methods without domain adaptation. From a sample perspective, two distinct methods emerge. The first method involves fine-tuning pre-trained models by utilizing a few high-quality samples from the target domain. This process enables the adaptation of the original model to the new distribution. For instance, in (Tong et al., 2020), deep models were refined for nationwide land cover classification by incorporating pseudo-labels with high confidence. Similarly, in (Hamrouni et al., 2021), new samples from the target domains were annotated to adjust RF classifiers through active learning. However, this approach often requires labelling additional samples, making it impractical for extensive area research. On the other hand, the second method, exemplified by the CropGAN (Wang et al., 2024), involves employing a GAN model to learn a mapping function. This function transforms time-series MSI sample data from the target domain to the source domain while preserving local structures. This transformation allows a pre-trained crop mapper classifier, utilizing source domain labelled data, to accurately process the transformed data, thereby enabling high-accuracy crop mapping without labelled data in the target domain.

However, the approaches discussed earlier, regardless of their perspective, face limitations arising from differences in label spaces between the target and source domains. As far as our knowledge extends, we are the pioneering contributors to tackling label space discrepancies from a sample perspective. This

is achieved through the utilization of a GAN model combined with specially designed identity losses.

3. Problem Statement

Let \mathbf{X} denote the time series remote sensing input data and \mathbf{Y} denote the GT. Each sample \mathbf{x} can be expressed as a temporal form $[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_t]$, where \mathbf{x}_i represents input at time i . \mathbf{x}_i can be further expanded as $[\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{ib}]$, containing multi-spectral bands information from band 1 to band b . Each GT label, denoted as \mathbf{y} , is represented as a one-hot vector comprising four elements, which correspond to the categories of corn, soybean, spring wheat, and other crops. Let \mathbf{X}_t denote the time series input target data, \mathbf{X}_s denote the time series input source data, and \mathbf{Y}_s denote the GT for \mathbf{X}_s . Each sample \mathbf{x}_s has a corresponding \mathbf{y}_s .

Our objective is to identify target crops to get target labels \mathbf{Y}_t in the target domain during their early growth stages, utilizing labelled source domain data ($\mathbf{X}_s, \mathbf{Y}_s$) and unlabelled target domain data (\mathbf{X}_t). This approach addresses challenges related to the cross-domain issue and the label space discrepancies issue. The source and target domains correspond to the study areas detailed in Section 3.1, while the specific challenges are elaborated upon in Sections 3.2 and 3.3.

3.1 Study Areas

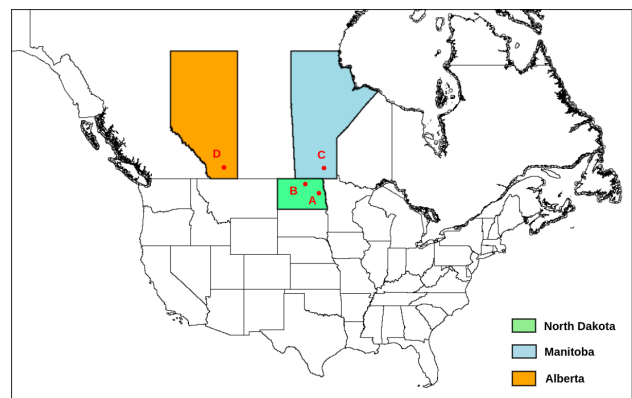


Figure 1. The Study Areas in the USA and Canada. The locations of our study areas are denoted by red dots.

The study areas are located in two countries: the USA and Canada. As depicted in Figure 1, Study Areas A (Traill County) and B (Cavalier County) are located in North Dakota, the USA, whereas Study Areas C and D are in Manitoba and Alberta, Canada, respectively. The geographic coordinates for Study Area C range from longitudes -97.97 to -96.59 and latitudes 49.22 to 49.59 . Study Area D is defined by longitudes -112.96 to -112.53 and latitudes 49.79 to 50.06 .

3.2 Cross Domain Issue

The Normalized Difference Vegetation Index (NDVI) calculated as a ratio between the red (Red) and near-infrared (NIR) values by

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

, is usually used to quantify vegetation greenness. Figure 2 displays the average NDVI value curves for the target crops and other crop types throughout their growth stages in different study areas. It is important to recognize that variations in

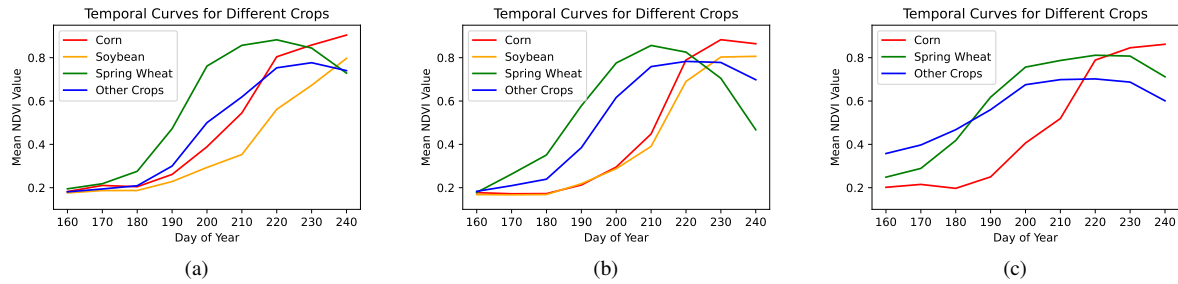


Figure 2. The Temporal NDVI Curves of Different Crops during Their Growing Period in Study Areas. (a) Study Area A and B. (b) Study Area C. (c) Study Area D.

