
Marcus Vinícius Coelho Vieira da Costa 1,2, Osmar Luiz Ferreira de Carvalho 3, Alex Gois Orlandi 1,2, Issao Hirata 1, Anesmar Olino de Albuquerque 4, Felipe Vilarinho e Silva 1, Renato Fontes Guimarães 2, Roberto Arnaldo Trancoso Gomes 2, and Osmar Abilio de Carvalho Júnior 2,4

1 Superintendency of Information Technology, Brazilian Electricity Regulatory Agency, Brasilia 70.910-900, Brazil; marcus.imagem@aneel.gov.br (M.V.C.V.d.C.); alexorlandi@aneel.gov.br (A.G.O.); issaohirata@aneel.gov.br (I.H.); felipesilva.imagem@aneel.gov.br (F.V.e.S.)
2 Department of Geography, University of Brasilia, Brasilia 70.910-900, Brazil; anesmar@ieee.org (A.O.d.A.); renatofg@unb.br (R.F.G.); robertogomes@unb.br (R.A.T.G.)
3 Department of Computer Science, University of Brasilia, Brasilia 70.910-900, Brazil; osmarcarvalho@ieee.org
4 Correspondence: osmarjr@unb.br

Abstract: Brazil is a tropical country with continental dimensions and abundant solar resources that are still underutilized. However, solar energy is one of the most promising renewable sources in the country. The proper inspection of Photovoltaic (PV) solar plants is an issue of great interest for the Brazilian territory’s energy management agency, and advances in computer vision and deep learning allow automatic, periodic, and low-cost monitoring. The present research aims to identify PV solar plants in Brazil using semantic segmentation and a mosaicking approach for large image classification. We compared four architectures (U-net, DeepLabv3+, Pyramid Scene Parsing Network, and Feature Pyramid Network) with four backbones (Efficient-net-b0, Efficient-net-b7, ResNet-50, and ResNet-101). For mosaicking, we evaluated a sliding window with overlapping pixels using different stride values (8, 16, 32, 64, 128, and 256). We found that: (1) the models presented similar results, showing that the most relevant approach is to acquire high-quality labels rather than models in many scenarios; (2) U-net presented slightly better metrics, and the best configuration was U-net with the Efficient-net-b7 encoder (98% overall accuracy, 91% IoU, and 95% F-score); (3) mosaicking progressively increases results (precision-recall and receiver operating characteristic area under the curve) when decreasing the stride value, at the cost of a higher computational cost. The high trends of solar energy growth in Brazil require rapid mapping, and the proposed study provides a promising approach.

Keywords: solar panel; deep learning; semantic segmentation

1. Introduction

Solar energy is one of the most promising renewable energy sources, being crucial for sustainable development in places with intense sunlight. Several studies have shown that solar energy systems allow for economic and efficiency gains, driven by technological and productive development that enables cost reduction to overcome technical barriers [1,2]. According to Sampaio and González [3], the main advantages of solar energy systems are reliability, low costs of operation and servicing, low maintenance, a free energy source, clean energy, high availability, generation closer to the consumer, a low environmental impact, potential to mitigate greenhouse gas emissions, and noiselessness. In contrast, the main disadvantages are a high initial cost, large installation area, high dependence on technology development, and climatic conditions (solar irradiation). The benefits of solar technology provided an exponential increase in installed solar energy capacity between 1992 and 2020 [4,5]. This detected growth of solar energy was not foreseen in previous scenarios of the Intergovernmental Panel on Climate Change’s fifth assessment report [6].
Creutzig et al. [7] considered that the cause of underestimating the potential of solar energy was rapid technological learning and political support only for specific technologies. In 2019, China led the Photovoltaic (PV) solar energy capacity, followed by the European Union and the United States of America, where together they hold more than 64% of the world’s total capacity.

Brazil offers good prospects for net-zero carbon energy due to its abundance of renewable energies: hydropower, bioenergy, wind, and solar [8]. Hydroelectric power is the primary generator of electric energy in Brazil. However, thermal energy is still needed to supply domestic demand in periods of prolonged drought [9–12]. Therefore, the challenge is to increase renewable energy production to supply the growing energy demand due to population growth and new technologies. One problem is that hydroelectric expansion prospects are in the Amazon region, with substantial environmental restrictions, such as extensive areas of flooding by dam reservoirs, methane emissions, and ecological changes [13–16]. In addition, climate change scenarios in Brazil for the 2030s and 2080s predict a decrease in rainfall and an increase in temperature, resulting in a reduction in hydroelectric production and an increase in solar (slight) and wind (significant) energy potential [17]. Thus, national progress needs to intensify alternative energy sources such as combining wind and solar sources [18].

