# using Generative Adversarial Networks

2024 12th International Conference on Agro-Geoinformatics (Agro-Geoinformatics) | 979-8-3503-8060-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/Agro-Geoinformatics262780.2024.10660730

Bosen Shao
*Center for Spatial Information Science and Systems*
*George Mason University*
Fairfax, VA, USA
bshao2@gmu.edu

Liping Di
*Center for Spatial Information Science and Systems*
*George Mason University*
Fairfax, VA, USA
ldi@gmu.edu

Hui Li
*Center for Spatial Information Science and Systems*
*George Mason University*
Fairfax, VA, USA
hli47@gmu.edu

Chen Zhang
*Center for Spatial Information Science and Systems*
*George Mason University*
Fairfax, VA, USA
czhang11@gmu.edu

Ziao Liu
*Center for Spatial Information Science and Systems*
*George Mason University*
Fairfax, VA, USA
zliu23@gmu.edu

Liying Guo
*Center for Spatial Information Science and Systems*
*George Mason University*
Fairfax, VA, USA
lguo2@gmu.edu

***Abstract— High-resolution earth surface imagery is significant for surface object identification and edge detection, supporting field-level object-oriented classification and classification refinement. The U.S. National Agriculture Imagery Program (NAIP) provides high-definition images covering the Conterminous United States during the agricultural growing season (June-August). Such an annual drone imagery product possesses spatial resolution as high as 0.6m, covering cloud-free red, green, blue, and near-infrared channels, acting as a premium data source to promote applications of field-level crop identification. However, the NAIP imageries are unavailable in other nations, leading to a challenge to global precise crop mapping.. This study explores a near-real high-resolution generation workflow to produce NAIP-like images with high spatial resolution by leveraging the capabilities of Generative Adversarial Networks (GANs). By employing the Sentinel-2 imagery as source data and NAIP data as destination data, we utilized the Pix2Pix GAN to produce high-resolution imagery with high-definition object context and edge like NAIP. We used pixel-to-pixel differences and correlations between processed and actual NAIP images in three bands to evaluate the model performance. This progress showcases the potential of GANs in translating satellite imagery to NAIP-like imagery. Moreover, it offers a clue to refine field-level crop-type mapping and boundary extraction in other nations without NAIP.***

***Keywords: GAN, Pix2Pix, NAIP, Sentinel-2, Field level crop mapping***

## I. INTRODUCTION

The field-level crop-type mapping is well-known for agriculture management and relevant decision-making. The successful cases, including the United States Department of Agriculture's (USDA) Cropland Data Layer (CDL), In-season CDL [1], [2], [3], and refined CDL was proposed and published [4]. The existing CDL products usually have 30m spatial resolution with well performance for large crop fields within the U.S.; however, the mixing pixels is unavoidable for certain places with small and fragment fields like Kenya and India. This issue introduces a challenge to global crop mapping. The National Agricultural Imagery Program (NAIP), administered by the USDA Farm Service Agency (FSA), conducted high-resolution airborne imagery, provided a clue to eliminate this issue. This program encompasses the Continental United States (CONUS) during the agricultural major growing season (June-August), providing high-resolution imagery (0.6 meters) as an invaluable resource for field-based agricultural research. NAIP data can easily extract field edges to conduct the object classification or segment fields after classification. Meanwhile, it can also provide basic data for agriculture management research like irrigation schedule decision-making [5], [6]. However, the NAIP-like data products are still limited in other nations due to being costly and time-consuming, causing a challenge for global precise crop mapping.

To obtain NAIP-like data during summer growing season, we proposed a novel approach to generate synthetic NAIP images from Sentinel-2 inputs. This approach depended on conditional generative adversarial networks (cGAN), explicitly employing the pix2pix cGAN model, Using NAIP images as the training set and the pix2pix model as the training architecture, our goal is to leverage Sentinel-2 imagery to produce high-resolution fake-NAIP images. The efficacy of cGANs in learning mappings from input to output images while concurrently optimizing a tailored loss function has been substantiated in the seminal work on diverse image-to-image translation tasks[7].