S	T(K)	P(mm/h)	E(mm/h)	R(kJ/m ²)	EI (m)
A	287.37	2.04	-1.01	4794.21	287.60
B	284.85	1.95	-0.93	4368.04	470.35
C	285.61	2.12	-0.95	4376.90	237.58
D	285.90	0.68	-0.64	5079.23	896.98

Table 1. The Environmental Conditions of the Study Areas. "S" denotes study areas, "T" annual average temperature, "P" average hourly precipitation, "E" average hourly evaporation rate, "R" surface net solar radiation, and "EI" elevation.

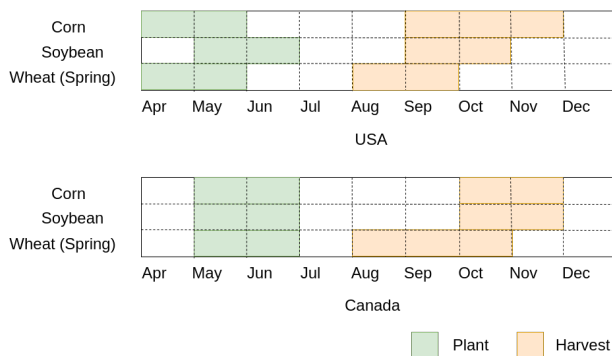


Figure 3. The Crop Calendar delineates the planting and harvesting schedules for target crops in the USA (above) and Canada (below).

NDVI curves for a specific crop across different regions can be attributed to different environmental conditions, as detailed in Table 1, and varying crop calendars, as illustrated in Figure 3. These factors contribute to distinct phenological characteristics in crops within these regions. Consequently, when a model trained on source domain data is directly applied to the target domain, it frequently underperforms, termed as the cross-domain issue.

3.3 Label Space Discrepancies Issue

Current cross-domain methods aim to align the data distribution between the source and target domains to mitigate the domain shift problem, assuming that the crop types between these domains are identical. However, in the context of real-world applications, these domains often display a variety of crop types. For instance, Table 2 illustrates the diverse crop types present in the study areas. The absence of certain crop types in each domain leads to a significant difference between the data distributions, which the current cross-domain methods cannot resolve, known as the label space discrepancies issue.

Crops	A	B	C	D
Corn	✓	✓	✓	✓
Spring wheat	✓	✓	✓	✓
Soybean	✓	✓	✓	-
Sunflower	✓	✓	✓	-
Barley	✓	✓	✓	✓
Canola	✓	-	✓	✓
Peas	✓	-	-	✓
Sugarbeets	-	✓	-	✓
Dry Beans	✓	✓	-	-
Pasture and Forages	-	-	✓	✓
Canary seed	-	-	✓	-
Oats	-	-	✓	-
Rye	-	-	✓	-
Flax seed	-	-	✓	-
Potatoes	-	-	-	✓
Lentils	-	-	-	✓

Table 2. Crop Types in Study Area A, B, C, and D. "✓" denotes the presence of a crop, while "-" indicates its absence.

4. Methodology

To address the cross-domain issue and label space discrepancies, the MultiCropGAN model is proposed. It transforms the target domain data into the source domain while preserving the characteristics of the target domain data related to crop types absent in the source domain. Figure 4 illustrates our proposed MultiCropGAN model, consisting of a generator, a discriminator, and a classifier. The primary goal of the generator is to transform data from the target domain into the source domain. Meanwhile, the discriminator's objective is to differentiate between the transformed target data and the original source data. The classifier and the discriminator are structurally identical and share weights, with the exception of the output layer. Trained with source data and labels, the classifier's purpose is to categorize crop types in the target domain using the transformed target data produced by the generator.

