Deep Learning model to detect and differentiate installed solar panels using global satellite imagery from Portugal

LUIS MONTEIRO, Instituto Superior Tecnico, Portugal

This dissertation explores the use of deep learning models to detect and differentiate installed solar panels using global satellite images of Portugal.

The challenge lies in the incomplete documentation and registration of solar PV installations, which are not centrally registered and vary between countries. Taking advantage of aerial and satellite images, this study addresses the need for automatic identification of photovoltaic solar installations in wide geographic areas.

In this work, two models are used, one for identification and the other for differentiation, and with this strategy, we achieved some good results. For the identification of solar panels, we obtained precision and recall of 80,11 percent and 83,21percent, respectively. For the differentiation between solar panels, we also obtained a good result, although more modest, with an accuracy of 78,92 percent, but this value opens some doors for future work. These findings contribute to our understanding of solar panel installation. Using deep learning models and satellite imagery, this research improves the identification and monitoring of solar photovoltaic installations, as well as solar and thermal installations, facilitating sustainable energy systems.

Additional Key Words and Phrases: deep learning ; renewable energy ; computer vision ; satellite imagery; solar panels

ACM Reference Format:
Luis Monteiro. 2018. Deep Learning model to detect and differentiate installed solar panels using global satellite imagery from Portugal . *ACM Trans. Graph.* 37, 4, Article 111 (August 2018), 5 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION AND RELATED LITERATURE

The transition to clean energy is a global priority as the energy industry has been a major contributor to greenhouse gas emissions. Governments worldwide are striving to decrease climate risks and promote a shift to clean energy technologies while ensuring energy security. Key milestones include limiting global temperature rise, achieving net-zero carbon emissions by 2050, and improving air quality. The United Nations' Sustainable Development Goals for energy are also important, aiming to ensure access to affordable and sustainable energy for all by 2030.

According to the International Renewable Energy Agency (IRENA) [3], solar photovoltaic power is expected to be a major contributor to the renewable energy mix, with an estimated increase from 2 percent in 2018 to 25 percent in 2050 global electricity generation. However, not all solar PV installations are accurately documented or registered in central databases, leading to significant gaps in understanding their distribution and capacity. In the Netherlands, a study carried out by the company Stedin in collaboration with Sobolt estimates that one in four homes with solar panels is not registered [4]. Satellite and aerial images are increasingly being explored to automatically identify and map solar PV installations across wide geographic areas.

The lack of a comprehensive and accurate database of solar PV systems poses challenges to grid management and energy balance. Excess production in specific regions can lead to congestion and voltage fluctuations, requiring improved monitoring and operation of energy balance management. Policymakers and solar companies also require reliable data on solar PV systems to make informed decisions and develop effective policies. Accurate mapping and information on solar PV installations can facilitate market projections and policy design for supported and encouraged solar energy.

The use of satellite and aerial images for identifying solar PV systems has seen significant development in recent years. The field of solar panel identification saw its first contribution in 2015 with Stephen Lee's research using support vector machines (SVMs) and small datasets. Lee's work marked the introduction of neural network architecture, specifically convolutional neural networks (CNNs), for image classification.

Jiafan Yu's DeepSolar research [9] utilized transfer learning and CNNs for both image classification and semantic segmentation, achieving a precision of 93.2 percent and a recall of 88.9 percent. The paper by Roberto Castello [1] focused on mapping solar panel rooftop installations in Switzerland using semantic segmentation techniques with a U-Net CNN architecture.

In one of the latest works [5], Mayer converted the original TensorFlow model from Yu's research into PyTorch and achieved a precision of 92 percent and a recall of 98 percent using a Google Maps-based model. Dataset creation involved a four-step strategy to optimize solar panel system classification models and reduce dataset size.

Despite the fact key milestones have already been reached, there are still certain challenges to be addressed, such as the requirement for high-resolution images, the need to sort all training images manually, the difficulty of categorizing images from different locations, and the non-clear distinction among solar PV panels and solar thermal panels.

Various issues are linked to the difficulties of categorizing images from other locations. For example, while solar panels are the same from country to country, there are other structures on rooftops that make categorization difficult. Another example is the usage of Google-based images, because of the quality varies greatly from city to city, even within the same country. For a more concrete example, the company acting in the energy sector that despite the attempting to use different images and using distinctive threshold values, it was feasible to determine that with all of these alterations, the inferred

Author's address: Luis Monteiro, luis.m.monteiro@tecnico.ulisboa.pt, Instituto Superior Tecnico, Av. Rovisco Pais 1, Lisbon, Portugal, 1049-001.