The application of generative adversarial networks (GANs) for enhancing and synthesizing satellite imagery has gained significant traction in recent years. Several studies have demonstrated the utility of conditional GANs (cGANs) in image-to-image translation tasks across various domains. The pix2pix framework [7]employs cGANs for various image translation tasks, such as converting aerial photos to maps. This approach highlights the effectiveness of using a U-Net-based generator architecture combined with a PatchGAN discriminator, which penalizes structures at the patch scale rather than globally. This architecture has proven effective in maintaining local coherence while allowing large-scale transformations. Subsequent research has built upon the pix2pix framework to address specific remote sensing and satellite imagery challenges. The CycleGAN framework [8]extends the pix2pix model by introducing cycle consistency to enable unpaired image-to-image translation. This approach has benefited applications where paired training data is scarce, enabling more flexible and broad applications of image translation without requiring exact pairings of input and output images. Integrating domain knowledge from physical sciences into deep learning models has proven essential for ensuring physical consistency in synthetic satellite imagery generation. Another research introduced the Earth Intelligence Engine[9], which is a generative vision pipeline that synthesizes physically consistent satellite imagery for flood risk visualization. This work leverages physics-based models to improve the trustworthiness and interpretability of generated imagery, bridging the gap between data-driven methods and physical modeling.

Moreover, advancements in edge-enhanced GANs have refined the ability to super-resolve remote sensing images, addressing issues such as detail loss and blurring [10]. These techniques ensure that the high-resolution images retain critical edge details, essential for accurate interpretation and analysis. These foundational works collectively underscore the versatility and efficacy of GANs, particularly cGANs, in addressing various challenges in satellite imagery analysis and synthesis. The ongoing advancements continue to expand the capabilities and applications of these models in remote sensing and beyond.

## II. METHDOLOGY

### A. Study Area and Data Collection

In this study, we set a place in the Northern Missouri state area as our study area, which was covered by both Sentinel-2 and NAIP images. TABLE I presents the image ID for both Sentinel-2 and NAIP images we used in this project.

TABLE I DATA COLLECTION

| | **Data Collection** | |
|---|---|---|
| | ***Data type*** | ***Image Id*** |
| 1 | Sentinel-2 | COPERNICUS/S2_SR/20220622T164849_20220622T170254_T15SW |
| 2 | Sentinel-2 | COPERNICUS/S2_SR/20220620T165901_20220620T171040_T15SWB |
| 3 | Sentinel-2 | COPERNICUS/S2_SR/20220620T165901_20220620T171040_T15SWC |
| 4 | Sentinel-2 | COPERNICUS/S2_SR/20220620T165901_20220620T171040_T15SWC |
| 5 | NAIP | USDA/NAIP/DOQQ/m_3609214_sw_15_060_20220621 |
| 6 | NAIP | USDA/NAIP/DOQQ/m_3709250_sw_15_060_20220621 |
| 7 | NAIP | USDA/NAIP/DOQQ/m_3809228_ne_15_060_20220620 |
| 8 | NAIP | USDA/NAIP/DOQQ/m_3809230_ne_15_060_20220620 |

We chose the area because it has cloud-free Sentinel-2 and NAIP at the similar time in the end of June 2022. The area includes various landform features, such as cropland, architecture, natural vegetation, and water content.

The Sentinel-2 and NAIP data for this study was collected using Google Earth Engine (GEE) within a specific timeframe and geographic region. Fig.1 includes the whole Missouri state covered by NAIP 2022 image composite, the main study area, and the pixels we picked for the experiment. The process involved creating composites of both image types to ensure consistency and clarity for subsequent analysis and model training.