4.1 Data Preprocessing

The preprocessing, shown in Figure 5, aims at providing complete time-series remote sensing data by filling gaps between MSI images due to cloud cover, atmospheric interference, or sensor limitations. There are four dataset used in the preprocessing:

- The Sentinel-2 MSI images, widely recognized and employed in numerous agricultural applications within the scientific community (Wang et al., 2021a, Blickensdörfer et al., 2022), are utilized as remote sensing data. Sentinel-2 captures high-resolution MSI images up to 10 meters,

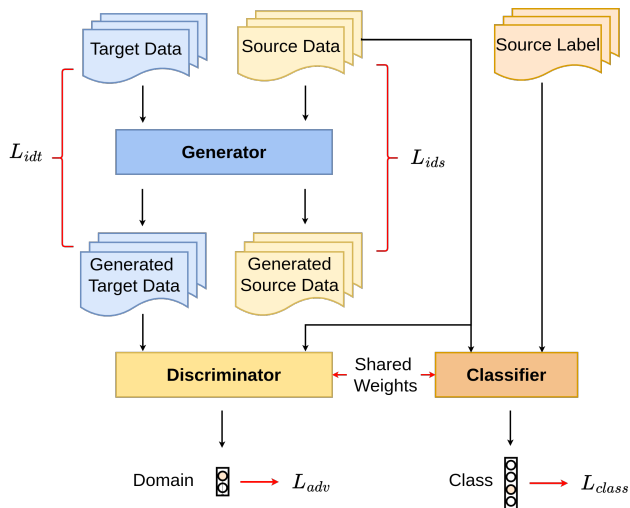


Figure 4. The MultiCropGAN Structure with Training Dataflow. It comprises three essential components: the generator, the discriminator, and the classifier.

with a 5-day revisit time, enabling frequent crop growth monitoring.

- The Dynamic World dataset (Brown et al., 2022) is utilized, providing detailed class probabilities and labels with a 10-meter resolution for nine distinct land categories, notably including cropland.
- The CDL (Boryan et al., 2011), a crop-specific land cover raster map accessible for the entire contiguous U.S. land area at a 30-meter resolution, is provided by the USDA. It serves as the training GT for the USA.
- Similarly, the Canada Agriculture and Agri-Food Canada (AAFC) Annual Crop Inventory (Fisette et al., 2013), with a 30-meter spatial resolution, is employed as the GT for Canada in the test process.

The procedure, shown in Figure, starts by acquiring MSI images with six bands, including B2 (Blue), B3 (Green), B4 (Red), B8 (NIR), B11 (Shortwave Infrared 1), and B12 (Shortwave Infrared 2), encompassing the entire designated study areas. These images originate from the Sentinel-2 Dataset and are captured at regular 10-day intervals as an image sequence. Our primary goal is to identify target crops with label space discrepancies and cross-domain problems during an early growth stage. Consequently, the time series remote sensing data should start after the planting period and end before the onset of crop harvest. As depicted in Figure 3, the earliest target crop harvesting season commences in August. As a result, nine image groups are compiled, spanning from May 1st to July 30th. Simultaneously, to eliminate non-agricultural lands in the MSI images, the process extracts the cropland mask during the crop growing season from the Dynamic World dataset and reprojects it to maintain a consistent 30-meter resolution. After excluding non-agricultural lands using the cropland mask extracted from the Dynamic World dataset, the MSI images in the nine image groups specifically focus on agricultural lands. Within each image group, Sentinel Hub’s cloud detector (Skakun et al., 2022) is utilized to implement cloud masks on the Sentinel-2 MSI images. This merging process retains the mean values of the cloud-free MSI images, resulting in a composite MSI image at

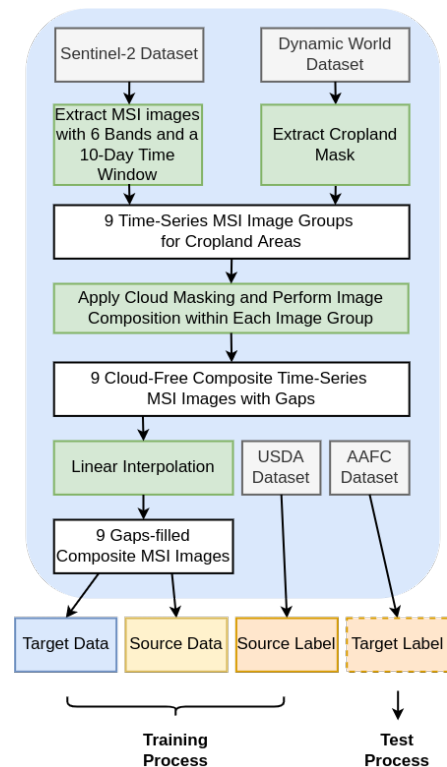


Figure 5. The Preprocessing.

a clear, 30-meter resolution. Consequently, a set of nine high-quality time-series MSI images is generated and characterized by their absence of clouds and gaps. These images encompass six bands, all with a spatial resolution of 30 meters.

