ZURICH UNIVERSITY OF APPLIED SCIENCES

DEPARTMENT OF MANAGEMENT AND LAW

INSTITUTE FOR INFORMATION TECHNOLOGY

Master Thesis:

Assessing placement efficiency of photovoltaic installations using Mask R-CNN

Tim Claude Bänziger

Primary supervisor: Dr. Elena Gavagnin

Secondary supervisor: Dr. Mario Gellrich

Management Summary

Photovoltaic (PV) energy production has experienced strong growth over the past years and is forecasted to greatly contribute to the successful transition to renewable energy production as demanded by Switzerland's Energy Strategy 2050. Several studies attempted to estimate the national PV potential on building rooftops but arrived at strongly varying results ranging from 15 to 53 TWh annually. To a vast extent, the differences can be explained by the application of varying rooftop utilization ratios which were extrapolated by all previous studies. Moreover, no comparison of the placement of existing PV installations to the suitability categorization from the sonnendach.ch project was yet carried out. Therefore, the aim of this master thesis was to develop and evaluate a prototype methodology to close the research gaps regarding rooftop utilization ratio and the efficiency of PV panel placement. The prototype methodology to answer these questions was developed in Python and leverages publicly available data from the Swiss government in conjunction with a Mask R-CNN for the accurate segmentation of PV panels on high resolution aerial imagery. A total of 1130 individual images of building rooftop were thereby collected in the canton of Aargau of which 974 were used to train the Mask R-CNN model. After four training iterations with varying dataset sizes, the segmentation performance of the Mask R-CNN achieved an iou_score of 0.74. Overall, the rooftop utilization ratio found in this thesis equated to 29%, suggesting that all PV potential studies systematically overestimate the extent of rooftop utilization. Moreover, the findings of this thesis suggest that the more suitable a rooftop area is, the greater its extent of utilization whereas previous studies assumed a uniform distribution of utilization ratio across all suitability categorizations. From the assessed building rooftops, 2.8% have their PV panels suboptimally placed and therefore fail to efficiently exploit solar radiation. 71% of which were successfully detected by the model. Overall, the findings of this thesis proved that an automated, large-scale assessment of PV placement efficiency is technically feasible. This information could support national energy planning as well as PV incentive decision making. However, the segmentation performance of the Mask R-CNN achieved with the resources available to this thesis is currently insufficient for detailed quantitative analyses. Consequently, further studies to improve the Mask R-CNN performance should be conducted before applying the prototype methodology on a large scale.

Table of contents

List of figures

List of tables

List of equations

List of abbreviations

.csv	-	Comma Separated Value file
.geojson	-	Geographical JavaScript Object Notation file
.gpkg	-	GeoPackage file
.jpg	-	Joint Photographic Group file
.tif	-	Tagged Image File Format file
CAGR	-	Continuous Annual Growth Rate
CNN	-	Convolutional Neural Network
EPA	-	ElectricityProductionPlants dataset
FCN	-	Fully Convolutional Network
FPN	-	Feature Pyramid Network
GIS	-	Geographic Information System
GPU	-	Graphical Processing Unit
gt_mask	-	Ground Truth Mask
GWh	-	Gigawatt Hours
GWR	-	Federal Register of Buildings and Dwellings
HTML	-	Hypertext Markup Language
HVAC	-	Heating, Ventilation and Air Conditioning
KWh	-	Kilowatt Hours
MWh	-	Megawatt Hours
pred_mask	-	Prediction Mask
PV	-	Photovoltaic
RAM	-	Random-Access Memory
R-CNN	-	Regional Based Convolutional Neural Network
Mask R-CNN	-	Mask Region Based Convolutional Neural Network
RGB	-	Red, Green, Blue
ROI	-	Region of Interest
RPN	-	Region Proposal Network
RQ	-	Research Question
TWh	-	Terawatt Hours
UAV	-	Unmanned Aerial Vehicles

1. Introduction

1.1. Background

Due to the decreases in the cost of PV cells, solar energy generation has experienced substantial growth in recent years. As such, PV installations have become an increasingly vital cornerstone to meeting the carbon emission goals in an effort reduce global climate change (Gupta et al., 2021). Despite being less radical compared to neighboring countries such as Germany, Switzerland set ambitious targets to reduce CO_2 emissions. In early 2017, Switzerland revised its energy law to subsidize renewable energies along with its Energy Strategy 2050 (Assouline et al., 2018). The goal is to compensate for the decommissioning of nuclear power, planned by 2035, with renewable energy and an overall reduction of CO_2 emissions of 50 to 80% by 2050 (UEVK, 2017). Over 40% of Switzerland's current overall energy demand results from the heating, ventilation, and air conditioning (HVAC) of buildings (Bundesamt für Energie, 2021b). Also, roughly 32% of the national electricity demand is caused by HVAC and lighting. To achieve the goals set by the Energy Strategy 2050, it is therefore critical that buildings become more energy-efficient and the remaining demand is satisfied by renewable sources (Assouline et al., 2018). PV installations on building rooftops have shown to be a reliable large-scale energy source for urban areas that can significantly contribute to decentralized renewable energy supply. In addition, PV installations require little to no changes in building infrastructure, therefore, being seen as an optimally suited solution to address urban electrification and boost renewable energy supply on a local level (Lonergan & Sansavini, 2022). Subsequently, PV installations have the potential to drastically reduce CO_2 emissions and aid in achieving renewable energy goals (Walch et al., 2019)

Despite the promising technology and strong growth with a CAGR of 39% between 2010 and 2020, photovoltaic energy only accounted for 4.7% of Switzerland's total energy production in 2020 (Bundesamt für Energie, 2021c). While 73% of new PV installations in 2020 were built on single-family houses, only 7% were built on industrial buildings. Due to the larger rooftop area available, the total electrical output of new PV installations on industrial buildings exceeds the total output of single-family houses with 173 and 138 MWh, respectively. (Bundesamt für Energie, 2021c).