The Brazilian territory has a high solar incidence availability, with a vast area close to the equator and without significant variations in the day’s solar duration [19]. The semi-arid region has the most significant aptitude for installing solar power plants [20–25]. Between the two solar energy generation technologies, the Brazilian government has initially prioritized PV instead of concentrated solar power [26]. In 1995, the Hydroelectric Company of San Francisco developed the first PV system connected to Brazil’s grid in Recife [27]. The crisis in the Brazilian electric sector between the years 2013 and 2015 favored the decentralization and diversification of the electric matrix sources. Therefore, since 2014, Brazil’s solar energy has experienced a substantial expansion, with the first projects for PV Plants being contracted by way of public auctions. In the second half of 2015, solar energy production and distributed generation marked an inflection of growth driven by regulations and adoption of incentive changes [28]. Despite the various barriers to the development of solar energy (technological, economic, sociocultural, managerial, environmental, and political) [29–34], the current strong growth in PV energy brings optimistic perspectives for the electricity sector. Barbosa et al. [35] demonstrate, by modeling, that PV solar energy in Brazil will reach more than 36% of total electricity in 2050. This rapid expansion is mainly due to technology development, reducing investment costs, increasing the PV Panels capacity, and other enterprise cost reductions [36]. Furthermore, energy security policies and the eco-label design for improving air quality, by reducing greenhouse gas emissions, also contribute to solar energy growth [37].

Moreover, in developing countries such as Brazil, the PV solar plants are vital for ensuring energy security. Thus, inspecting solar plant constructions is important in order to carry out effective public policies. In Brazil, the Brazilian Electricity Regulatory Agency (ANEEL) is responsible for regulating the installed capacity expansion and monitoring the powerplant construction progress [38]. However, the inspection is manual, which will increase in complexity over time, requiring laborious work with skilled professionals and high costs for fieldworks and technical analysis. According to the ANEEL database, the growth expectancy for PV solar plant energy is considerable. For 2021, 32 new ventures are expected, and more than 140 for 2022. Furthermore, sustainable energy sources (e.g., wind and solar) tend to have many ventures with low energy production, increasing the number of processes to evaluate, urgently requiring automatic processes.

Remote sensing data (aerial photography and satellite imagery) enable inspection periodically, and have been widely used in the electrical sector for effective maintenance of electrical lines [39–41], thermal monitoring from nuclear power plants [42–45], environmental changes from hydroelectric dams [46–49], and energy consumption using nighttime light satellite imagery [50–52], among others. In solar energy, many studies use remote sens-
In automatic detection, Deep Learning (DL) emerges as a powerful method, especially in regards to computer vision problems using convolutional neural networks (CNN), due to its ability to process multi-dimensional arrays [69] with wide remote sensing applications [70–75]. Several reviews were carried out on the different DL methods, in which object detection, semantic segmentation, and instance segmentation were the most common approaches [76–78]. The method choice is highly dependent on the task objectives. When the main goal is to make a pixel-wise classification (as is the case with PV solar plants), semantic segmentation is a great alternative [79,80].

Previous studies in PV solar panel detection have shown promising results using the DL method, presenting very high accuracy. However, most studies consider urban PV panels using aerial or high-resolution satellite images [81–83], while PV solar plant mapping is still restricted [84]. This approach is an effective alternative to construction inspection, requiring periodic data and free satellite imagery. Previous studies on PV panel detection have not yet shown reasonable solutions for classifying large regions, and the use of mosaicking with sliding windows is a promising solution [85–87].

The primary motivation for this study is the development of a methodology based on remote sensing for the automatic monitoring of new installations of PV solar plants. In Brazil, the high growth of solar energy throughout the territory, with a continental dimension, prevents on-site inspection due to the financial and time cost, requiring the development of technological alternatives. Therefore, this research aims to evaluate the use of DL methods, representing the state of the art of computer vision, to identify and monitor PV solar power plants from ANEEL’s database using Sentinel-2 images. This methodology represents an innovation for the management and monitoring of installed solar energy structures on the Brazilian territory, and similar research does not exist in the country to date.