4.2 Model Structure

In the MultiCropGAN framework, the generator operates as an autoencoder. The encoder is composed of three residual blocks, each featuring two one-dimensional convolutional (Conv1D) layers followed by a MaxPooling layer. The Conv1D layers are characterized by 32 filters with a kernel size of 3. Meanwhile, the MaxPooling layers have a pool size of 3 and a stride of 1. The encoder is followed by a flatten layer and a sequence of fully connected layers with dimensions 96, 64, 64, 64, and 96, respectively. The decoder, mirroring the encoder, employs Conv1DTranspose layers, which are specifically designed for transposing a Conv1D layer. This technique reverses the transformation performed by the corresponding Conv1D layers in the encoder, thus aiding in reconstructing the original input features from their encoded state. The discriminator and classifier adopt the same encoder structure as the generator. This is followed by two fully connected layers, featuring dimensions of 32 and 16, respectively. The discriminator’s output layer is a fully connected layer with a dimension of 2, designed to distinguish between generated target data and original source data. Conversely, the classifier’s output layer is a fully connected layer with a dimension of 4, tasked with identifying target crops (such as corn, soybean, and spring wheat) and other crop types.

4.3 Losses

4.3.1 Adversarial Loss The adversarial loss function is defined by

	MCGAN (Ours)		CropGAN		STDAN		DACCN		TempCNN		RF	
	OA (%)	F1 (%)	OA	F1	OA	F1	OA	F1	OA	F1	OA	F1
S1	80.90	80.91	74.47	73.89	70.52	69.66	72.52	72.04	71.44	70.37	67.93	66.90
S2	81.86	81.87	70.90	71.29	72.68	71.00	68.06	67.05	71.06	70.19	68.23	67.11
S3	81.64	81.54	72.81	72.02	68.98	67.51	67.82	65.02	66.76	63.37	66.07	64.71
S4	82.96	82.93	70.48	69.28	73.68	72.76	71.68	71.03	71.80	70.39	66.95	65.59
S5	81.08	81.04	73.39	73.87	64.88	61.77	73.86	72.27	70.90	68.78	67.35	66.09
Avg	81.69	81.66	<u>72.41</u>	<u>72.07</u>	70.15	68.54	70.79	69.48	70.39	68.62	67.31	66.08

Table 3. Experiment Metrics for the First Experiment: Best metrics are indicated in bold, while the best metrics of the SOTA methods are underlined.

	MCGAN (Ours)		CropGAN		STDAN		DACCN		TempCNN		RF	
	OA (%)	F1 (%)	OA	F1	OA	F1	OA	F1	OA	F1	OA	F1
S1	79.96	80.27	70.00	69.25	64.19	63.06	67.24	68.52	71.68	72.35	76.78	76.93
S2	80.89	80.96	68.78	68.75	71.87	73.01	70.01	70.96	73.88	74.41	76.93	77.00
S3	79.53	80.55	70.94	72.97	69.40	70.76	66.89	65.25	72.05	73.26	<u>77.27</u>	<u>77.25</u>
S4	77.39	79.47	72.82	73.43	69.63	68.69	66.94	67.89	74.76	75.52	76.99	77.13
S5	80.68	81.45	72.03	72.37	68.65	69.92	67.07	67.07	75.27	76.07	77.09	77.12
Avg	79.69	80.54	70.91	71.35	68.75	69.09	67.63	67.94	73.53	74.32	<u>77.01</u>	<u>77.09</u>

Table 4. Experiment Metrics for the Second Experiment: Best metrics are indicated in bold, while the best metrics of the SOTA methods are underlined.

$$\mathcal{L}_{adv}(G, D, \mathbf{X}_t, \mathbf{X}_s) = \mathbb{E}_{\mathbf{x}_s \sim p_{data}(\mathbf{x}_s)} [\log D(\mathbf{x}_s)] + \mathbb{E}_{\mathbf{x}_t \sim p_{data}(\mathbf{x}_t)} [1 - \log D(G(\mathbf{x}_t))] \quad (2)$$

, which aims to drive the generator to produce transformed data within the target domain that closely resembles real data from the source domain. The discriminator’s objective is to distinguish the generated data as fake data, while the generator’s goal is to craft realistic data to deceive the discriminator.

4.3.2 Identity Loss The identity losses are defined by

$$\mathcal{L}_{idt}(G, \mathbf{X}_t) = \mathbb{E}_{\mathbf{x}_t \sim p_{data}(\mathbf{x}_t)} [\|G(\mathbf{x}_t) - \mathbf{x}_t\|_1] \quad (3)$$

$$\mathcal{L}_{ids}(G, \mathbf{X}_s) = \mathbb{E}_{\mathbf{x}_s \sim p_{data}(\mathbf{x}_s)} [\|G(\mathbf{x}_s) - \mathbf{x}_s\|_1] \quad (4)$$

, which encourage the generator to preserve the identity of the input data. It computes the difference between the generator output and the input data. The objective of minimizing this loss is to ensure that the generator does not make unnecessary alterations to the data and maintains its essential characteristics.

4.3.3 Class Loss The class loss, denoted as the multi-class cross-entropy loss, is defined by the equation:

$$\mathcal{L}_{class}(C, \mathbf{X}_s, \mathbf{Y}_s) = - \sum \mathbf{y}_s \cdot \log(C(\mathbf{x}_s)) \quad (5)$$

This loss function encourages the classifier to extract features that encompass class-related information. These features are shared with the discriminator since the classifier and the discriminator have shared weights. The discriminator utilizes these features to discern whether the data originates from the original source domain or not.