Figure 1 shows the growth in total solar electricity production in Switzerland since the turn of the millennium (Bundesamt für Energie, 2021c).
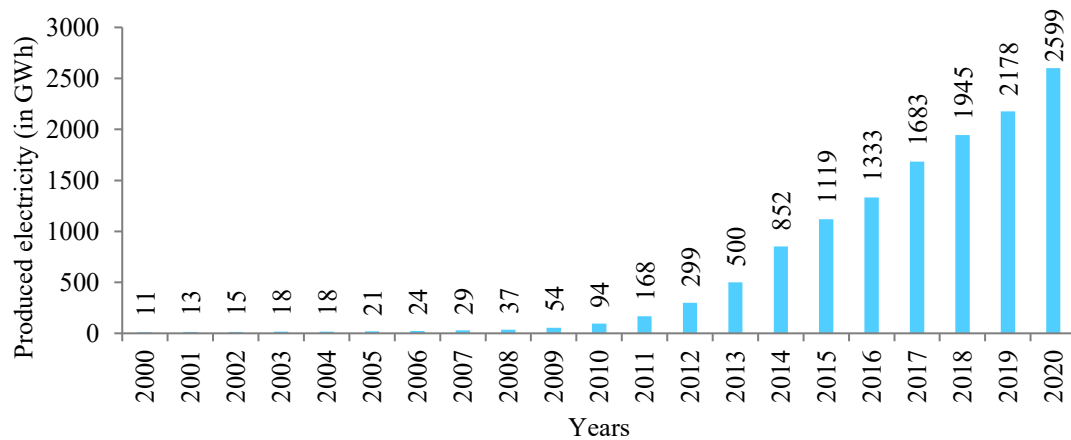
Figure 1: Solar electricity production growth in Switzerland (in GWh) between 2000 and 2019 (Bundesamt für Energie, 2021b)
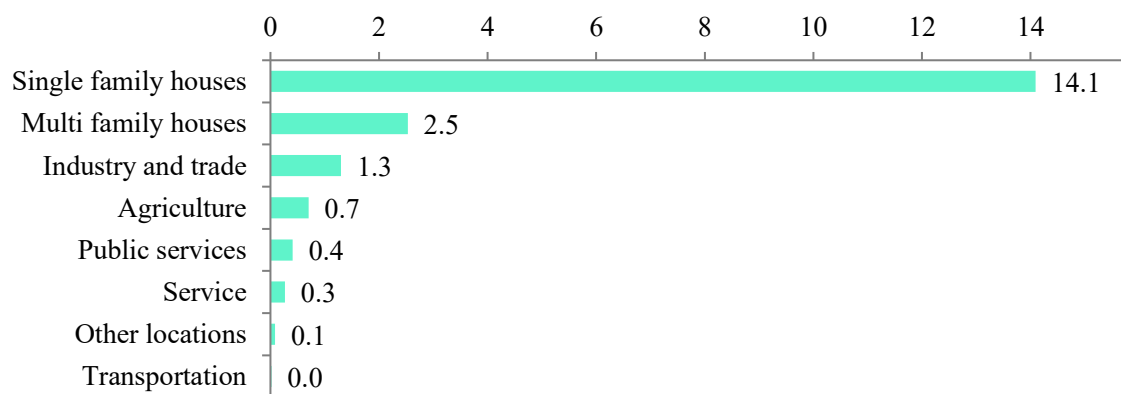
Figure 2: Number of newly built PV installations among different categories in 2020 (in thousands) (Bundesamt für Energie, 2021b)

These developments have raised the question of how large the potential solar energy production in Switzerland could be, should all available building rooftop areas be maximally utilized for PV energy production. Several studies tried to answer this question, yet the estimates for the PV potential in Switzerland show greatly varying results ranging between 15 to 53 TWh. These discrepancies can primarily be traced back to strong variation in estimated available building rooftop areas (Walch et al., 2019). Moreover, the estimated rooftop utilization ratio which was extrapolated by all PV potential studies also varies greatly between the studies. No approach to determine the extent to which PV installation owners exploit the available rooftop area on a large scale has yet been developed.

Additionally, no granular information regarding the placement of PV panels within the rooftop area is currently available. This information would allow examining whether the PV panels are placed in the most optimally suited region of a given rooftop (placement efficiency). Thereby, two research gaps could be identified: (1) the available rooftop area utilization ratio on a large scale and (2) the placement efficiency of current PV installations.

1.2. Problem definition

As aforementioned in the Background section, the overall PV potential of Switzerland's rooftop has been analyzed to a great extent. However, PV installation owners often do not fully utilize the entire available rooftop area in practice. As previous PV potential studies solely estimate the rooftop utilization ratio to account for partly used rooftop surfaces, they likely do not display a realistic picture. The negligence of this real-life factor might ultimately lead to problems in grid distribution or energy planning and thus threaten the achievement of the goals set by the Energy Strategy 2050. As part of the sonnendach.ch project, the Federal Office for Energy developed a model to identify the PV potential for every building rooftop in Switzerland. The model takes a variety of variables such as, inclination, orientation, and shading into account and calculates the potential PV energy output for all plane surfaces on building rooftops (Portmann et al., 2016). This information should help PV owners to find the most efficient placement for their PV panels. Despite a large amount of high-quality data available from the Swiss government, existing PV panels are solely mapped to individual addresses. No granular mapping with the exact shape and position (segmentation) on the rooftops has been conducted. Comparing the actual placement of existing PV panels to the optimal placement according to the sonndach.ch model would allow drawing conclusions on whether PV installation owners have placed their PV panels correctly to maximally exploit the solar radiation available on their rooftop.