0730-0301/2018/8-ART111 $15.00
https://doi.org/XXXXXXX.XXXXXXX

values were not reaching the precision and recall values obtained in the study [5]. These results are attached in Appendix A.

There are 2 major problems that were addressed in this work, the first was the classification of images with a resolution of 11cm/pixel, these being Portuguese images with typical characteristics of the country, the second problem addressed was the problem of differentiating thermal solar panels from photovoltaic solar panels.

2 PROTOTYPE FRAMEWORK

For the development of the prototype, two models were created, the Classification Model and the Differentiation Model.

The prototype works as follows: an image is first run through the Classification Model, which will categorize the images with one of two possible outputs "Negative" or "Positive". If no solar panel is identified, the output will be "Negative" the procedure finishes here, and the categorization of the image is without a solar panel. If a solar panel is identified, the image is classified as "Positive" and will be delivered to an additional algorithm for the Differentiation Model. The Differentiation Model classifies the solar panels present in the images as, "Solar Panel Thermal" or "Solar Panel Photovoltaic", As shown in the Figure 1.

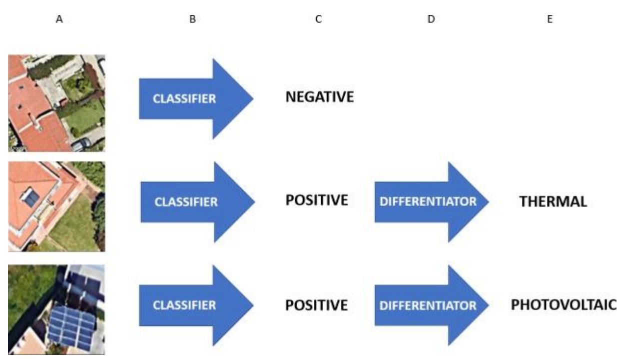

Fig. 1. Prototype Framework

A - The input satellite imagery obtained by Google Maps.
B - The Classification Model is applied.
C - The results returned by the Classification Model.
D - The Differentiation Model is applied.
E - The results returned by the Differentiation Model.

3 FILE STRUCTURE AND STORAGE OF PROJECT ARTIFACTS

All the material developed in this project could be found in a folder in my Google Drive "https://drive.google.com/drive/folders/1Rgz-LX1sLYRmYsJMrEWaH7-A5Bp5zG6-?usp=sharing" the architecture of the folder is made up of three folders, one to store the images another one to store the weights of the architectures and the last one to store the results and logs, in addition to the three folders there are two collab files, one with the Classification Model and another with the Differentiation Model with both training parts and test inside each file.

4 RESEARCH APPROACH

In the following section, I will talk about the techniques and architectures that were used in this project, we will discuss how they work and their strength. Let's start with the Data before turning our attention more toward the architectures and techniques used.

4.1 Data

The datasets provided by EDP had been analyzed, and after the cleaning, and the reclassification, the dataset revealed a lack of photographs containing solar panels, both in terms of quantity and diversity. These positive images represented only 3.4 percent of the images in the entire dataset, there was also a bias with the images as they were very similar in terms of the zones in which they were taken, as they were all taken in zones with essentially the same topography, which does not accurately represent Portugal as a whole.

So to fill in the lack of photographs of solar panels and the lack of diversity I made a plan to collect photographs from different regions of the country with different topographies, including Restelo, Faro, Birre, Belas, Benfica, and some areas with more enterprises, such as the Sport Lisboa e Benfica Pool Complex (Aquared) and the Ourika Solar Photovoltaic Centre, this way makes the model more resilient, trying to divide the dataset into urban, rural, fields and forests as was done in the study [5].

Two techniques were used for a photograph collection. The first involved using software provided by EDP, which required specifying geographic positions and zoom values to generate a grid of consecutive images with the same size, resulting in an uninterrupted map between the positions. This software had a cost per collected photograph and was mainly utilized for the test and validation datasets. The second technique involved manual collection using Google Maps, selecting different landscape types with and without solar panels, and ensuring a diverse dataset. The images collected through these techniques varied in size and resolution, enhancing the model's resilience.

It's important to note that all labeling was done manually, ensuring accuracy in dataset preparation.

