Calculating Shadows with U-Nets for Urban Environments

Dominik Rothschedl ✉ ⌂
dwh GmbH, Vienna, Austria

Franz Welscher ✉ iD
Insitute of Geodesy, Graz University of Technology, Austria

Franziska Hübl ✉ iD
Insitute of Geodesy, Graz University of Technology, Austria

Ivan Majic ✉ iD
Insitute of Geodesy, Graz University of Technology, Austria

Daniele Giannandrea ✉ ⌂
dwh GmbH, Vienna, Austria
Institute of Information Systems Engineering, TU Vienna, Austria

Matthias Wastian ✉ ⌂
dwh GmbH, Vienna, Austria

Johannes Scholz ✉ iD
Insitute of Geodesy, Graz University of Technology, Austria

Niki Popper ✉ ⌂
dwh GmbH, Vienna, Austria
Institute of Information Systems Engineering, TU Vienna, Austria

──── **Abstract** ────

Shadow calculation is an important prerequisite for many urban and environmental analyses such as the assessment of solar energy potential. We propose a neural net approach that can be trained with 3D geographical information and predict the presence and depth of shadows. We adapt a U-Net algorithm traditionally used in biomedical image segmentation and train it on sections of Styria, Austria. Our two-step approach first predicts binary existence of shadows and then estimates the depth of shadows as well. Our results on the case study of Styria, Austria show that the proposed approach can predict in both models shadows with over 80% accuracy which is satisfactory for real-world applications, but still leaves room for improvement.

1 Introduction

The production of renewable energy in urban environments is a crucial contribution to carbon neutrality. This requires the assessment of the solar energy potential that is reflected by the solar radiation on the earth's surface [1]. Of particular interest is the assessment of solar energy potential in urban environments, where almost 50% of the world's population is located. Besides photovoltaic systems mounted on roofs, there is additional potential for

photovoltaic systems on facades. As urban areas are covered by buildings that cast shadows on surrounding buildings, and the production of renewable energy with photovoltaics is influenced by shadows - the calculation of shadows is key to make informed decisions.

Contemporary Geographic Information Systems (GIS) are capable of representing shadows for 3D city models based on some type of surface information. The task of generating shadows is usually performed by strictly geometrical approaches such as GIS shadow calculation models. Such models have high accuracy, but usually come at high computational costs depending on spatial and temporal resolution of the data and calculation [11, 12]. One such model is used as a source of ground truth in this study as well [4, 5].

In order to address these issues, this paper presents an approach to calculate shadows using GeoAI methods. One approach, already using machine-learning libraries (tensor-based techniques) but simply optimizing the data preparation and computation time, was shown by[2]. Their urban test area is also represented by a digital surface model (DSM) with a spatial resolution of $1m$. They provide a proof-of-concept for binary shadow calculation, in contrast to our ML-approach, which is able to predict not only the binary value, but also the depth of the shadow. In detail, we present a method for the calculation of the shadow depth of tiles in Styria using U-Net [13]. With this machine-learning approach to solving this problem of physics, we strive to get results more quickly after a computationally intensive training [16, 17].

The U-Net, as basic structure of our network, was originally developed for segmenting biomedical images and is designed to get by with few training images and to be able to localise high-resolution features. These properties fit well for our shadow segmentation task, because as the shadow calculation depends on the position of the sun, we would have needed a large amount of training data.

2   Methodology

The architecture of a U-Net is made up by a contracting path followed by an expansion path, which is roughly symmetric to the contraction. Each contraction step consists of two convolutions and a subsequent max pooling as well as a doubling of the channel numbers. In every expansion step we have an upsampling, followed by a concatenation with the channels of the same size from the contracting path and two convolutions. The concatenations between contraction and expansion paths are key to allow for better localization of high-resolution features and thus more precise segmentation. On top of the U-Net we include residual connections [8] within the convolution net, concretely we insert identity mappings between every other multi-channel feature map. This further eases the training of our net, allowing us to use a larger number of layers in our net.