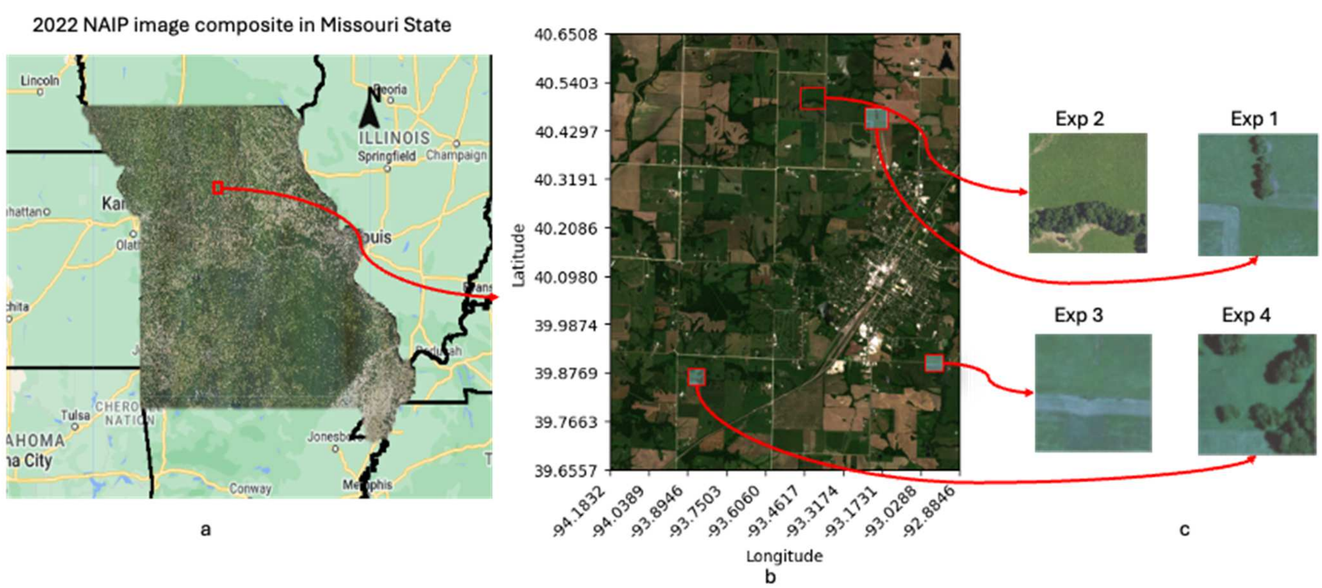


Fig. 1 Study area (*a. Missouri State NAIP 2022 image composite; b. Study area cover by NAIP 2022; c. experiment image chips*)

### B. Workflow

Figure 2 displays the workflow for generating high-resolution NAIP images from Sentinel-2 data involves several key steps. Initially, data is collected and preprocessed from Google Earth Engine (GEE), focusing on areas with cloud-free Sentinel-2 imagery and corresponding NAIP images. Sentinel-2 data is filtered to select bands B4, B3, and B2 for RGB composition, and specific date ranges are chosen to match the NAIP acquisition period. The selected date is exported and downloaded as GeoTIFF images. Sentinel-2 and NAIP chips are then created, with both sets resized to 256x256 pixels. These image chips are fed into the pix2pix GAN model to generate high-resolution NAIP images from the Sentinel-2 inputs. The generated NAIP images undergo validation to assess their accuracy and quality against the real NAIP images.

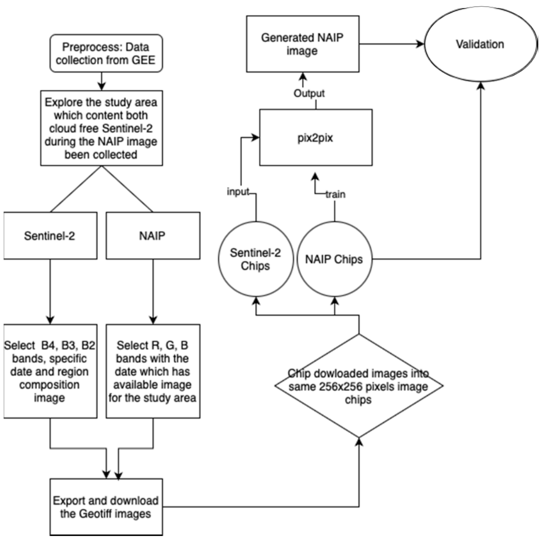


Fig. 2 Workflow

## C. *Multi-spectrum Data Pre-processing*

Sentinel-2 Data Collection: Sentinel-2 data was collected from June 20 to June 28, 2022. Relevant bands (B2, B3, B4, B8, B11, B12) were selected, and cloud masking was applied using a 10% cloud probability threshold. The median composite, filtered by date and geography, provided a clear, cloud-free image for further processing. NAIP Data Collection: The NAIP data used the images on June 28, 2022 which were the closet date image with available Sentinel-2. Composite Visualization: The Sentinel-2 composite was visualized in true color using bands B4, B3, and B2 with adjusted values for clarity. The NAIP composite used RGB bands for true color visualization. Both composites were added to the GEE map for visual inspection and validation.

The preprocessing steps involved aligning and normalizing Sentinel-2 and NAIP images to ensure compatibility for training a Generative Adversarial Network (GAN) model. The Sentinel-2 imagery, initially in a different spatial resolution and coordinate reference system (CRS), was reprojected and cropped to match the NAIP image specifications. The reprojecting process utilized the GDAL library through Rasterio, calculating the necessary transformation parameters to align the Sentinel-2 image to the CRS of the NAIP imagery. Specifically, the calculate_default_transform and reproject functions from Rasterio were employed, ensuring pixel alignment and maintaining image integrity during transformation.