1.3. Objective

The objective of this master thesis is to develop and evaluate a prototype methodology that leverages publicly available data from the Swiss government to close the research gaps identified in the Background section regarding the extent of rooftop surface area exploitation (rooftop utilization ratio) and PV panel placement efficiency.

Therefore, the prototype model must be able to automatically identify buildings with PV installations, segment the PV panels precisely and assess whether the PV panels of a given rooftop are placed correctly to exploit the available solar radiation maximally. The optimal PV panel placement is thereby determined by the categorization of the sonnendach.ch project (Portmann et al., 2016). The prototype shall lay the foundation for a model able to conduct a country-wide assessment and serve as a proof of concept that can be utilized by energy companies for commercial reasons.

1.4. Research question

The research question (RQ) of this thesis is divided into three Sub-RQs which subsequently resulted in the three Tasks structure found in the Methodology section:

- <u>RQ</u>: Can a large-scale prototype methodology be developed that detects and compares the placement of PV panels with the rooftop area suitability categorization according to the sonnendach.ch project?
    - o <u>Sub-RQ1</u>: Can a prototype method be developed that allows to automatically crop high resolution aerial imagery to the shape of individual buildings?
    - o <u>Sub-RQ2</u>: Can a Mask R-CNN be trained to accurately segment PV panels on rooftops?
    - o <u>Sub-RQ3</u>: Can the masks from the Mask R-CNN be compared to the categorization of the sonnendach.ch project?

1.5. Scope limitations

The development of this prototype methodology is limited to a narrowly bounded region. While the methodology was developed to work for the entirety of Switzerland, the data collected and processed is limited to a few regions within the canton of Aargau. As this thesis has a proof-of-concept nature focusing on the functionality rather than outcome, certain parts of the methodology have potential for further optimization yet not conducted. Examples thereof are research on the tuning of Mask R-CNN hyperparameters (see 3.2.2) or the reduction of data noise in the PV placement assessment (see 3.3.2). Moreover, the assessment is limited to the suitability categorization and does not take into account other variables form the sonnendach.ch project. These scope limitations are translated into recommendations for further research in the Discussion section.

1.6. Contribution and relevance

The primary relevance of this thesis is twofold. Firstly, it contributes to the existing body of academic knowledge by developing a prototype methodology to exactly measure the rooftop utilization ratio on a large scale. Thus, eliminating the need to rely on extrapolated rooftop utilization ratio as employed by previous PV potential studies. The thesis thereby contributes by delivering techniques and data needed to improve the accuracy of further potential estimates that are crucial to electrical planning and achieving the Energy Strategy 2050. Moreover, an accurate rooftop utilization ratio that can be exactly determined for any arbitrary region also aids in understanding the specific reasons why less than the available area is exploited. This information could help policy makers to strengthen inducements for PV owners. Secondly, the prototype methodology of this thesis allows catering to a range of commercial use cases related to the solar industry. By identifying the placement effectiveness and the extent to which a given address is exploiting the available area, (after-) sales initiatives can be launched. A possible use case could be that PV installer companies identify buildings that have suboptimally placed PV panels with the aim of selling their services.

1.7. Structure of the thesis

This thesis is divided into following structure: In the first two chapters, the state of research is shown on the basis of related work and the theoretical foundations needed for the development of the prototype methodology are discussed. In the following chapter, the developed methodology is explained in detail and respective Python code is linked. All data used as part of the prototype methodology is detailed following the Methodology chapter. The achieved outcomes are then presented in the Results chapter. Like the RQs, the Methodology as well as the Results chapter are divided into the three main Tasks: data collection and preprocessing, Mask R-CNN for PV segmentation, and lastly, the comparison to SolkatDach suitability categorization. Finally, the results are taken up again and brought into context in the Discussion chapter and condensed in the Conclusion.

2. Literature review

2.1. PV potential estimates for Switzerland

As mentioned in the Background section, the eight studies that estimated the total electrical PV potential in Switzerland show strongly varying results: In an early study that was conducted by the International Energy Agency (IEA) in 2002, the PV potential of several European countries was analyzed by roughly calculating the available rooftop area derived from the building ground floor area. This rather rudimentary method estimated the total PV potential of Switzerland to be roughly 15 TWh (International Energy Agency (IEA), 2002). Years later in 2016, Switzerland's Federal Office of Energy, Swisstopo, and MeteoSwiss funded the large-scale project sonnendach.ch. The project uses a geographic information system (GIS) based method jointly with rule of thumb approaches to estimate the PV suitability of building rooftops (Portmann et al., 2016). With relatively high a rooftop utilization ratio of 62%, the total PV potential of Switzerland was estimated to be 53 TWh (Walch et al., 2019) which is the highest estimate of all potential studies. The first analysis based on machine learning was carried out by EPFL in 2017. In this study GIS, machine learning algorithms, and support vector machines were used. The study conducted on a communal level estimated the countrywide PV potential to be 18 TWh (Assouline et al., 2017). Shortly after, Assouline et al. (2018) published another paper intending to improve the initial study of 2017. Random forest algorithms in combination with more accurate datasets were used. With an estimated PV potential of 16 TWh, the results are slightly below the initial study from 2017 (Assouline et al., 2018). Buffat et al. (2018) conducted another PV potential analysis with the primary goal to compare different PV technologies. The model, which also accounted for panel and inverter inefficiencies, estimated Switzerland's PV potential to 41 TWh. Once the data generated by Portmann et al. (2016) was available for the entirety of Switzerland the association Swisssolar conducted another study. The comparatively new paper estimated the total PV potential of Switzerland to be approximately 49 TWh of which 23 TWh can be used in the short term (Remund et al., 2019). Through a very detailed approach using GIS and machine learning methods, Walch et al. (2020) conducted yet another potential estimate. Based on the extrapolated estimates for of approximately 9.6 million rooftops, the study estimated the nationwide PV potential to be 24 (±9) TWh.