All images received by the model are passed through a filter that scales them to the same size. All images are resized to the chosen size which was "[255, 255, 3]", which denotes that the filter transforms the images into images of "255" by "255" pixels, and the "3" denotes that the images have three layers of color and use the RGB standard, indicating that the model is equipped to receive images in color.

The size that was chosen was "[255, 255, 3]" given that I'm not training the model from scratch I'm using a concept called transfer learning. It made perfect sense to utilize the same size considering the neural networks were trained using the "imagenet" dataset, which contained images of the same size. The basic goal of the transfer learning method is, to begin with a head start, but when I utilized a different dimension, I experienced worse outcomes immediately.

In addition to resizing the training datasets were manipulated using a technique called Augmentation [6]. In this particular case, the images were manipulated in terms of Rotation that rotates the image, also use manipulation known as the Shear Angle that slightly deformed the image fixating one axis when stretching the image at a specific, and Brightness makes the images darker and lighter.

These decisions were taken in an effort to preserve as much as possible the normal arrangement of solar panels in Portugal. The optimal orientation in Portugal for solar systems is facing south with an inclination of at least 10 degrees and up to a maximum of 45 degrees[2].

The data collected was divided between the Training Set, Validation Set, and Test Set as well as the division within each of these datasets, namely, whether or not the photographs have solar panels for the Classification model, for the Differentiation I will use the photographs with solar panels from the Classification model and I will divide this by type of solar panel if the photograph shows a thermal solar panel or a photovoltaic solar panel.

The classification model dataset consists of a total of 9105 photographs, which are separated nos three datasets: the training set which is the bigger with 80 percent, the validation set with 10 percent, and the test set with 10 percent.

As the Differentiation model uses the positive images of the total number of images it is only 1138 photographs and is also divided into three datasets: the training set which is also the bigger one, like in the Classification Model, with 70 percent, the validation set with 20 percent, and the test set with only 10 percent.

4.2 Implementation

To build the classification model, the Inception-ResNet-v2 architecture [7] was chosen due to its strong performance. However, training the entire model with its numerous parameters proved time-consuming. To overcome this, a two-phase training approach was adopted. In the first phase, only a few lower layers were trained with a learning rate of 0.001, gradually reduced using the Reduce Learning Rate On Plateau technique, as the name indicates when the model is stuck on a plateau the learning rate drops and so the model can get better results. In the second phase, the model's weights from the previous phase were used, unlocking all layers and reducing the learning rate to 0.0001. Only if the results improved, the weights were saved.

Although this technique yielded positive results, it suffered from drawbacks such as slow training, taking approximately 10 hours for both phases. To address this, a switch to the EfficientNetV2-S[8] model was deemed necessary. The EfficientNetV2-S model, being smaller yet still capable of achieving good results, allows for training to be completed in a single phase. This not only reduced computational costs but also saved time.

4.3 Techniques

The BinaryCrossentropy is used to calculate the Loss, this loss is also influenced by the regularization techniques, and in the classification model, the loss function is also influenced by the Class Weights, penalizing the class with the largest size more. Both Adam optimizer and RMSProp were used because both always had very similar values.

4.4 Framework for Analysis

The visualization of the results goes through two main phases, the first was the results that the model obtained from the validation dataset during the course of training, and the second was when the model was already trained and the test dataset was utilized.

To visualize the results for the validation dataset, I used several strategies, the first and simplest of all was through the train function provided by Keras, which throughout each training session gives feedback on the results and at the end of each epoch, makes the respective results available for that epoch. The other two strategies derived from these results can then be viewed from the tensorboard where graphs are created from each value, the other takes these results and saves them in tables(for example: excel), this way it is easier to view all data from once and thus be able to identify relationships between them.

It wasn't until the findings were satisfactory that it moved on to this second phase. The second phase, which was with the test dataset, didn't have any progression graphs as the validation phase did; instead, it merely provided the values in a table.

5 PERFORMANCE EVALUATION METRICS

The classification model will be evaluated in a test set composed of photographs from a few places across Portugal, to make it as close to the Portuguese reality as possible. The metrics to determine the success of the algorithm were not easy to choose.

Normally, accuracy is a good assessment criterion, but it performs better when he has an even class distribution, which is not our case.

Therefore, for the case of the Classification Model that has a dataset imbalance that was chosen that uses are precision (Equation 1), recall (Equation 2), and F1 score (Equation 3).

For the Differentiation Model which has a dataset that is more traditional with a balanced dataset, we also used accuracy (Equation 4) as a metric.