Following reprojection, the Sentinel-2 image was cropped to match the NAIP image bounds using the mask function in conjunction with a bounding box derived from the NAIP image metadata. This ensured that the spatial extents of both datasets were identical, facilitating direct comparison and model training. The resulting reprojected and cropped Sentinel-2 image was then resampled to a resolution of 256x256 pixels using OpenCV's resize function, a critical step to match the input dimensions required by the GAN model.

Normalization of the pixel values was also performed, scaling the data to a range between -1 and 1, which is optimal for the tanh activation function employed in the generator network. The preprocessed Sentinel-2 and NAIP images were subsequently used to train the GAN, where the generator model synthesized high-resolution NAIP images from the Sentinel-2 inputs. Once trained, the generator model was capable of producing synthetic NAIP imagery, which was validated against actual NAIP data.

This preprocessing pipeline ensured that the input Sentinel-2 and target NAIP images were spatially and spectrally aligned, facilitating practical training of the GAN model and improving the fidelity of the generated outputs. The alignment and normalization procedures are critical for maintaining the geometric and radiometric integrity of the datasets, which are paramount for the success of image-to-image translation tasks in remote sensing applications.

## D. *Training*

This research used the Pix2Pix to translate satellite imagery from Sentinel-2 to high-resolution NAIP images. The model consists of three main components: the discriminator, the generator, and the composite model, which combines both the generator and discriminator for joint training. The discriminator, designed to differentiate between real and synthesized images, was constructed using a series of convolutional layers with leaky ReLU activations and batch normalization, culminating in a patch-based output layer to classify image patches as real or fake. The generator employed an encoder-decoder architecture with skip connections, where the encoder progressively reduces the spatial dimensions of the input image through convolutional layers, and the decoder restores the original dimensions through transposed convolutions, integrating skip connections to preserve high-resolution details.

The encoder blocks in the generator comprised convolutional layers with leaky ReLU activations. In contrast, the decoder blocks incorporated transposed convolutional layers, batch normalization, and optional dropout layers, concluding with a tanh activation function to generate the output image. The generator model was trained to minimize the L1 loss between the generated and real images, promoting accurate reconstruction. In contrast, the discriminator was trained using binary cross-entropy loss to distinguish real images from generated ones.

For training, the dataset consisting of paired Sentinel-2 and NAIP images was preprocessed to normalize pixel values between -1 and 1 and split into patches suitable for input to the GAN. The training process involved alternately updating the discriminator and generator, where the discriminator was first trained on real and fake image pairs, followed by updating the generator to improve the quality of the synthesized images based on the discriminator's feedback. The Adam optimizer with a learning rate 0.0002 and a beta1 parameter of 0.5 was employed to train both models. The model's performance was periodically evaluated by generating samples from the validation set and saving the generator's weights, ensuring the progressive improvement of image synthesis quality throughout the training epochs. This approach enabled the model to learn intricate mappings from Sentinel-2 to NAIP imagery, facilitating the generation of high-fidelity synthetic NAIP images from Sentinel-2 inputs.

## E. *Validation*

The validation of the generated NAIP images was conducted by comparing the digital values of each pixel between the generated (fake) and real NAIP images. This comparison involved converting both sets of images into binary representations using a thresholding process, where pixel values above a certain threshold were set to one and those below were set to zero. Precision, recall, and F1 score metrics were then computed to quantify the model's performance. Precision measured the ratio of correctly identified relevant pixels to the total pixels identified as relevant in the generated images. Recall assessed the ratio of correctly identified relevant pixels to the total actual relevant pixels in the real images. The F1 score, which combines precision and recall, provided a comprehensive evaluation of both the accuracy and completeness of the generated images. High scores across these metrics indicated a close match between the generated and real NAIP images, confirming the effectiveness of the pix2pix GAN model in accurately replicating high-resolution satellite imagery.

## III. Expirement and Result

### A. Experiment design and results

The experiment was conducted four times, each with a different Sentinel-2 image from distinct study areas input into the trained model. Figure 3 presents the results of these experiments, showcasing the original Sentinel-2 image, the generated NAIP image, and the real NAIP image side by side. Both visual inspection and statistical analysis were employed to evaluate the model's performance.