The most recent study was published by the ZHAW Institute of Natural Resource Sciences in early 2021 which focuses to accurately estimate the rooftop utilization ratio. By manually assessing 99 rooftops in the Zurich area, the study estimated rooftop utilization ratio to be 50% which resulted in a countrywide PV potential of 39 TWh (Moro et al., 2021). As the rooftop utilization ratios and thus total potential estimates vary significantly between the studies, it becomes evident that no consensus has yet been reached. Table 1 provides an overview summary of the PV potential studies conducted in Switzerland.

Table 1: Comparison of Swiss PV potential estimates

Study	Available rooftop area (in km^2)	Rooftop utilization ratio (in %)	Total potential (in TWh)
International Energy Agency (2002)	251	55.0	15.0
Portmann et al., (2016)	314	61.9	53.1
Assouline et al., (2017)	328	60.5	17.9
Assouline et al., (2018)	252	60.5	16.3
Buffat et al., (2018)	485	70.1	41.3
Remund et al., (2019)	252	50.0	49.1
Walch et al., (2020)	267	56.4	24.6
Moro et al., (2021)	231	50.0	38.8

2.2. Detection and segmentation of PV installations

Several studies investigating the feasibility of detecting PV panels on building rooftops using machine and deep learning indicate promising results (Golovko et al., 2017; Li et al., 2020; Malof et al., 2016). However, PV detection is no topic of further analysis of this thesis as the existing PV installations in Switzerland are already mapped by the Federal Office of energy and the data is publicly available. Therefore, no detection is needed and solely the precise delineation of PV panels as polygons through image segmentation is of interest. The following section focuses on research regarding the segmentation of PV panels though machine learning approaches.

In an early study conducted by Malof et al. (2016), images from satellites with a resolution of 0.3 m were used to detect PV panels. However, the study also examined to what extent the model can segment PV sizes and shapes with an object-based approach using a random forest classifier. With a training set of approx. 2.7 thousand manually annotated PV panels, no sufficient results could be established. Another detection study that also assessed the segmentation of PV panels was conducted by Li et al. (2020).

As part of the SolarFinder project, the authors leveraged a k-means algorithm to segment individual rooftop images into objects and contours followed by a hybrid approach using support vector machine radial basis function and a convolutional neural network (CNN) model. Through this sophisticated model supported by a large-scale data set, the characteristics (size and shape) of PV panels could be accurately profiled. In a very recent study focusing on the extraction of PV panel characteristics, Souffer et al. (2022) used RGB as well as thermal imagery from unmanned aerial vehicles (UAV) combined with a random forest classifier to segment PV panels. Through the comparatively high data quality, a strong result with an average F1-score of 0.99 was achieved. By solely depending on thermal images and using a different approach, Wang et al. (2021) achieved a slightly lower F1-score of 0.86. He and Zhang (2020) developed a model to detect and segment PV panels using an R-CNN architecture (see 2.5.1). A training set of almost 6 thousand rooftop images with a resolution of 0.3 m was used. The study conducted in the urban region of Manchester, UK claims to successfully delineate PV panels but provides no proof of performance for image segmentation tasks. Jiang et al. (2021) conducted a study using three different types of areal footage. As such, satellite images with a spatial resolution of 0.8 m and 0.3 m alongside high-resolution images (0.1 m) taken by UAVs were analyzed. The three datasets were then tested using the U-Net, RefineNet, and DeepLab v3+ algorithms. The Authors indicate that for PV segmentation on rooftops, the footage with a resolution of 0.1 m is most suited. The obtained F1 scores ranged from 0.86 (U-Net) to 0.91 (DeepLab v3+) (Jiang et al., 2021). The training data of this study is freely available but will be of limited use as the buildings and surroundings in Switzerland vary from China.

It must be noted that a direct comparison of the above studies is not feasible due to the large differences in data as well as algorithm choice and architecture. Nevertheless, it can be concluded that the field of PV panel detection was subject to extensive research while PV panel segmentation is mostly analyzed as a secondary by-product, not included in the main research focus. However, a study that focuses on segmenting PV panels from satellite imagery was published by Liang et al. (2020). The authors utilized the satellite imagery dataset provided as part of the DeepSolar project of Stanford University and manually annotated 2733 samples. The data was used to fine-tune a pre-trained Mask R-CNN that was trained on the COCO dataset.

The overall performance of the Mask R-CNN was satisfactory, but the model had difficulties to accurately segment the edges of the rectangular shaped PV panels. This problem was eliminated through the application of a right-angle polygon fit algorithm. An in-dept literature review of Mask R-CNNs can be found in section 2.5. Also relying on a Mask R-CNN algorithm, researchers of the Institute of Geomatics at the University of Applied Sciences of Northwestern Switzerland conducted a study with very similar objectives as this thesis. The authors manually annotated 7'839 image tiles with a total 31'401 polygons representing roughly 22'000 PV panels (averaging 20 images/ hour). Through the extensive amount of ground truth data, a relatively high performance was achieved (no comparable metrics published). The computational load associated with this amount of data, was distributed on four graphical processing units (GPU) (Nvidia Tesla V100 SXM2) supported by a HPE Apollo 6500 Deep Learning GPU server. With this hardware at hand, the authors estimated the total inferencing time for the entirety of Switzerland to be approximately ten days (Meyer, 2020). No other information regarding the results is publicly available. Because of this and the study's similarity in terms of objectives and methodology, the authors were contacted multiple times in hopes for a knowledge exchange, but no answers have been received.