For our purposes, a U-Net with 5 up- and downsampling layers and a depth of 256 in the bottleneck layer was implemented, where the input tensor consists of the geographic data for each pixel in a 64×64 tile. More specifically, for each pixel, the input is defined as the elevation information, i.e. *surface height* and the *surface height plus the average height of the objects* in the pixel, the terrain information *slope* and *aspect*, and *sun angle* and *azimuth* at a certain time. The input tensor therefore has the size $64 \times 64 \times 6$. Let's take as an example a tile whose centre has the coordinates $(47.0867407955596, 15.423575649486619)$. The pixel in the upper right corner of the tile has a surface height of 352.3 metres, with the objects on the surface 354.59 metres. The terrain has an aspect of $0.23°$ and a slope of $15.77°$. Assuming that the shadow is to be calculated on 19.02.2022 at 12 noon, the angle of the sun is $39.28°$ and the azimuth is $240.971°$. Hence the input tensor of this pixel is $(352.3, 354.59, 15.77, 0.23, 39.28, 240.97)$.

The output differs between the two models. The binary model calculates whether a pixel is shaded or not, whereas the shadow depth model tries to predict the shadow depth, or more precisely the degree of shading of objects with consideration of the shadow depth in a pixel. The degree of shading is divided into eleven classes, where the first class indicates that 0% of the object surface is shaded, the second class indicates that 10% of the object surface is shaded, the third class indicates that 20% of the object surface is shaded, and so on. This is therefore a multi-class net.

For the training of the networks, training areas in Styria (Austria) were defined and divided into 64×64 tiles, where one pixel corresponds to $1m^2$. While the elevation and terrain information for the input layer were derived from the digital service model, a random day of the year 2022 at 12 o'clock was chosen for each tile to determine the position of the sun, from which the azimuth and the angle of the sun were calculated. A total of 449152 tiles were generated and further augmented, i.e., rotated 90, 180 and 270 degrees to increase the size of training data. This dataset was split into 66% training tiles, whereas the rest serve as validation tiles. For each of these tiles, the ground truth was calculated to train the nets. For this purpose, a QGIS Terrain Shading plugin was used to calculate the shadow depth over the DSM [5]. The next step is to transform the result of the QGIS plugin to make it comparable to the output of the nets. For the binary model, a pixel is not shaded, i.e. it has the value 0, if the shadow depth is zero, otherwise its value is 1. For the multi-class net, the degree of shading of an object p_{shaded}, if there is any shading, is calculated from the depth of shading $d_{shadow} \in \mathbb{R}^-$ and the object height, the difference between surface height $h_{surface} \in \mathbb{R}^+$ and ground level $h_{ground} \in \mathbb{R}^+$.

$$p_{shaded} = \begin{cases} 0 & , d_{shadow} = 0 \text{ and } h_{surface} - h_{ground} = 0 \\ 1 & , d_{shadow} < 0 \text{ and } h_{surface} - h_{ground} = 0 \\ \lfloor \frac{d_{shadow}}{h_{surface} - h_{ground}} \rceil & , \text{else} \end{cases} \tag{1}$$

Hence, $p_{shadow} \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$ which is equivalent to eleven categories.

In each training step, different evaluation metrics were applied to the current results to check whether the neural net is learning. One of those is the Jaccard index [6], also known as intersection over union, and the other one is the Dice score [7]

$$\mathcal{J}(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}, \tag{2}$$

$$\mathcal{C}(A, B) = \frac{2|A \cap B|}{|A| + |B|}, \tag{3}$$

where A and B are any two batches of tiles to be compared. Both scores determine the similarities of sets and are common [10, 15].