2. Materials and Methods

The present research had the following methodological steps (Figure 1): (2.1) data preparation; (2.2) DL models; (2.3) DL accuracy analysis; (2.4) mosaicking; and (2.5) mosaicking accuracy analysis.

![Figure 1. Methodological flowchart.](image-url)
2.1. Data Preparation

2.1.1. Study Area

Brazil has a large and diverse territory, presenting different solar energy incidence [21]. Nevertheless, many areas are extremely suitable for the installation of PV panels. Therefore, we selected 24 areas to conduct this experiment (Figure 2). There are limited PV plants installed in the Brazilian territory and currently no open datasets considering Sentinel-2 data [88]. However, the development of methodologies and expansion of databases is a fundamental strategy for monitoring large-scale PV with a high growth trend.

Figure 2. Study Area.
2.1.2. Image Acquisition and Annotations

We obtained Sentinel-2 cloudless images with four channels (Red, Green, Blue, and near infra-red) for each region containing PV solar power plants. For each image, a specialist manually annotated ground truth (GT) masks considering two classes: background and PV solar plant. The background class presents a wide variety of spectral behaviors, including the different soil and vegetation compositions present in a large-scale country such as Brazil. The research considered the difference in the light incidence and the construction of panels in each region for DL model training.

2.1.3. Data Split

After preparing each tile with their respective annotations, we separated the dataset into training, validation, and testing sets. For each area of interest that may contain more than one PV solar plant, we cropped at least seven 256 × 256-pixel tiles. Table 1 lists the distribution of areas and images for training, validation, and testing.

Table 1. Data split in training, validation, and test sets.

<table>
<thead>
<tr>
<th>Set</th>
<th>Number of Areas</th>
<th>Number of Images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train</td>
<td>15</td>
<td>210 (75%)</td>
</tr>
<tr>
<td>Validation</td>
<td>5</td>
<td>40 (14.28%)</td>
</tr>
<tr>
<td>Test</td>
<td>4</td>
<td>30 (10.71%)</td>
</tr>
</tbody>
</table>

2.2. DL Models

2.2.1. Architectures and Backbones

Semantic segmentation allows for a pixel-wise classification, being highly suitable for many remote sensing applications [74]. Most semantic segmentation networks include an encoder/decoder structure. The encoder aims to extract features, whereas the decoder restores the image’s original dimensions. In the last few years, many architectures were proposed to increase performance in this task (e.g., U-net [89], SegNet [90], Feature Pyramid Network (FPN) [91], DeepLab [92,93], and Pyramid Scene Parsing Network (PSPNet) [94], and backbones (e.g., ResNet [95], ResNeXt [96], and Efficient-net [97]). This study evaluated four commonly used architectures (U-net, DeepLabv3+, FPN, and PSPNet) and four backbones (ResNet-50 (R-50), ResNet-101 (R-101), Efficient-net-b0 (Eff-b0), and Efficient-net-b7 (Eff-b7)). We used models from the Semantic Segmentation repository [98], which provides different architectures and backbones in Pytorch.

2.2.2. Model Configurations

In addition to choosing the appropriate models, it is crucial to make fine adjustments for the task at hand. The first problem is the reduced number of available samples. Therefore, in addition to obtaining at least seven frames from each location, we applied two augmentations in the training process: random horizontal flip and random vertical flip (both with a probability of 0.5). The second problem is class distribution (there are many more background pixels than solar panel pixels). Thus, we used a loss function that minimizes this effect, the Dice Loss:

\[
\text{Dice Loss} = \frac{2 \times (\text{pred} \cap \text{GT})}{|\text{pred}| + |\text{GT}|},
\]

in which pred is the DL prediction, and GT is the ground truth mask. In addition, we used transfer learning with Imagenet [99] pre-trained weights for faster convergence; to avoid overfitting, we applied callbacks, saving the model with the lowest Dice Loss in the validation set. Regarding hyperparameters, we used: (a) 300 epochs; (b) Adam optimizer; (c) 5 × 10-3 learning rate (lr); and (d) batch size of 5.
2.3. DL Accuracy Analysis

Accuracy analysis is a fundamental step for DL model evaluation. Since semantic segmentation models provide a pixel-wise mask, the metrics compare the predicted mask and the GT mask through confusion matrix metrics. The confusion matrix (Table 2) has four quadrants in binary tasks: True Negatives (TN), True Positives (TP), False Positives (FP), and False Negatives (FN).