4.3.4 Total Loss The total loss is defined by

$$\mathcal{L}_{total} = \alpha \mathcal{L}_{adv}(G, D, \mathbf{X}_t, \mathbf{X}_s) + \beta (\mathcal{L}_{idt}(G, \mathbf{X}_t) + \mathcal{L}_{ids}(G, \mathbf{X}_s)) + \sigma \mathcal{L}_{class}(C, \mathbf{X}_s, \mathbf{Y}_s) \quad (6)$$

, where α, β, σ are the weight parameters. The total loss can be expressed as a minimax function:

$$\min_{G, C} \max_D \mathcal{L}_{total}(G, D, C, \mathbf{X}_t, \mathbf{X}_s, \mathbf{Y}_s) \quad (7)$$

, where the generator seeks to minimize the loss while the discriminator aims to maximize it.

5. Experiments

In this section, our proposed MultiCropGAN is compared with other state-of-the-art (SOTA) algorithms, including CropGAN, STDAN, DACCN, TempCNN, and RF, on two experiments of cross-domain time series early crop classification with discrepancies in label spaces. The classification encompasses four labels, comprising three types of target crops (corn, soybean, and spring wheat) and a category for other crop types.

5.1 Experimental Setup

Two experiments are set up in this paper:

- The first experiment employs study areas A and B as the source domain, while study area C is designated as the target domain. Despite the alignment of target crop types with those in the source domain, variations exist in the types of other crops.
- In the second experiment, study areas A and B function as the source domain, and study area D serves as the target domain, specifically lacking soybean. The variations in the types of other crops are retained.

In the experiments, Our method and SOTA methods are categorized into two groups. MultiCropGAN, CropGAN, STDAN, and DACCN fall under the cross-domain methods category, while TempCNN and RF are classified as direct methods.

5.2 Training Setup

The experiment utilizes data sampled from 2019. From the source domain, 5,000 data points are randomly selected for

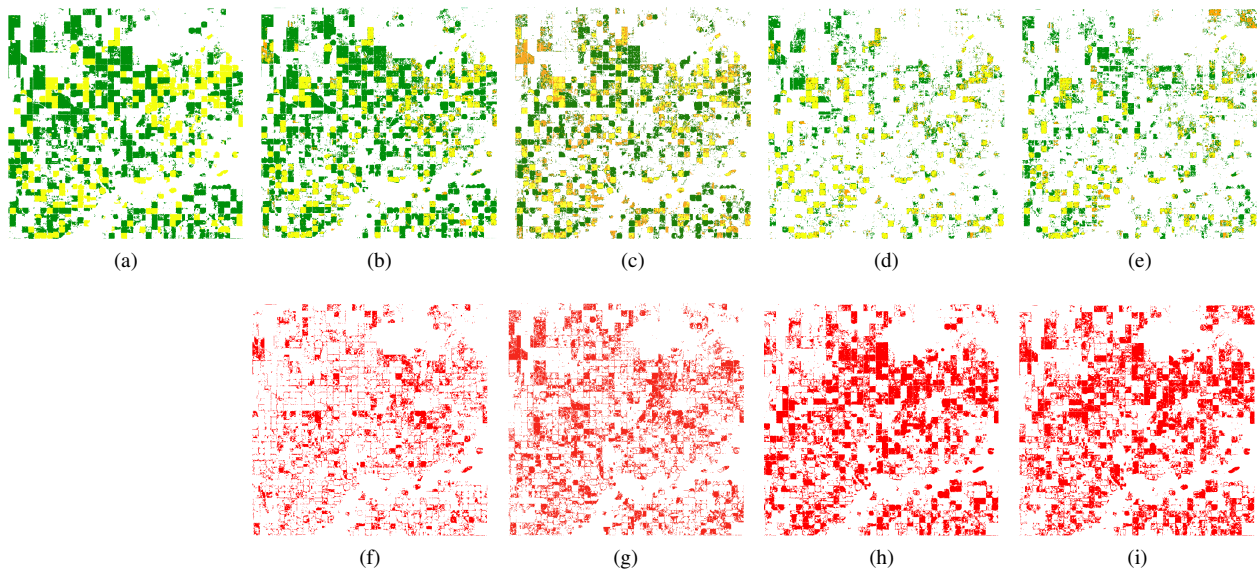


Figure 6. Visualization of the Second Experiment Results Employing Cross-Domain Deep Learning Methods. In this visualization, yellow denotes corn, orange signifies soybean, green indicates spring wheat, and white represents other crops. (a) displays the GT for crop types. The crop mapping results are depicted in (b) for MultiCropGAN, (c) for CropGAN, (d) for STDAN, and (e) for DACCN. The corresponding error images are illustrated in panels (f) through (i) for MultiCropGAN, CropGAN, STDAN, and DACCN, respectively. Red highlights the misclassified pixels.