3 Results

3.1 Binary Model

As already discussed in Section 2 the binary model predicts whether a pixel is shadowed or not. To measure the learning behaviour of the network, we will use the Jaccard index (IoU), Dice Score and Accuracy, and then perform a threshold analysis.
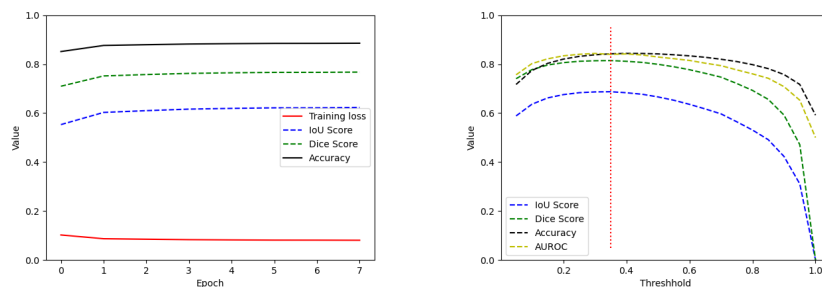
■ **Figure 1** The left figure shows the training metrics of the binary model. Due to early stop algorithm, training was cancelled after 8 epochs. The right figure shows the IoU, Dice, Accuracy and AUROC [3] for validation dataset at different thresholds, where the vertical line indicates the threshold with the best results.
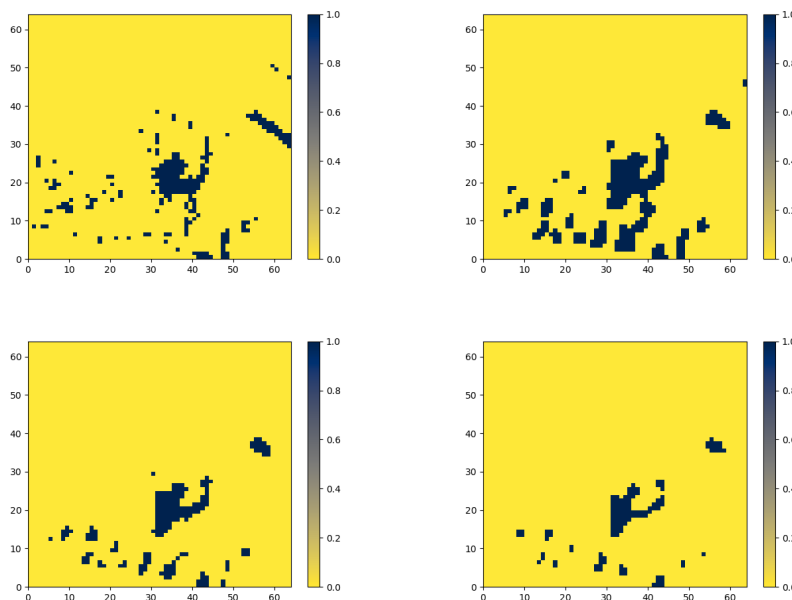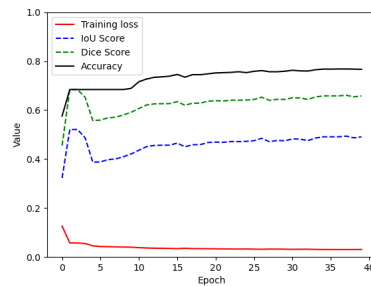

■ **Figure 2** The top left figure shows **the original shadowing** computed with the QGIS plugin. The top right figure shows the **shading predicted** by the binary model **with threshold** 0.2. The bottom left figure shows the **shading predicted** by the binary model **with threshold** 0.35. The bottom right figure shows the **shading predicted** by the binary model **with threshold** 0.5.

The model learning is basically achieved in the first 4 epochs. This is depicted in the left of Figure 1, and can be explained with the number of tiles used for training and the property of the U-Net to learn quickly with a small data set. One may observe in the right of Figure 1, the best results are obtained when the threshold is chosen at 0.35. If the choice is too low, the transitions can also be predicted as shaded. It is also remarkable that in this particular study the MSE-Loss performs best. In the literature[14][9], the cross-entropy loss is mostly used.

3.2   Shadow Depth Model

This model is more complex than the binary model, due to the nature of a multi-class net. Generally, it is expected that premature termination will not occur in this case, which can be also seen in Figure 3. Figure 4 shows that this net accurately calculates the shaded/non-shaded areas, but the transition between them is not sharp as in reality, which is why the result looks blurred. This can be remedied by an additional algorithm that sharpens the results.

Figure 3 The figure shows the training metrics of the shadow depth model for 40 epochs. Unlike the binary model, the accuracy increases over the epochs so that the training was not terminated earlier.
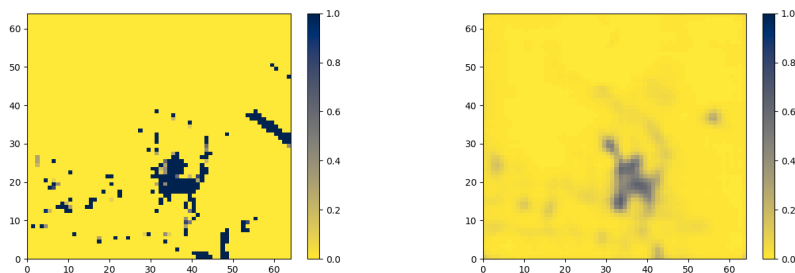
Figure 4 The left figure shows **the original shadowing** computed with the QGIS plugin. The right figure shows the **shading predicted by the shadow depth model**.