Table 2. Confusion matrix.

<table>
<thead>
<tr>
<th>Ground truth</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>TN, FP</td>
</tr>
<tr>
<td>1</td>
<td>FN, TP</td>
</tr>
</tbody>
</table>

The model outputs probability, whereas the GTs are integers. Thus, it was necessary to establish a cutoff point for the threshold metrics. A stricter threshold tends to reduce the commission errors, while a more permissive threshold tends to reduce omission errors. Thus, we applied a commonly intermediate threshold of 0.5 for three metrics (overall accuracy, F-score, and IoU):

\[
\text{Overall Accuracy} = \frac{TP + TN}{TP + TN + FP + FN},
\]

\[
\text{IoU} = \frac{TP}{TP + FP + FN},
\]

\[
\text{F-score} = \frac{TP}{TP + \frac{1}{2}(FP + FN)}.
\]

2.4. Mosaicking

The 256 × 256 pixel tiles used in training may not represent an entire scene, requiring a postprocessing stage. Mosaicking using a sliding window algorithm is a very promising solution. However, combining frames side by side to reconstruct a scene may also induce errors in the single frame edges. A way to minimize this effect is to apply a sliding window with overlapping pixels, where the final pixel will be the average from the overlapped pixels. Thus, we compared six different stride values for the mosaicking strategy: 8, 16, 32, 64, 128, and 256 (adjacent frames). Figure 3 shows four images with consecutive frames using different stride values. The smaller the stride value, the more overlapping pixels (which tends to reduce errors in the frame edges).

Figure 3. Four examples of different stride values of two consecutive frames in ascending order, where the stride value \( a < b < c < d \).
2.5. Mosaicking Accuracy Analysis

To evaluate the mosaicking, we analyzed the ranking metrics Receiver Operating Characteristic Area Under the Curve (ROC AUC) and Precision-Recall (PR) AUC, considering six stride values: 8, 16, 32, 64, 128, and 256. The ROC curve considers the true positive rate (TP/(TP + FN)) and false positive rate (FP/(TN + FP)) and the PR curve considers the precision (TP/(TP + FP)) and recall (TP/TP + FN). From the points generated, it is possible to calculate the area under these curves.

3. Results

3.1. DL Metrics Results

Overall, the different architectures and backbones presented good results (Table 3). The U-net presented the best metrics results regarding the different architectures, followed by DeepLabv3+, FPN, and PSPNet. Despite the higher complexity of the DeepLabv3+ architecture, the U-net presented better results as the targets do not present a high variance in scaling, one of the most significant benefits of this model. Moreover, although PSPNet provided the worst results, the difference is not extremely large, and the training period is considerably lower (less than half the time to train the Eff-b7 using the U-net architecture, and nearly one-fifth of the period for training on the DeepLabv3+ architecture). When analyzing the different backbones, apart from Eff-b0 with the PSPNet architecture, the results did not change significantly. Moreover, metrics-wise, the accuracy score shows high values among all models (<3% variation), possibly due to the fact that there are many more pixels corresponding to the background class than the panels class. The IoU and F-score provide much more meaningful results. The Eff-b7 using the U-net architecture had the best IoU and F-score results, and an intermediate computational cost.

Table 3. Semantic segmentation evaluation (accuracy, IoU, F-score, and epoch period) using three architectures (U-net, DeepLabv3+, and PSPNet), and four backbones (Efficient-net-b7 (Eff-b7), Efficient-net-b0 (Eff-b0), ResNet-101 (R-101), and ResNet-50 (R-50)).

