Data Augmentation with Generative Adversarial Network for Solar Panel Segmentation from Remote Sensing Imagery

Justinas Lekavičius[1,], Valentas Gružauskas[1]*

[1] Institute of Computer Science, Vilnius University, Lithuania, me@justinaslekavicius.com, valentas.gruzauskas@mif.vu.lt
* corresponding author
doi: 10.5281/zenodo.11548410

Abstract: With the increasing popularity of solar energy in the electricity market, demand arises for data such as precise locations of solar panels for efficient energy planning and management. However, this data is not easily accessible; information such as precise locations sometimes does not exist. Furthermore, existing data sets for training semantic segmentation models of photovoltaic installations are limited, and the manual annotation of remote sensing imagery is time-consuming and labour-intensive. Therefore, the pix2pix generative adversarial network (GAN) is used to create additional remote sensing data, enriching the original resampled training data of varying ground sampling distances without compromising its integrity. Experiments with the DeepLabV3 model, ResNet-50 backbone, and pix2pix generative adversarial network architecture were conducted to discover the advantage of using GAN-based data augmentations for a more accurate remote sensing imagery segmentation model. The result is a fine-tuned solar panel semantic segmentation model, trained using transfer learning and an optimal amount – 60% of GAN-generated RS imagery for additional training data. The findings demonstrate the benefits of using GAN-generated images as additional training data, addressing the issue of limited data sets, and increasing IoU and F1 metrics by 2% and 1.46%, respectively, compared to classic augmentations. The improved semantic segmentation model allows for better solar panel detection in remote sensing images and the potential development of a regional photovoltaic installation map for better electricity network planning and risk management.

Keywords: solar panels; semantic segmentation; data augmentation; generative adversarial network; remote sensing.

1 Introduction

The usage of renewable energy is becoming more widespread every year due to its increasing availability for homeowners and energy production possibilities for electricity providers. The growing expansion of solar panel usage also increases the need for data such as precise panel locations, types, and specifications for effective energy grid planning and management (Guangul and Chala 2019). Nevertheless, such data is limited due to various factors such as privacy concerns, relaxed requirements, and the time-consuming and resource-intensive nature of collecting this data. The combination of machine learning and remote sensing emerges as a solution for this problem, specifically using satellite imagery for solar panel detection and analysis. Semantic segmentation using convolutional neural networks such as U-Net (Ronneberger et al. 2015) and FCN (Long et al. 2015) allows for accurate solar panel identification in remote sensing images. Furthermore, the solar panel segmentation task can be achieved with better results with the introduction of even more advanced networks such as RU-Net (Li and Lau 2022) and EfficientNet-B5 (Ge et al. 2022). Despite these advances, the issue of the need for annotated and diverse datasets for effective model training remains (Sun et al. 2021). To solve this problem, data augmentation techniques are helpful, including classic augmentations such as rotations and scaling adjustments, and generative adversarial networks for domain translations and brand-new data creation. One of these networks for image-to-image translation is pix2pix (Isola et al. 2018), introducing the possibility of creating new realistic images from already existing limited data. This additional data can be used as supplementary imagery for semantic segmentation model training, increasing its performance. The goal of this study is to improve the solar panel detection accuracy of the DeepLabV3 semantic segmentation model, using the pix2pix generative adversarial network for additional training data creation. Transfer learning and fine-tuning are also applied for the DeepLabV3 model, and the issue of a limited dataset and manual labelling is addressed by creating new training data as an alternative to classic augmentations, which do not create entirely new images.

2 Materials and methods

A collection of various remote sensing images and solar panel segmentation mask pairs was used for the training of both the semantic segmentation model and the generative adversarial network. These images are of various spatial resolutions (0.8m, 0.3m, 0.2m, 0.1m) and sizes (1024x1024, 400x400, 256x256). For each spatial resolution, 640 image-mask pairs were selected, and the total amount of data was 2560 pairs, with an 80/10/10 split for training, validation, and testing subsets. To solve the issue of imagery resolution differences and to avoid scale discrepancies, the images were resampled to 512x512 target size and 0.1m target spatial resolution for consistency during model training. The width and height of the image and its segmentation mask are resampled as shown in equation (1).

Then the newly resampled images and masks are either padded or cropped depending on whether the size is smaller or larger than the target size. If they are smaller, padding is applied, as shown in equation (2). If they are larger, cropping in the region of interest (center of largest solar panel object in segmentation mask) is applied. This is done to prevent as much information loss as possible.

The pix2pix generative adversarial network is trained using the same training data as for the DeepLabV3 semantic segmentation model, for the task of translating semantic segmentation masks (input data) to real remote sensing images (target data). In the process, the generator attempts to generate realistic images that are faithful to the target data, while the discriminator tries to tell apart real images from fake ones. The D_real and D_fake losses are back-propagated to the discriminator, indicating how well the discriminator can identify real and fake generated images. The G_L1 and G_GAN losses are back-propagated to the generator and signal the difference between the original and synthetic image, and how well the generator can fool the discriminator. The training process is displayed in Figure 1.