4   Conclusion

In this paper, the calculation of shading by a U-Net with residual layers was discussed and trained using selected test areas in Styria. As the results have shown, satisfactory values for the metrics, especially for accuracy, were obtained for both the binary net and the shadow depth net. This is an improvement over the current state of the art because there are currently no approaches that can predict the non-binary depth of the shadow. However, there are still a number of questions that are still open and are in need of further investigation. For example, the nets perform best with MSE loss as the training loss. However, the present state of affairs provides satisfactory results that may serve for further studies. Another aspect that needs to be further investigated and developed is the calculation time. Contrary to the literature, the approach shown is a factor of ten slower in the calculation of 10000 tiles than the QGIS plug-in with the traditional method.

In summary, with the approach of a U-Net as a basis for calculating the shadow depth, a suitable basis for further developments could be created.

References

1 Athanasios Angelis-Dimakis, Markus Biberacher, Javier Dominguez, Giulia Fiorese, Sabine Gadocha, Edgard Gnansounou, Giorgio Guariso, Avraam Kartalidis, Luis Panichelli, Irene Pinedo, et al. Methods and tools to evaluate the availability of renewable energy sources. *Renewable and sustainable energy reviews*, 15(2):1182–1200, 2011.

2 Sukriti Bhattacharya, Christian Braun, and Ulrich Leopold. An Efficient 2.5D Shadow Detection Algorithm for Urban Planning and Design Using a Tensor Based Approach. *ISPRS International Journal of Geo-Information*, 10(9):583, September 2021. `doi:10.3390/ijgi10090583`.

3 Andrew P Bradley. The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern recognition*, 30(7):1145–1159, 1997.

4 Zoran Cuckovic. Enhancing terrain cartography with natural shadows, 2019. URL: `https://landscapearchaeology.org/2019/qgis-shadows/`.

5 Zoran Cuckovic. Terrain shading: a qgis plugin for modelling natural illumination over digital terrain models, 2021. URL: `https://github.com/zoran-cuckovic/QGIS-terrain-shading`.

6 Luciano da F. Costa. Further generalizations of the jaccard index. *CoRR*, abs/2110.09619, 2021. `arXiv:2110.09619`.

7 Lucas Fidon, Wenqi Li, Luis C Garcia-Peraza-Herrera, Jinendra Ekanayake, Neil Kitchen, Sébastien Ourselin, and Tom Vercauteren. Generalised wasserstein dice score for imbalanced multi-class segmentation using holistic convolutional networks. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, 2017, Revised Selected Papers 3*, pages 64–76. Springer, 2018.

8 Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015. `arXiv:1512.03385`.

9 Fabian Isensee, Jens Petersen, Andre Klein, David Zimmerer, Paul F. Jaeger, Simon Kohl, Jakob Wasserthal, Gregor Koehler, Tobias Norajitra, Sebastian Wirkert, and Klaus H. Maier-Hein. nnu-net: Self-adapting framework for u-net-based medical image segmentation, 2018. `arXiv:1809.10486`.

10 Dominik Müller, Iñaki Soto-Rey, and Frank Kramer. Towards a guideline for evaluation metrics in medical image segmentation. *BMC Research Notes*, 15(1):1–8, 2022.

11 Jesús Polo, Nuria Martín-Chivelet, and Carlos Sanz-Saiz. BIPV Modeling with Artificial Neural Networks: Towards a BIPV Digital Twin. *Energies*, 15(11), 2022. `doi:10.3390/en15114173`.

12 P. Redweik, C. Catita, and M. Brito. Solar energy potential on roofs and facades in an urban landscape. *Solar Energy*, 97:332–341, 2013. `doi:10.1016/j.solener.2013.08.036`.

13 Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015. `arXiv:1505.04597`.

14 Nahian Siddique, Sidike Paheding, Colin P. Elkin, and Vijay Devabhaktuni. U-net and its variants for medical image segmentation: A review of theory and applications. *IEEE Access*, 9:82031–82057, 2021. `doi:10.1109/ACCESS.2021.3086020`.

15 Abdel Aziz Taha and Allan Hanbury. Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool. *BMC medical imaging*, 15(1):1–28, 2015.

16 Sander van der Hoog. Deep learning in (and of) agent-based models: A prospectus, 2017. `arXiv:1706.06302`.

17 Yuan Yin, Vincent Le Guen, Jeremie Dona, Emmanuel de Bezenac, Ibrahim Ayed, Nicolas Thome, and Patrick Gallinari. Augmenting physical models with deep networks for complex dynamics forecasting. *Journal of Statistical Mechanics: Theory and Experiment*, 2021(12):124012, December 2021. `doi:10.1088/1742-5468/ac3ae5`.