2.3. Image Segmentation

As mentioned before, the intent of this master thesis is not to detect PV panels but to determine their exact size and placement on the rooftops, so their placement can be compared to the suitability categorization from the sonnendach.ch project. This problem can be solved using image segmentation which is defined as a set of specific image processing techniques applied to divide an image into one or more meaningful regions. In other words, image segmentation defines separate semantic entities within an image. On a more technical level, image segmentation is a sub-class of computer vision that assigns a label to each pixel with respect to its semantic meaning (Ghosh et al., 2020). Despite being less well researched as the field of object detection, image segmentation is used in multiple real-world applications such as traffic surveillance (autonomous driving), medical image processing (e.g., localization of abnormalities such as melanomas or aneurysms), defense guidance systems or forensics (Ghosh et al., 2020).

2.4. Deep learning techniques for image segmentation

While several traditional approaches for image segmentation have been available for years, only with the surfacing of deep learning techniques fundamentally new segmentation algorithms evolved. Through the application of CNNs, that learn through back-propagation of convolutional kernels, the segmentation performance was greatly improved and, thus, new applications could be served (Ghosh et al., 2020). The progress in performance was supported, to a large extent, by powerful baseline systems such as the Fast or Faster Regional-CNN and Fully Convolutional Network (FCN) (He et al., 2017). In image segmentation, a convolution can be defined as a function that calculates a sum of product between input values and kernel weights while convoluting a smaller kernel over a larger image. The convolutional kernels of a conventional CNN thereby tend to generate activation maps with regards to specific properties of the objects. Because of the way activations work, they can be thought of as segmentation masks for object-specific features. Therefore, these output activation matrices already include the key to generating requirement-specific segmentation. The majority of image segmentation algorithms exploit this CNN capability to build the segmentation masks needed to solve the problem (Ghosh et al., 2020). Appendix 10.1 summarizes major deep learning-based segmentation algorithms in chronological order.

2.5. Mask R-CNN

2.5.1   Origins

To understand functionality and architecture of the Mask R-CNN, it is pivotal to understand its origins which date back to the advent of CNNs. Back then, the question whether networks like AlexNet can be re-designed to detect the presence of multiple objects arose. The answer to this question was found in the development of R-CNNs (Ghosh et al., 2020). First introduced by Ross Girshick, R-CNNs utilize a selective search algorithm to extract region proposals that are then passed though the convolutional network before a category is assigned for a given image crop. As one forward pass for every region proposal (can be hundreds or thousands per image) is required, R-CNNs suffer from high computational costs. Moreover, as these region proposals vary in size and shape, achieving uniform-size features is impossible (Ghosh et al., 2020). With the introduction of the Fast R-CNN, Girshick addressed the downsides of traditional R-CNNs through the utilization of region-of-interest (ROI) pooling.

As such, ROIs are pooled dynamically to generate feature outputs of a fixed spatial extent (e.g., 6 x 6), leaving the selective search technique as the sole remaining bottleneck (Girshick, 2015). This approach also significantly improved detection speed. Shortly after inventing Fast R-CNN, Faster-R-CNN, which relies on intermediate activation maps to propose bounding boxes instead of external features, was invented (Ren et al., 2017). Faster R-CNN at its core, is composed of two stages. The first being the region proposal network (RPN) which proposes region object bounding boxes. The second stage, being somewhat identical to Fast R-CNN, extracts features using ROI pooling from each bounding box and performs a classification (He et al., 2017). The Faster R-CNN pathed the way for the development of today's state of the art image segmentation algorithm: Mask R-CNN.