<table>
<thead>
<tr>
<th>Architecture</th>
<th>Backbone</th>
<th>Accuracy (%)</th>
<th>IoU (%)</th>
<th>F-Score (%)</th>
<th>Epoch Period (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-net</td>
<td>Eff-b7</td>
<td>98.08</td>
<td>91.17</td>
<td>95.38</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Eff-b0</td>
<td>98.05</td>
<td>90.97</td>
<td>95.27</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>R-101</td>
<td>97.96</td>
<td>90.58</td>
<td>95.06</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>R-50</td>
<td>97.98</td>
<td>90.70</td>
<td>95.12</td>
<td>4</td>
</tr>
<tr>
<td>DeepLabv3+</td>
<td>Eff-b7</td>
<td>97.83</td>
<td>89.98</td>
<td>94.73</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Eff-b0</td>
<td>97.77</td>
<td>89.82</td>
<td>94.64</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>R-101</td>
<td>97.46</td>
<td>88.47</td>
<td>93.88</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>R-50</td>
<td>97.02</td>
<td>86.63</td>
<td>92.84</td>
<td>6</td>
</tr>
<tr>
<td>PSPNet</td>
<td>Eff-b7</td>
<td>97.35</td>
<td>88.03</td>
<td>93.64</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Eff-b0</td>
<td>96.73</td>
<td>85.43</td>
<td>92.14</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>R-101</td>
<td>97.06</td>
<td>86.98</td>
<td>93.04</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>R-50</td>
<td>97.23</td>
<td>87.60</td>
<td>93.39</td>
<td>3</td>
</tr>
<tr>
<td>FPN</td>
<td>Eff-b7</td>
<td>97.38</td>
<td>87.99</td>
<td>93.61</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Eff-b0</td>
<td>97.45</td>
<td>88.21</td>
<td>93.73</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>R-101</td>
<td>97.58</td>
<td>89.21</td>
<td>94.30</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>R-50</td>
<td>97.25</td>
<td>87.74</td>
<td>93.47</td>
<td>5</td>
</tr>
</tbody>
</table>

Figure 4 shows three examples from the test set, and three examples from the validation set with their corresponding original images (RGB channels), GT, and prediction. Despite some errors in the edges of the objects, these results suggest a correct identification of the target, with few errors.
Figure 4. Three examples from the test set and three examples from the validation set with their corresponding original image, ground truth (GT), and prediction.
3.2. Mosaicking Results

Table 4 shows the ROC AUC scores using the 1536 × 768 area, using six different stride values (8, 16, 32, 64, 128, and 256). The analysis only considered the best model (U-net with Eff-b7 backbone). When the stride value decreases, results progressively improve in both metrics. Nevertheless, decreasing the stride value increases the computational cost needed, becoming a significant limitation, especially for practical applications.

Table 4. ROC AUC, PR AUC, and processing time for 8, 16, 32, 64, 128, and 256 stride values.

<table>
<thead>
<tr>
<th>Stride</th>
<th>ROC AUC</th>
<th>PR AUC</th>
<th>Processing Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>99.42</td>
<td>97.85</td>
<td>2829</td>
</tr>
<tr>
<td>16</td>
<td>99.25</td>
<td>97.56</td>
<td>734</td>
</tr>
<tr>
<td>32</td>
<td>98.89</td>
<td>96.99</td>
<td>193</td>
</tr>
<tr>
<td>64</td>
<td>98.66</td>
<td>96.42</td>
<td>63</td>
</tr>
<tr>
<td>128</td>
<td>98.36</td>
<td>95.39</td>
<td>15</td>
</tr>
<tr>
<td>256</td>
<td>98.16</td>
<td>94.49</td>
<td>4</td>
</tr>
</tbody>
</table>

Figure 5 shows the original image, its corresponding GT, and the prediction using U-net with Eff-b7 backbone and 8-pixel stride value on a 1532 × 768-pixel image. This mosaicking strategy enables the classification of areas with large dimensions, outputting images with no discontinuity.

Figure 5. Mosaic representation on a 1536 × 768-pixel image with the original image, the corresponding ground truth (GT), and prediction using the U-net with Efficient-net-b7 backbone.
4. Discussion

The best result of our study was the U-net with the Eff-b7 backbone, although the other methods also reach high or adequate values. However, an unexpected result is that U-net outperformed DeepLabv3+ by a slight margin. This result is probably because the input images do not present multi-scale objects—one of the main contributions of the DeepLabv3+ method. Therefore, these results show that simpler structures may be well suited in some scenarios, highlighting the importance of testing different architectures.

Other solar panel detection studies using DL methods have demonstrated high accuracy in different locations. However, studies carried out on PV solar plants are still much lower than residential PV solar panels. Considering the large-scale solar plants, Hou et al. [84] proposed a study in China with one thousand images achieving 95% IoU from the U-net model. They used a much more significant amount of data, and the results were not dissimilar to ours (92% IoU).