Because the new remote sensing images are generated from existing training data (2048 image-mask pairs), the result is an additional 2048 pairs of new data that provide more variety compared to regular augmentation applications. An example of generated images compared to original data can be seen in Figure 2.

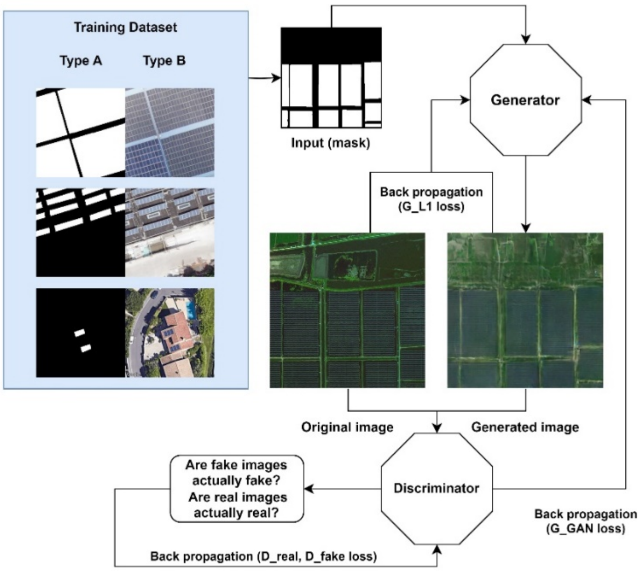


Figure 1. Diagram of Pix2pix generative adversarial network training process for creating new remote sensing images from segmentation masks.

$$\text{new dimension} = \text{original dimension} * \left(\frac{\text{spatial resolution}}{\text{target spatial resolution}}\right) \quad (1)$$

$$\text{padding dimension} = \left(\frac{\text{target dimension} - \text{new dimension}}{2}\right) \quad (2)$$

Figure 2. Original remote sensing images (top) and newly GAN-generated data (bottom).

Six different models were trained to test the benefits of using GAN-generated data as additional data for semantic segmentation model training versus using regular augmentations. The regular data augmentations performed for the training data were random horizontal flip with a 50% chance of it being applied, random rotation of 5 degrees,

random perspective change with 0.05 distortion scale, and 50% chance of it being applied, and random application of Gaussian blur (5x5 kernel size and standard deviation of 0.1 min and 2.0 max) with 50% chance of it being used. The trained models are:

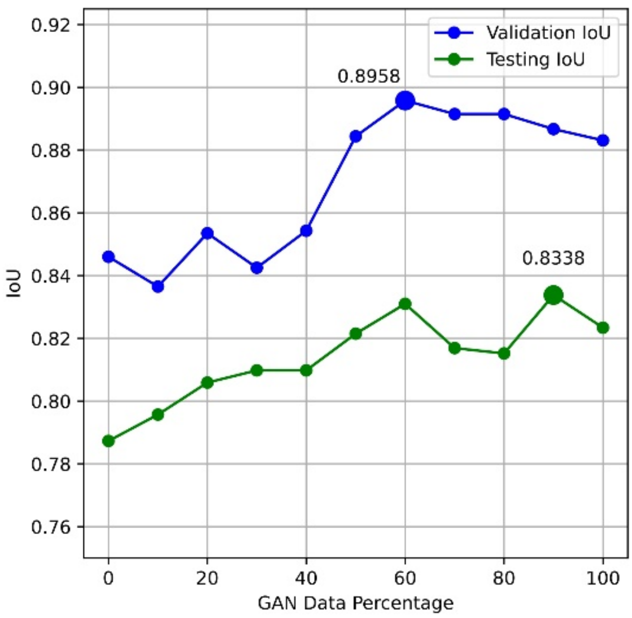
- No_a – trained with the original training subset.
- Basic_a – trained with additional applied classic augmentations.
- Gan25 – trained with 25% GAN-generated images as extra training data.

- Gan25a – trained with additional applied classic augmentations.
- GanO – trained with the optimal amount of GAN-generated images as extra training data.
- GanO_a – trained with additional applied classic augmentations.

3 Results and discussion

To determine the optimal GAN-generated data amount for training the semantic segmentation model, sensitivity analysis was performed, training the DeepLabV3 model using transfer learning and incremental amounts of additional GAN-generated data. This method finds the optimal amount of generated synthetic data before the model is impacted negatively. Validation and testing subset IoU and loss metrics were compared, and the changes with additional GAN data usage are detailed in Figure 3. Because the criteria used for determining the best-performing variation of the model was validation IoU, 60% was selected as the best number of GAN-generated images to use for appending the original training data.