Precision is defined as the proportion of accurately predicted positive observations to all expected positive observations. The Recall is the ratio of accurately predicted positive observations to all positive observations in the actual positive class. The F1 Score is calculated as the weighted average of Precision and Recall. As a result, this score considers both false positives and false negatives. The accuracy measures the proportion of correct predictions made by the model over the total number of predictions.

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \tag{1}$$

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \tag{2}$$

$$F1score = 2 \frac{Recall * Precision}{Recall + Precision} \tag{3}$$

$$Accuracy = \frac{TP + TN}{FP + TN + TP + FN} \tag{4}$$

True positives (TP) refer to successfully predicted positive samples. False positives (FP) refer to incorrectly predicted positive samples. False negatives (FN) refer to incorrectly predicted negative samples. True negatives (TN) refer to successfully predicted negative samples.

6 RESULTS

As explained in subsection 4.4 only those experiences that showed relevant results with the validation dataset were actually tested with the test dataset.

6.1 Classification Model Results

For the classification model we can highlight the best training. The graph on the left is represented by precision and the graph on the right is represented by the Recall graph.

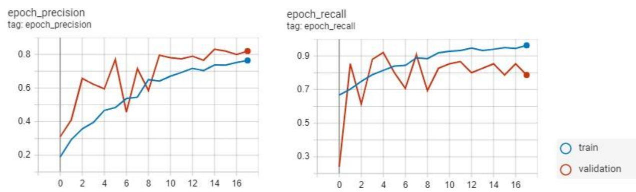

Fig. 2. Graphs of the best results of the classification model

On the x-axis are the epochs in this case the training took place in 17 epochs, on the y-axis is the numerical value the value of 0.8 represents 80 percent.

The saved weights were from the best result that was obtained in the validation dataset, with a precision value of 0.8242 and a recall value of 0.8533.

In the test dataset we can see Figure 3, it was used with the same weights as the validation training.

Region	Precision	Recall	F1-Score
Restelo	0,7943	0,8267	0,8102
Belas	0,7762	0,8406	0,8071
Cascais	0,8074	0,8058	0,8066
B2B	0,8521	0,8633	0,8577
Birre	0,7829	0,8385	0,8097
Faro	0,7938	0,8176	0,8055

Fig. 3. Table of test dataset results

Similar results were obtained on average with a precision value of 0.8011 and a recall value of 0.8321.

6.2 Differentiation Model Results

For the differentiation model I will also present the best training that corresponds to the best result. We can see in Figure 4, both the x-axis and the y-axis have the same values as in Figure 2

The weights saved were from the best result that had an accuracy of 79,35 percent in the validation dataset.

In the test dataset, which was used with the same validation training weights, obtained an accuracy of 78,92 percent.

7 ANALYSIS

To carry out the analysis, it also made sense to have it divided into a classification model and a differentiation model, as each of them has its own characteristics.

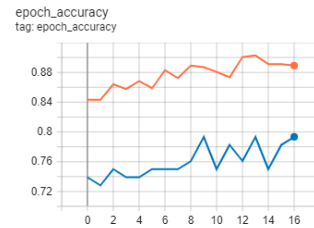

Fig. 4. Graphs of the best results of the Differentiation Model

7.1 Classification Model Analysis

Taking into account the results obtained by EDP present in Appendix A, there was a great improvement in terms of the classification of the Birre and Belas datasets. B2b obtained smaller improvements, but even so. In the other datasets that I added for the test dataset to be more complete, they also obtained good results.

It should be noted that these good results I am referring to here in Portugal because the results of the [5] algorithm in the sei paper show higher values, these higher values can also be due to the quality of the images because the images used in my model was of reduced quality.

7.2 Differentiation Model Analysis

As we can see in the graph of the Differentiation Model referring to accuracy, the training line and the validation line are still a little separated, so it would be advantageous to do more fine-tuning, but the big problem with this model is clearly the very small number of the dataset because although augmentation was used, it is only possible to advance up to a certain point. In order to be able to see the difference that more images make in relation to a better one, since giving 200 images to the training dataset the accuracy of the model increased by 5 percent.

Another thing that has been proven because this speculation already existed, the model is very good at distinguishing solar panels that have a thermosyphon, those without a thermosyphon do not have the best results but many of them can be measured very accurately, which is very good news bearing in mind how difficult it is for a normal person to differentiate.