2.5.2   *Functionality and architecture*

Five years ago, He et al. (2017) presented the conceptually simple Mask R-CNN framework that remains the most widely accepted technique for image segmentation to date (Liang et al., 2020). The framework which leverages the Faster R-CNN method, was enhanced with a parallel branch that provided object specific binary classification on a pixel-based level, generating accurate segments (He et al., 2017). The first of the two-stage approach (see 2.5.1) used in Faster R-CNN remains unchanged while the second stage was adapted. In Faster R-CNN, the second stage generates two outputs for each object (bounding-box and class label) while in Mask R-CNN another branch to accommodate the object mask is added. Figure 3 illustrates the architectural evolution from a simple R-CNN to a Mask R-CNN as used in this thesis.
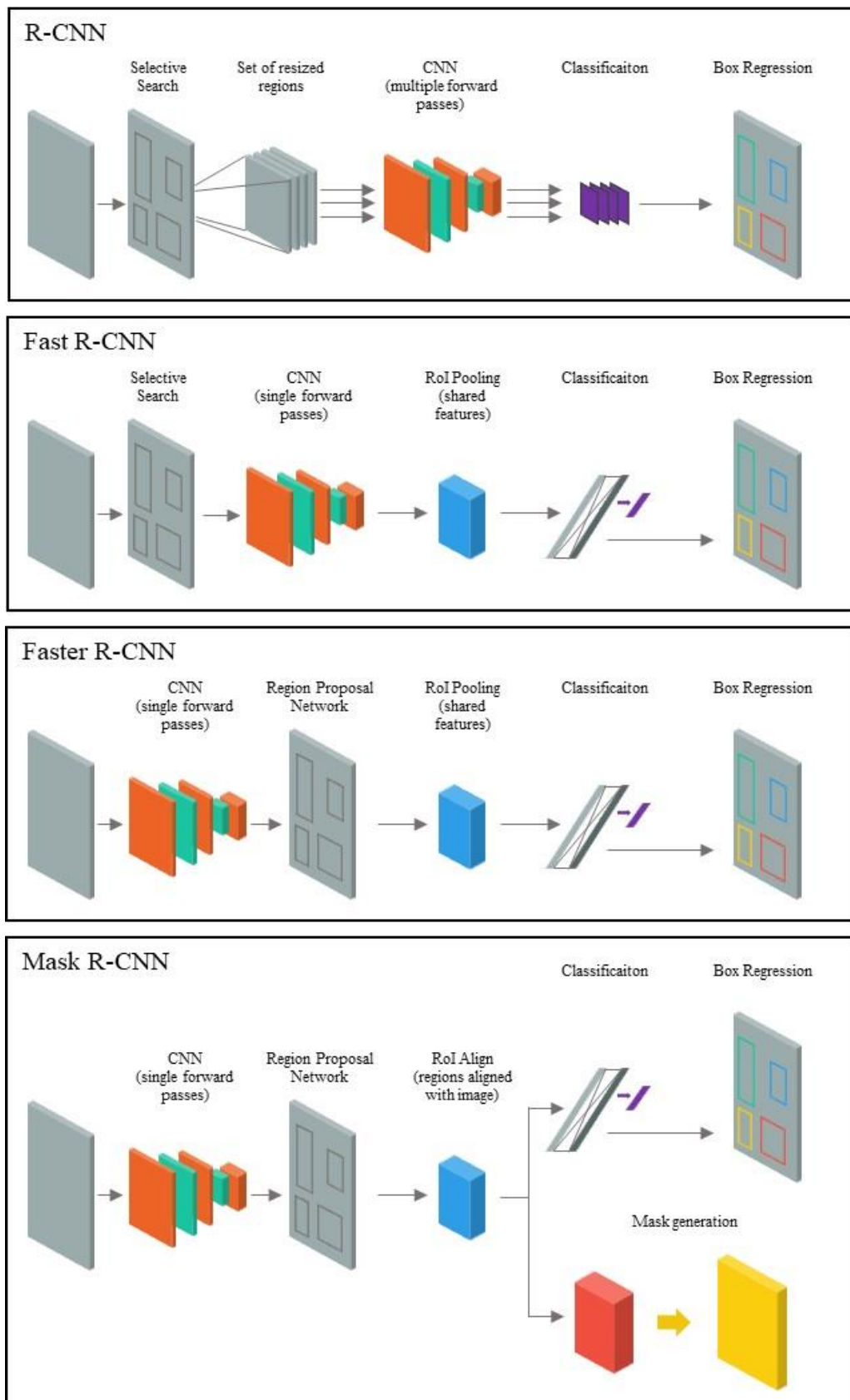
Figure 3: Mask R-CNN architecture evolution illustrated (adapted from Ghosh et al., 2020)

3. Methodology

Following the three-split structure of the Sub-RQs (see 1.4), the Methodology is also divided into three main Tasks that can encompass multiple Subtasks. These Tasks are (1) data collection and processing, (2) training a Mask R-CNN for PV segmentation and finally (3) the comparison to the SolkatDach suitability categorization originating from the sonnendach.ch project.
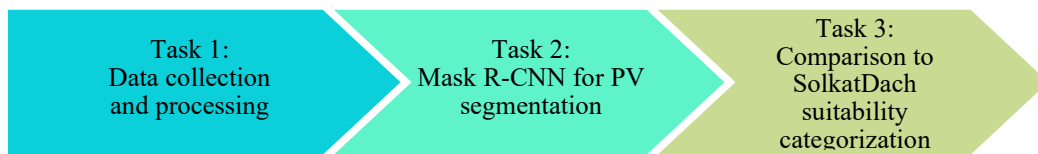
Figure 4: High-level process of the developed prototype methodology