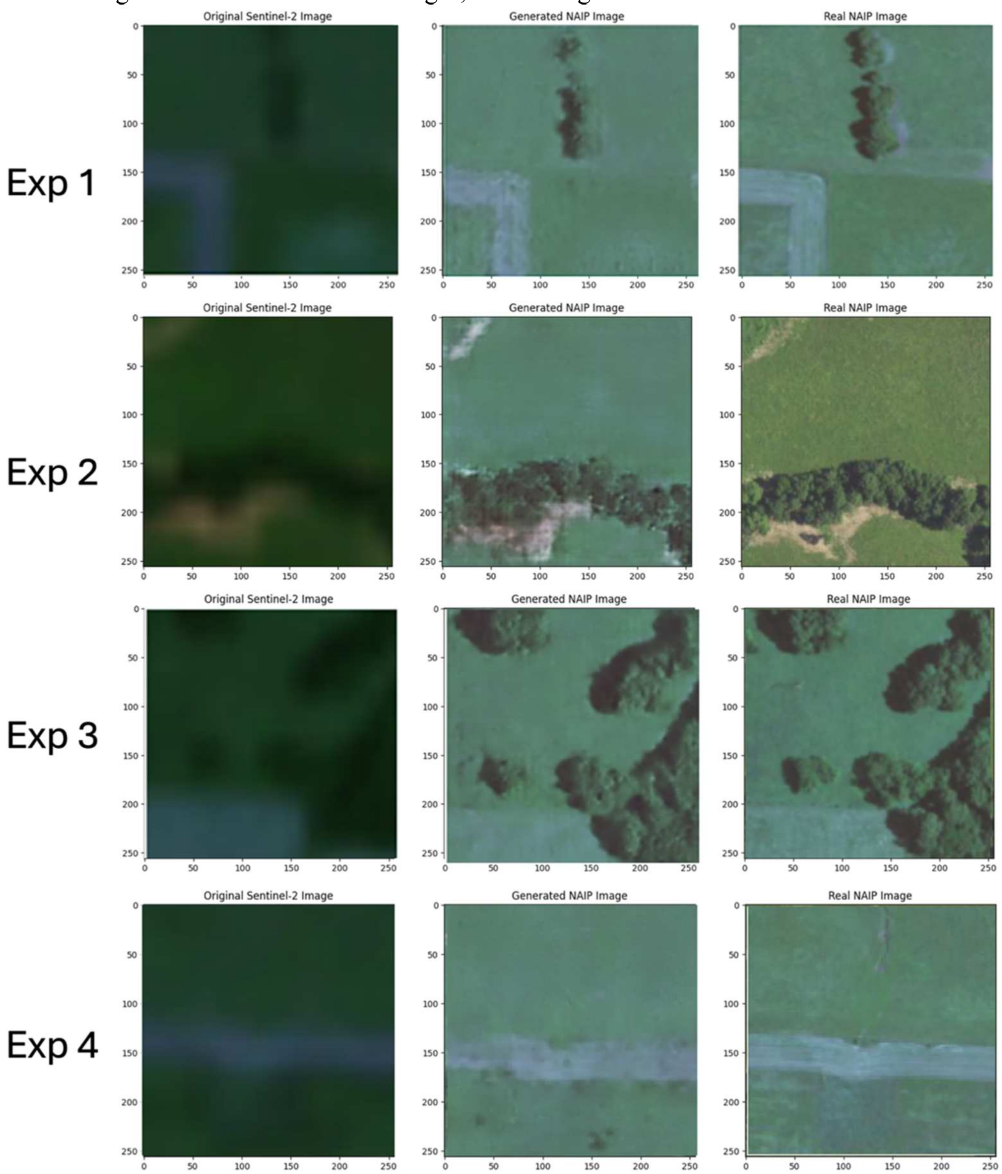


Fig. 3. Experiment 1 to 4 Sentinel-2, Generated NAIP, and Real NAIP image comparison

### B. Visual Comparison

When the generated NAIP images are compared with the original Sentinel-2 inputs, there is a significant improvement in resolution, specifically in the clarity of the edges of each object on the ground. When comparing the generated NAIP with the real NAIP, we found that both experiments performed well in enhancing the edges of objects. However, in the generated NAIP by experiment 2, we found a significant color difference compared to the real NAIP.

### C. Statistical Analysis

Table II presents the validation results of the generated NAIP images using precision, recall, and F1 score metrics.