Generally, accuracy results are lower in residential PV solar panels due to their smaller dimension and higher susceptibility to noise interference. Yuan et al. [100] applied a simple ConvNet for large-scale solar panel mapping from aerial images, and evaluated their model in the cities of Boston and San Francisco using completeness (0.84 and 0.87) and correctness (0.81 and 0.85) metrics. Yu et al. [101] proposed DeepSolar with a substantial amount of training data using high-resolution satellite images, obtaining 93.1% recall and 88.5% precision, results very similar to our F-score (95%). Zhuang et al. [83] applied the U-net in satellite images for residential panels, achieving 74% IoU. Recently, Jie et al. [82] combined a U-net model with edge detection networks. The authors showed that the edge detection increased performance on two city panel datasets by nearly 2% IoU. This effect may be even less prominent in large solar plants since it is easier to detect borders, as shown in our study. Even though these studies trained with smaller PV solar panels, the results show an excellent ability to segment panels even with simpler models.

Thus, the results of our and previous studies suggest that the mapping of PV solar panels should be addressed in a data-driven, rather than model-driven, perspective, i.e., the DL models do not present a significant difference, and the most important endeavor is to obtain a reliable source of generating good annotations. Moreover, the present study showed significant results using data augmentation despite a limited amount of data.

The mosaicking procedure enables the classification of areas of indefinite and large sizes. We have shown that using a smaller stride value increases performance, but also the computational cost. The stride value for a practical application should take both factors into consideration. Regarding the mosaicking technique on semantic segmentation models, de Albuquerque et al. [86] performed a comparative analysis using different stride values, presenting progressively better ROC AUC scores for lower stride values, a result also verified in our research.

This research presents many possibilities for future studies. A first proposition would be to estimate energy production using the mapping of the photovoltaic solar panel from DL, and the level of solar incidence in a specific region. Another relevant test would be evaluating radar images due to cloud cover and atmospheric interference in optical images. Although synthetic aperture radar (SAR) images are noisy, they can be useful in some scenarios. Studies comparing the frame sizes according to the proposal by Bem et al. [102] can also be valuable in understanding the model’s differences in various tasks (e.g., binary and multiclass) and object scales.

5. Conclusions

The survey and monitoring of PV solar power plants are extremely important for energy management and planning. The high growth of solar energy in Brazil, a country with continental dimensions, generates an increase in inspection processes for ANEEL that is only possible through technological innovation. Thus, this paper presented a comparison between DL models for the classification of PV solar plants using Sentinel-2 images with four spectral bands (RGB and near infra-red), comparing four architectures
(U-net, DeepLabv3+, FPN, and PSPNet) with four backbones (ResNet-50, ResNet-101, Eff-b0, and Eff-b7), totaling 16 combinations. Additionally, we used augmentation and transfer learning. The PV panel spectral and shape characteristics facilitate the accurate detection of the panels. Results were satisfactory using the different backbones and architectures, but U-net with Eff-b7 backbone presented the best results with 98% accuracy, 92% IoU, and 95% F-score. We estimate that the most critical factors when mapping PV solar panels are a reliable source of data and their possible applications. For the classification of large regions, the image mosaicking procedure significantly improves when using more overlapping pixels, minimizing edge errors. The results are also expressive when analyzing the ROC AUC score and PR AUC score, in which the results progressively increase whilst decreasing the stride value. However, the computational cost may be a significant challenge for practical applications, since the processing time significantly increases with the stride value reduction. This methodology has many applications and satisfies the conditions for automatically classifying PV solar plants using free Sentinel-2 imagery, allowing for a significant advance in monitoring the implanted infrastructure.


Funding: This research was funded by the following institutions: National Council for Scientific and Technological Development (434838/2018-7) and Coordination for the Improvement of Higher Education Personnel (Finance Code 001).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The following research materials are available online at https://github.com/osmarluiz/large-scale-solar-panel (accessed on 12 May 2021).

Acknowledgments: The authors are grateful for financial support from CNPq fellowship (Osmar Abílio de Carvalho Júnior, Roberto Arnaldo Trancoso Gomes, and Renato Fontes Guimarães). Special thanks are given to the research group of the Laboratory of Spatial Information System of the University of Brasilia for technical support. The authors thank the researchers from the ANEEL, who encouraged research with deep learning. Finally, the authors acknowledge the contribution of anonymous reviewers.

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

References

8. Gilks, H.C.; Simon, S.; Soria, R. 100% Renewable energy supply for Brazil-The role of sector coupling and regional development. *Energies* 2017, 10, 1859. [CrossRef]