In this section, the three main Tasks as well as the Subtasks are explained in detail. The vast majority of steps in this prototype methodology were performed through Python scripts that were executed in the Google Colab environment directly mounted to google drive where the data was stored. The Python scripts can be accessed through the links to the GitHub repository provided at the beginning of each Subtask. The first two Tasks were both executed over four iterations. This means that for each iteration additional data was collected to train a Mask R-CNN by executing Task 1 again. The data from the first iteration was then included in the dataset for the second iterations and so forth. Unless stated otherwise, a Tesla P100-PCIE-16GB GPU was used for all steps of this master thesis. Figure 5 provides an overview of the code structure and all data required to execute the methodology within the google directory. The entire prototype methodology developed in this thesis is illustrated in Figure 7.

```
\---MA
  +---1_0_data_collection                              -> Folder for task 1: Data collection and processing
  |  +---1_1_obtain_high_res_aerial_imgs               -> Folder for subtask 1.1 & 1.2
  |  |  +---data_col_folder1-n                         -> Folders to save images
  |  |  |   download_and_merge_imgs.ipynb              -> Code for subtask 1.2
  |  |  |   download_links_SWISSIMAGE10CM_2_meta.csv   -> Download links for image tiles
  |  |
  |  \---1_2_save_individual_imgs_of_rooftops_with_pv  -> Folder for subtask 1.3
  |      +---data_col_folder_1-n                       -> Folder to save images
  |      |   EPA_GWR_merged.csv                        -> EPA & GWR combined (see subtask 1.1)
  |      |   image_puncher_v2.ipynb                    -> Code for subtask 1.3
  |      |   solkat_ca_Aargau.geojson                  -> geojson of SolkatDach (region≈Aargau)
  |      |   tif_to_jpeg.ipynb                         -> Code to convert .tifs to .jpgs (see subtask 1.4)
  |
  +---2_0_model_development                            -> Folder for task 2: Mask R-CNN for PV segmentation
  |  \---iteration_1-n                                 -> Folders for each iteration
  |      +---dataset                                   -> Dataset folder
  |      |  +---train                                  -> Train images and annotation masks
  |      |  \---val                                    -> Validation images and annotation masks
  |      \---Mask_RCNN                                 -> Mask R-CNN folder (by Matterport)
  |         |   mask_rcnn_coco.h5                      -> Coco weights
  |         +---logs                                   -> Folder where model logs are saved
  |         |  +---pv_timestamp                        -> Folders where model logs are saved
  |         |  |   |   events.out.tfevents.xyz         -> Tensorflow event files
  |         |  |   |   mask_rcnn_pv_n.h5               -> Model weights saved as .h5
  |         |  |
  |         +---mrcnn                                  -> M-RCNN package(by Matterport)
  |         |  |   config.py                           -> Config module for hyperparameters
  |         |  |   model.py                            -> Model module
  |         |  |   utils.py                            -> Utils model for miscellaneous useful functions
  |         |  |   visualize.py                        -> Visualize model to plot images and graphs
  |         |  |
  |         \---PVseg                                  -> Custom Folder
  |            |   PVseg.py                            -> Module for key functionality (by Matterport)
  |            |   PVseg_data_inspect.ipynb            -> Code to inspect data (by Matterport)
  |            |   PVseg_model_inspect_TCB.ipynb       -> Code to inspect model (adapted from Matterport)
  |            |   PVseg_run.ipynb                     -> Code to set environment and run model
  |            |   PV_seg_weight_instpect.ipynb        -> Code to inspect model weights (by Matterport)
  |            |   Tensorboard.ipynb                   -> Code to inspect model performance via Tensorboard
  |
  \---3_0_assessment                                   -> Folder for task 3: Comparison to SolkatDach suitability categorization
     +---3_1_inferencing_and_saving_masks              -> Folder for subtask 3.1
     |  +---pred_and_gt_masks                          -> Folder to save pred_masks and gt_masks (see subtask 3.1)
     |  |   Save_pred_masks.ipynb                      -> Code for inferencing and saving the prediction (as well as gt) masks
     |
     \---3_2_pv_placement_assessment                   -> Folder for subtask 3.2
        +---pred_and_gt_test                           -> Folder containing pred_masks and gt_masks of test dataset
        +---pred_and_gt_train                          -> Folder containing pred_masks and gt_masks of tain dataset
        +---all_tifs                                   -> Folder containing .tif images of the entire dataset
        |   pv_placement_assessment.ipynb              -> Code for comparing SolkatDach to pred_masks and gt_masks
```
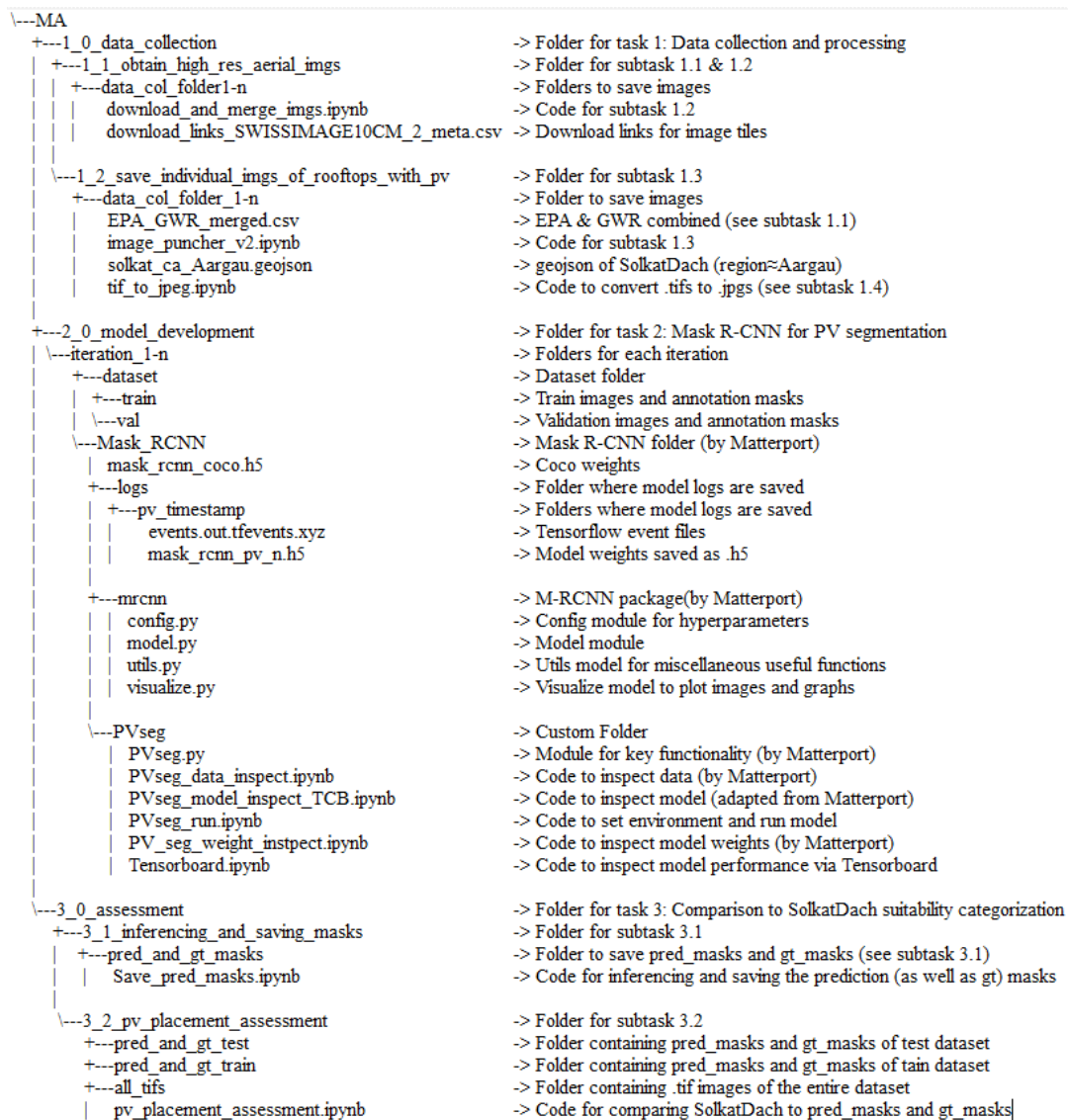
Figure 5: Directory tree of key elements as mounted in Google Colab