The analysis indicates that the generated images are highly similar to the real NAIP images. Precision scores range from 0.71 to 0.96, with an average of around 88% accuracy, demonstrating the model's ability to identify relevant features within the imagery accurately. Recall scores average 0.91, reflecting the model's effectiveness in capturing the comprehensive set of relevant features. The F1 scores, which balance precision and recall, consistently fall between 0.74 and 0.96, with an average of 0.89, highlighting the robustness of the model in generating high-fidelity NAIP images. These results confirm the efficacy of the pix2pix GAN model in enhancing Sentinel-2 data, producing high-resolution images that closely align with the real NAIP images. The validation process substantiates the reliability of the generated outputs, validating the model's performance across multiple image sets. This comprehensive assessment underscores the potential of the pix2pix GAN model in practical applications requiring high-resolution satellite imagery translation.

TABLE II VALIDATION RESULTS

| **Experiments** | | | |
|---|---|---|---|
| | ***Precision*** | ***Recall*** | ***F1 Score*** |
| Exp 1 | 0.92 | 0.98 | 0.94 |
| Exp 2 | 0.71 | 0.77 | 0.74 |
| Exp 3 | 0.94 | 0.92 | 0.93 |
| Exp 4 | 0.96 | 0.97 | 0.96 |
| Avg. | 0.88 | 0.91 | 0.89 |

## IV. CONCLUSION

The study introduced a method to utilize Sentinel-2 imagery for creating synthetic NAIP (National Agriculture Imagery Program) images with a significantly improved resolution, from 10-meter pixels to 0.6-meter pixels. The improved resolution and sharpness of the generated NAIP images facilitate in-depth monitoring and analysis of agricultural fields. Future work will involve testing and refining the model with Sentinel-2 images from locations outside the United States, with the goal of establishing a global database of NAIP-like imagery. This expansion could further improve agricultural monitoring and research on a global scale.


## ACKNOWLEDGMENT

This study was supported by grants from NSF (Grant #: 2228000 and 2345039, PI: Dr. Liping Di)


## REFERENCES


[1] C. Boryan, Z. Yang, R. Mueller, and M. Craig, “Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program,” Geocarto Int., vol. 26, no. 5, pp. 341–358, Aug. 2011, doi: 10.1080/10106049.2011.562309.

[2] C. Zhang et al., “Towards automation of in-season crop type mapping using spatiotemporal crop information and remote sensing data,” Agric. Syst., vol. 201, p. 103462, Aug. 2022, doi: 10.1016/j.agsy.2022.103462.

[3] H. Li et al., “Automated In-season Crop-type Data Layer Mapping without Ground Truth for the Conterminous United States based on Multisource Satellite Imagery,” IEEE Trans. Geosci. Remote Sens., pp. 1–1, 2024, doi: 10.1109/TGRS.2024.3361895.

[4] L. Lin et al., “Validation and refinement of cropland data layer using a spatial-temporal decision tree algorithm,” Sci. Data, vol. 9, no. 1, p. 63, Dec. 2022, doi: 10.1038/s41597-022-01169-w.

[5] Haoteng Zhao; Liping Di; Liying Guo; Lin Li; Chen Zhang; Eugene Yu; Hui Li, “Optimizing Irrigation Scheduling Using Deep Reinforcement Learning,” presented at the 2023 11th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Wuhan, China, 2023, pp. 1–4. Accessed: Jun. 17, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10233673

[6] H. Zhao, L. Di, L. Guo, C. Zhang, and L. Lin, “An Automated Data-Driven Irrigation Scheduling Approach Using Model Simulated Soil Moisture and Evapotranspiration,” Sustainability, vol. 15, no. 17, Art. no. 17, Jan. 2023, doi: 10.3390/su151712908.

[7] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-Image Translation with Conditional Adversarial Networks.” arXiv, Nov. 26, 2018. Accessed: Jun. 07, 2024. [Online]. Available: http://arxiv.org/abs/1611.07004

[8] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired Image-To-Image Translation Using Cycle-Consistent Adversarial Networks,” presented at the Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2223–2232. Accessed: Jun. 19, 2024. [Online]. Available: https://openaccess.thecvf.com/content_iccv_2017/html/Zhu_Unpaired_Image-To-Image_Translation_ICCV_2017_paper.html

[9] B. Lütjens et al., “Physically-Consistent Generative Adversarial Networks for Coastal Flood Visualization.” arXiv, Feb. 21, 2023. doi: 10.48550/arXiv.2104.04785.

[10] K. Jiang, Z. Wang, P. Yi, G. Wang, T. Lu, and J. Jiang, “Edge-Enhanced GAN for Remote Sensing Image Superresolution,” IEEE Trans. Geosci. Remote Sens., vol. 57, no. 8, pp. 5799–5812, Aug. 2019, doi: 10.1109/TGRS.2019.2902431.