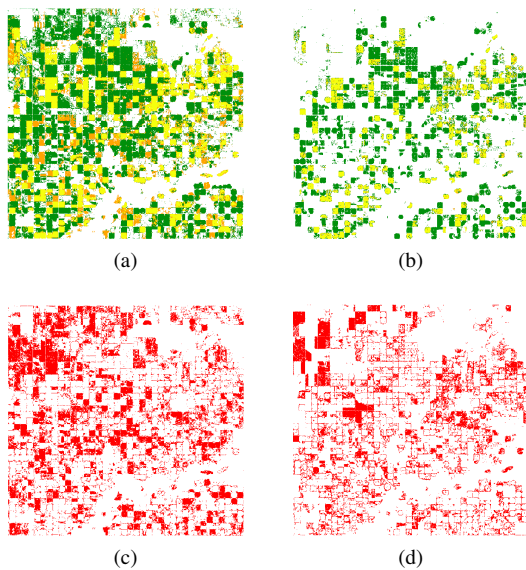


Figure 7. Visualization of the Second Experiment Results Employing the TempCNN and RF. The crop mapping results are depicted in (a) for TempCNN, and (b) for RF. The corresponding error images are illustrated in (c) for TempCNN, and (d) for RF.

each of nine existing crop types, as detailed in Table 2, amounting to a total of 45,000 data points. These are balanced and accompanied by source domain labels. In contrast, for each target domain, 150,000 data points were randomly sampled, lacking label information and resulting in unbalanced data. For the training and evaluation of models, the source data was distributed as follows: 80% was dedicated to training, and the remaining 20% was set aside for evaluation, specifically for the application of early stopping criteria. Notably, all target data were utilized in the training process of the cross-domain methods and testing process of all methods. All methods are

trained solely on the training sets to acquire well-trained models for these two experiments. Every compared method was repeatedly trained on each subset from scratch 5 times with the same training configuration. Our MultiCropGAN was trained with $\alpha = 1, \beta = 20, \sigma = 1$ in the loss function. The RF classifier was trained using specific parameters: `tree_num = 50` and `leaf_size = 15`. The optimizer employed in all deep learning methods underwent a substitution, being replaced by the Adam optimizer, initialized with a learning rate of 0.0005 and configured with $\beta = (0.9, 0.998)$. The RF classifier underwent training using the sklearn library, while all deep models were constructed using Tensorflow. Training persisted until the completion of 500 epochs or upon convergence, as determined by an early stopping criterion set at 50 epochs.

5.3 Experimental Results

Two metrics, the Overall Accuracy (OA) and Weighted F1 Score, are utilized to assess the performance of the proposed Multi-CropGAN model. Tables 3 and 4 present the OA and F1 score results of our method in comparison to the SOTA. Across two experiments, our method demonstrates the highest average OA and F1 score.

In the first experiment, our approach achieved an average OA of 81.69% and an average F1 score of 81.66%. This marked a notable improvement, with a +9.28% increase in OA and a +9.59% boost in F1 score when compared to the CropGAN model's performance, which had an OA of 72.41% and an F1 score of 72.07%, securing the second position in the rankings. Among the cross-domain methods, DACCN claimed the third position, surpassing STDAN in performance. Within the category of direct methods, TempCNN exhibited slightly better results than STDAN but lagged behind DACCN. Notably, RF demonstrated the weakest performance among all the methods, with an OA of 67.31% and an F1 score of 66.08%.

In the second experiment, our approach achieved an OA of 79.69% and an F1 score of 80.54%. In contrast, the RF model

secured the second position, attaining an OA of 77.01% and an F1 score of 77.09%. Noteworthy is the performance of the direct method TempCNN, which secured the third position with an OA of 73.53% and an F1 score of 74.32%, outperforming the other three cross-domain methods. Among these cross-domain methods, CropGAN demonstrated superior performance with an OA of 70.91% and an F1 score of 71.35% compared to STDAN, which achieved an OA of 68.75% and an F1 score of 69.09%. Furthermore, STDAN outperformed DACCN, which attained an OA of 67.63% and an F1 score of 67.94%.

The second experiment presents a more intricate scenario concerning discrepancies within the label space, in contrast to the first experiment. This complexity is characterized not only by a diversity of other crop types but also by the conspicuous absence of soybeans among the target crop categories. As depicted in Figure 6(a), it is evident that soybeans (depicted in yellow) are not present in the target domain, although they exist in the source domain. Figures 6 and 7 provide visual representations of the results obtained through both our methodology and SOTA approaches, specifically tailored to address the challenges of the second experiment. In the results, it becomes apparent that CropGAN and TempCNN misclassify parts of corn, spring wheat, or other crop types as soybean pixels. Moreover, STDAN, DACCN, and RF exhibit trends of misclassification of corn and spring wheat as other crops.