The testing results of all six models are compared in Table 1. Compared to the baseline model No_a, the model GanO displayed improvements across all metrics, with pixel accuracy increasing by 0.78%, precision – 3.41%, sensitivity – 2.49%, F1 – 2.71%, IoU – 3.19%, and the loss decreasing by 0.0282. Based on the IoU metric, which is important in determining the percentage of correctly segmented solar panels, the 256 images in the testing subset were categorized by good segmentation (At least 50% of the solar panel found), poor segmentation (less than 50% of the solar panel found), and by no solar panels found. The model GanO, trained with an additional optimal amount (60%) of GAN-generated data, also correctly segmented more solar panels in remote sensing images compared to the baseline No_a model.



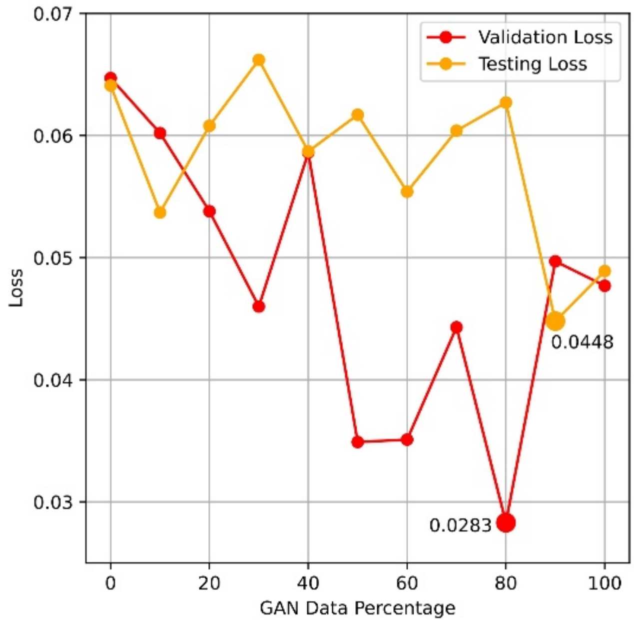


Figure 3. Sensitivity analysis of DeepLabV3 model training using transfer learning and various percentages of additional GAN-generated training data.

Table 1. Testing results of various models trained during the experiments.

Model	**Avg Acc (%)**	**Avg Prec (%)**	**Avg Rec (%)**	**Avg F1 (%)**	**Avg IoU (%)**	**Avg Loss**	**Segmented Solar Panels (IoU)**		
							Good ≥ 0.5	Poor < 0.5	None = 0
No_a	97.89	86.72	85.62	85.25	80.13	.0650	229	12	15
Basic_a	97.88	89.63	86.51	86.50	81.32	.0547	235	13	8
Gan25	98.09	89.11	85.94	85.71	80.42	.0586	229	16	11
Gan25a	97.91	88.84	87.25	87.08	81.41	.0550	238	8	10
GanO	98.67	90.13	88.81	87.96	83.32	.0368	237	11	8
GanO_a	98.04	89.82	87.69	87.77	82.90	.0611	238	9	9

4 Conclusions

In summary, using the generative adversarial network is beneficial for training dataset augmentation with brand new remote sensing images from existing limited data, especially when the optimal amount of GAN-generated data is determined with a sensitivity analysis. Expanding a limited data set with newly generated pictures helps train the solar panel segmentation model for better accuracy and segmentation quality compared to using classic data augmentations. Furthermore, having new images produced using the generative adversarial network eliminates the need for time-consuming manual annotations for training data creation With even more improvements to not only the segmentation model hyperparameters but also the generative adversarial network parameters, fine-tuned semantic segmentation model may be used for even more precise solar panel detection in remote sensing images and development of a regional solar panel map, allowing for more accessible solar panel market analysis and risk management.

5 References

Ge, F., Wang, G., He, G., Zhou, D., Yin, R., Tong, L., 2022. A Hierarchical Information Extraction Method for Large-Scale Centralized Photovoltaic Power Plants Based on Multi-Source Remote Sensing Images. Remote Sensing 14, 4211.

Guangul, F.M., Chala, G.T., 2019. Solar Energy as Renewable Energy Source: SWOT Analysis, in: 2019 4th MEC International Conference on Big Data and Smart City (ICBDSC). Presented at the 2019 4th MEC International Conference on Big Data and Smart City (ICBDSC), pp. 1-5.

Isola, P., Zhu, J.-Y., Zhou, T., Efros, A. A., Image-to-Image Translation with Conditional Adversarial Networks. arXiv, Nov. 26, 2018. http://arxiv.org/abs/1611.07004 (Accessed 11 June, 2024).

Li, L., Lau, E., 2022. RU-Net: Solar Panel Detection From Remote Sensing Image, in: 2022 IEEE Green Energy and Smart System Systems (IGESSC). Presented at the 2022 IEEE Green Energy and Smart Systems Conference (IGESSC), IEEE, Long Beach, CA, USA, pp. 1-6.

Long, J., Shelhamer, E., Darrell, T., 2015. Fully Convolutional Networks for Semantic Segmentation.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation.

Sun, X., Wang, B., Wang, Z., Li, Hao, Li, Hengchao, Fu, K., 2021. Research Progress on Few-Shot Learning for Remote Sensing Image Interpretation. IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing 14, 2387-2402.