8 CONCLUSION

In conclusion, given the lack of a comprehensive central database and the disparate levels of knowledge regarding distributed solar photovoltaic energy across Portugal, this study focused on overcoming the difficulties associated with the identification and recording of solar photovoltaic installations.

Utilizing satellite and aerial images to automatically identify and locate solar PV systems across large geographic areas is becoming more and more popular as a solution to these problems. By enabling the recognition and separation of installed solar panels using global satellite images from Portugal, the deep learning models created in this study support this effort.

The classification model's results showed considerable improvements, especially in the Birre and Belas datasets, while the differentiation model showed promise for differentiating solar panels with accuracy. However, it should be emphasized that the caliber of the photos utilized has an impact on how well the models function. The use of satellite and aerial pictures becomes essential in precisely determining the specific locations and capabilities of solar PV projects due to the lack of precise documentation.

The models used in this study represent a development in automatic identification and differentiation, helping to achieve the overall objective of creating a thorough understanding of the deployment of solar PV on a bigger scale. Future studies should continue to investigate and improve these methods while taking into account the distinctive conditions and changes in the deployment of solar PV across various locations. In the end, the results of this study increase the tracking and monitoring of solar PV installations, which supports broader efforts to promote sustainable energy systems.

REFERENCES

[1] Roberto Castello et al. "Deep learning in the built environment: Automatic detection of rooftop solar panels using Convolutional Neural Networks". In: *Journal of Physics: Conference Series*. Vol. 1343. 1. IOP Publishing. 2019, p. 012034.

[2] Sociedade Portuguesa de Energia Solar. *Em Portugal qual deve ser a orientação e inclinação dos paineis solares*. url: https://energiasolare.blogs.sapo.pt/80383.html.

[3] IRENA. *FUTURE OF SOLAR PHOTOVOLTAIC*. url: https://irena.org/-/media/Files/IRENA/Agency/Publication/2019/Nov/IRENA_Future_of_Solar_PV_2019.pdf.

[4] *Kwart van de zonnepanelen niet in beeld*. url: https://www.stedin.net/over-stedin/pers-en-media/persberichten/kwart-van-de-zonnepanelen-niet-in-beeld.

[5] Kevin Mayer et al. "DeepSolar for Germany: A deep learning framework for PV system mapping from aerial imagery". In: *2020 International Conference on Smart Energy Systems and Technologies (SEST)*. IEEE. 2020, pp. 1–6.

[6] Connor Shorten and Taghi M Khoshgoftaar. "A survey on image data augmentation for deep learning". In: *Journal of big data* 6.1 (2019), pp. 1–48.

[7] Christian Szegedy et al. "Inception-v4, inception-resnet and the impact of residual connections on learning". In: *Thirty-first AAAI conference on artificial intelligence*. 2017.

[8] Mingxing Tan and Quoc V. Le. "EfficientNetV2: Smaller Models and Faster Training". In: *CoRR* abs/2104.00298 (2021). arXiv: 2104.00298. url: https://arxiv.org/abs/2104.00298.

[9] Jiafan Yu et al. "DeepSolar: A machine learning framework to efficiently construct a solar deployment database in the United States". In: *Joule* 2.12 (2018), pp. 2605–2617.

Appendix A: Simulation results using the Pytorch model [5]

Threshold 20, training with images of the original model (Inference with B2B images) Precision: 0.8507 Recall: 0.8531 F1 score: 0.8519

Threshold 20, training with images of the original model (Inferência com imagens de Belas) Precision: 0.7714 Recall: 0.3776 F1 score: 0.5070

Threshold 20, training with images of the original model (Inferência com imagens de Birre) Precision: 0.2993 Recall: 0.6197 F1 score: 0.4036

Threshold 50, training with images of the original model (Inference with B2B images) Precision: 0.8912 Recall: 0.8333 F1 score: 0.8613

Threshold 50, training with images of the original model (Inference with images of Belas) Precision: 0.8033 Recall: 0.3427 F1 score: 0.4804

Threshold 50, training with images of the original model (Inference with images of Birre) Precision: 0.2993 Recall: 0.6197 F1 score: 0.4036

Threshold 20, re-training with Birre images (Inference with images of Belas) Precision: 0.7714 Recall: 0.3775 F1 score: 0.5069

Threshold 50, re-training with Belas images (Inference with images of Birre) Precision: 0.2993 Recall: 0.6197 F1 score: 0.4036

Threshold 20, re-training with Belas+Birre images (Inference with B2B images) Precision: 0.7124 Recall: 0.8955 F1 score: 0.7935