6. Discussion

When there are moderate differences in label spaces across various cross-domain methods, such as those observed in the first experiment, cross-domain deep learning methods exhibit superior performance over TempCNN and RF. Notably, when DACCN incorporates the MMD loss to extract the invariant features more strictly, it demonstrates higher performance than STDAN. CropGAN outperforms both STDAN and DACCN, possibly due to its utilization of the identity loss calculated from the source domain data. Particularly noteworthy is the exceptional performance of MultiCropGAN, surpassing all the other methods, thereby establishing itself as the most effective method in this context.

However, in scenarios characterized by substantial disparities among target crop labels, as evident in the second experiment, TempCNN and RF demonstrated superior performance compared to several cross-domain methods (such as CropGAN, STDAN, and DACCN). This could be attributed to the attempts made by these cross-domain methods to align data distributions between the target and source domains, despite the inherent differences caused by the absence of soybeans in the target domain. Despite DACCN's more rigorous attempts to extract invariant features by employing MMD loss compared to STDAN, it exhibited less effective performance. Notably, CropGAN consistently demonstrated superior performance compared to both STDAN and DACCN. As expected, MultiCropGAN maintained its position as the top-performing approach, surpassing even TempCNN and RF. One of the key contributing factors to MultiCropGAN's success is its use of two identity losses, which ensure that the generator retains crucial data characteristics without unnecessary alterations.

7. Conclusions

This paper introduces MultiCropGAN, an innovative generative adversarial neural network designed to tackle the issue of do-

main shift resulting from diverse regions and variations in label spaces during early crop mapping, leveraging multi-temporal multispectral input data. The MultiCropGAN model comprises three key components: a generator, a discriminator, and a classifier. Additionally, we introduce identity losses for both target and source domain data to ensure that the generator maintains essential data characteristics throughout the transformation process. The classifier is trained using source domain data and their respective labels, sharing model weights with the discriminator to enable the latter to leverage class-related features for distinguishing between generated target domain data and original source domain data.

Finally, our MultiCropGAN is evaluated against several SOTA methods in scenarios marked by variations in crop label spaces. Experiments conducted across the USA and Canada demonstrated that MultiCropGAN outperforms various SOTA methods, including CropGAN, STDAN, DACCN, TempCNN, and RF. The comparative analysis shows MultiCropGAN achieving the highest classification metrics, particularly effective in handling divergent label spaces in target and source domains. The results show that MultiCropGAN notably enhances classification outcomes for the target domain without utilizing any label information specific to the target domain.

8. Acknowledgements

I would like to express my deepest gratitude to my advisors, Hui HUANG and Radu STATE, for their unwavering support, guidance, and mentorship throughout the duration of this research project. Their expertise, encouragement, and invaluable insights have been instrumental in shaping the course of my academic journey.

I would also like to extend my heartfelt appreciation to my family for their endless love, encouragement, and sacrifices. Their unwavering belief in my abilities has been a constant source of motivation, and I am profoundly grateful for their unyielding support.

Additionally, I want to thank all my colleagues, friends, and fellow researchers who provided valuable assistance, feedback, and inspiration during this research endeavor. Your contributions have enriched the quality of this work and have been immensely valuable.

References

- Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., Marchand, M., 2014. Domain-adversarial neural networks. *arXiv preprint arXiv:1412.4446*.
- Blickensdörfer, L., Schwieder, M., Pflugmacher, D., Nendel, C., Erasmí, S., Hostert, P., 2022. Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany. *Remote sensing of environment*, 269, 112831.
- Boryan, C., Yang, Z., Mueller, R., Craig, M., 2011. Monitoring US agriculture: the US department of agriculture, national agricultural statistics service, cropland data layer program. *Geocarto International*, 26(5), 341–358.
- Brown, C. F., Brumby, S. P., Guzder-Williams, B., Birch, T., Hyde, S. B., Mazzariello, J., Czerwinski, W., Pasquarella, V. J.,

- Haertel, R., Ilyushchenko, S. et al., 2022. Dynamic World, Near real-time global 10 m land use land cover mapping. *Scientific Data*, 9(1), 251.
- Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sen-gupta, B., Bharath, A. A., 2018. Generative adversarial networks: An overview. *IEEE signal processing magazine*, 35(1), 53–65.
- Crisóstomo de Castro Filho, H., Abílio de Carvalho Júnior, O., Ferreira de Carvalho, O. L., Pozzobon de Bem, P., dos Santos de Moura, R., Olino de Albuquerque, A., Rosa Silva, C., Guimaraes Ferreira, P. H., Fontes Guimarães, R., Trancoso Gomes, R. A., 2020. Rice crop detection using LSTM, Bi-LSTM, and machine learning models from Sentinel-1 time series. *Remote Sensing*, 12(16), 2655.
- Duro, D. C., Franklin, S. E., Dubé, M. G., 2012. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote sensing of environment*, 118, 259–272.
- Fisette, T., Rollin, P., Aly, Z., Campbell, L., Daneshfar, B., Filyer, P., Smith, A., Davidson, A., Shang, J., Jarvis, I., 2013. Aaaf annual crop inventory. *2013 Second International Conference on Agro-Geoinformatics (Agro-Geoinformatics)*, IEEE, 270–274.
- Ge, S., Zhang, J., Pan, Y., Yang, Z., Zhu, S., 2021. Transferable deep learning model based on the phenological matching principle for mapping crop extent. *International Journal of Applied Earth Observation and Geoinformation*, 102, 102451.
- Hamidi, M., Safari, A., Homayouni, S., 2021. An auto-encoder based classifier for crop mapping from multitemporal multispectral imagery. *International Journal of Remote Sensing*, 42(3), 986–1016.
- Hamrouni, Y., Paillassa, E., Chéret, V., Monteil, C., Sheeren, D., 2021. From local to global: A transfer learning-based approach for mapping poplar plantations at national scale using Sentinel-2. *ISPRS Journal of Photogrammetry and Remote Sensing*, 171, 76–100.
- Hao, P., Di, L., Zhang, C., Guo, L., 2020. Transfer Learning for Crop classification with Cropland Data Layer data (CDL) as training samples. *Science of The Total Environment*, 733, 138869.
- He, T., Xie, C., Liu, Q., Guan, S., Liu, G., 2019. Evaluation and comparison of random forest and A-LSTM networks for large-scale winter wheat identification. *Remote Sensing*, 11(14), 1665.
- Konduri, V. S., Kumar, J., Hargrove, W. W., Hoffman, F. M., Ganguly, A. R., 2020. Mapping crops within the growing season across the United States. *Remote Sensing of Environment*, 251, 112048.
- Kwak, G.-H., Park, N.-W., 2022. Unsupervised domain adaptation with adversarial self-training for crop classification using remote sensing images. *Remote Sensing*, 14(18), 4639.
- Mathur, A., Foody, G. M., 2008. Crop classification by support vector machine with intelligently selected training data for an operational application. *International Journal of Remote Sensing*, 29(8), 2227–2240.
- Pelletier, C., Webb, G. I., Petitjean, F., 2019. Temporal convolutional neural network for the classification of satellite image time series. *Remote Sensing*, 11(5), 523.
- Pittman, K., Hansen, M. C., Becker-Reshef, I., Potapov, P. V., Justice, C. O., 2010. Estimating global cropland extent with multi-year MODIS data. *Remote Sensing*, 2(7), 1844–1863.
- Singha, C., Swain, K. C., 2016. Land suitability evaluation criteria for agricultural crop selection: A review. *Agricultural reviews*, 37(2), 125–132.
- Skakun, S., Wevers, J., Brockmann, C., Doxani, G., Aleksandrov, M., Batič, M., Frantz, D., Gascon, F., Gómez-Chova, L., Hagolle, O. et al., 2022. Cloud Mask Intercomparison eXercise (CMIX): An evaluation of cloud masking algorithms for Landsat 8 and Sentinel-2. *Remote Sensing of Environment*, 274, 112990.
- Tariq, A., Yan, J., Gagnon, A. S., Riaz Khan, M., Mumtaz, F., 2023. Mapping of cropland, cropping patterns and crop types by combining optical remote sensing images with decision tree classifier and random forest. *Geo-Spatial Information Science*, 26(3), 302–320.
- Tong, X.-Y., Xia, G.-S., Lu, Q., Shen, H., Li, S., You, S., Zhang, L., 2020. Land-cover classification with high-resolution remote sensing images using transferable deep models. *Remote Sensing of Environment*, 237, 111322.
- Waldner, F., Fritz, S., Di Gregorio, A., Defourny, P., 2015. Mapping priorities to focus cropland mapping activities: Fitness assessment of existing global, regional and national cropland maps. *Remote Sensing*, 7(6), 7959–7986.
- Wang, Y., Feng, L., Zhang, Z., Tian, F., 2023a. An unsupervised domain adaptation deep learning method for spatial and temporal transferable crop type mapping using Sentinel-2 imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 199, 102–117.
- Wang, Y., Huang, H., State, R., 2023b. Early Crop Mapping Using Dynamic Ecoregion Clustering: A USA-Wide Study. *Remote Sensing*, 15(20), 4962.
- Wang, Y., Huang, H., State, R., 2024. Cross Domain Early Crop Mapping using CropGAN and CNN Classifier. *arXiv preprint arXiv:2401.07398*.
- Wang, Y., Zhang, Z., Feng, L., Ma, Y., Du, Q., 2021a. A new attention-based CNN approach for crop mapping using time series Sentinel-2 images. *Computers and electronics in agriculture*, 184, 106090.
- Wang, Z., Zhang, H., He, W., Zhang, L., 2021b. Phenology alignment network: A novel framework for cross-regional time series crop classification. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, American Association for Artificial Intelligence, 2940–2949.