3.1. Task 1: Data collection and processing

As with any other machine learning model, a training dataset containing ground truth information is required to train the Mask R-CNN. In this case, the training dataset consists of images of building rooftops that have PV panels installed. Additionally, each image must have a mask that delineates the PV panels precisely. To generate the dataset as described, the images were downloaded and processed using multiple Python scripts and manually annotated. To ensure the actuality of the images, the data was collected within the canton of Aargau for which the most recently taken aerial images are available (see Figure 10). The exact location can be depicted in Appendix 10.2. The four steps required to fulfil this Task, are as follows:

3.1.1 Subtask 1: Identification of rooftops with PV panels

To identify buildings that have PV panels installed, the dataset "ElectricityProductionPlants" (EPA) provided by the Federal Office of Energy was utilized (Bundesamt für Energie, 2021a). The dataset includes valuable information such as the coordinates, beginning of operation, or total power output of existing PV installations. Unfortunately, the EPA dataset lacks a vital information that is required for following Subtasks, the EGID (German: Eidgenössischer Gebäudeidentifikator, short EGID) (see 3.1.3). The EGID is a unique identification number that the Federal Office for Statistics assigns every building within the country which can be found as part of the Federal Register of Buildings and Dwellings (German: Gebäude- und Wohnungsregister, short GWR) dataset (Bundesamt für Statistik, 2021). Both datasets, EPA and GWR were then merged to obtain one single ".csv" file through the spatial position coordinates. The final dataset (EPA_GWR_merged) functions as the filter to determine which rooftops have PV panels installed and should thus be taken into account.

3.1.2 Subtask 2: Obtaining high-resolution aerial imagery

Link to code

Before taking snapshots of individual rooftops with PV panels, high-resolution satellite imagery must be acquired. The digital color orthophoto mosaic of by the Federal Office for Topography (Swisstopo) covers the entirety of Switzerland with aerial images with a resolution of <10 cm (Bundesamt für Landestopografie, 2022). For the determined area, the individual tiles of one km^2 were downloaded then merged to one large photomosaic image. The process of downloading individual orthophotos and merging them to one single ".tif" file was automated through a Python script. The merged photomosaic serves as the base image for collecting multiple individual rooftop images as described in the following Subtask.

3.1.3 Subtask 3: Save individual image of each rooftop with PV panels

Link to code

To generate individual images of building rooftops with PV panels installed, the precise shape and location of each building rooftop must be determined. This information can be derived from the dataset "SolkatDach" which resulted from the sonnendach.ch project (Portmann et al., 2016).

The dataset in form of a ".gpkg" file was then cropped to the dimensions defined in Subtask 2 and saved as a ".geojson" file using the QGIS Desktop application (Sutton & Dassau, 2022). The SolkatDach dataset does not contain one polygon per rooftop but a single polygon for each rooftop surface (one or multiple per building). For a given rooftop, which is identified through the EGID, all individual rooftop surface polygons must thus be merged iteratively. Once a single polygon per rooftop is created, an image can be cropped from the photomosaic obtained in Subtask 2. In order obtain a rectangular sized image, padding is added on the edges surrounding the rooftop structure. The information resulting from Subtask 1 (EGID of rooftops with PV panels) as well as the area dimensions defined in Subtask 2 thereby serve as filters to determine for which rooftops should be executed in Subtask 3. Similarly to the previous Subtask, this process was automated through a Python script.

3.1.4   Subtask 4: Annotate the images individually to obtain image masks

Link to annotator

After collecting the images, it is necessary to create the annotation masks to establish a reliable ground truth. This process is time consuming since it must be undertaken manually. As The Mask R-CNN model solely processes ".jpg" image files, the individual rooftop images were converted from ".tif" to ".jpg" before annotating (link to code). For the annotation, the HTML based image annotation software VGG Image Annotator (VIA) was used. The images were annotated with one or multiple polygons, all belonging to the same class entitled "PV". Once the manual process was completed, the annotations were extracted in the VIA-JSON format. These steps were carried out for the training and validation dataset.

3.2. Task 2: Mask R-CNN for PV segmentation

Link to the original Matterport implementation

Link to the code of the adapted version

To segment PV panels on individual rooftop images the Mask R-CNN architecture was chosen. This decision is supported by two primary reasons: Firstly, the Literature review shows that Mask R-CNN is the is most widely used algorithm in both academia and practice and can thus be considered state-of-the-art for image segmentation tasks. Secondly, the algorithm was successfully used by researchers that conducted studies with very similar objectives (Liang et al., 2020; Meyer, 2020).

The model used for this thesis is a slightly adapted version of the well-established Matterport Mask R-CNN Architecture (Waleed, 2018). Thereby, the model is very similar to the original implementation published by He et al. (2017). Apart from hyperparameters, the adaptations made to the code from Matterport's Mask R-CNN are limited to structural changes and changes in terminology and designations. Subsequently, the model used in this thesis is functionally identical to Matterport's original version. The model was trained multiple times with different datasets. Thereby a total of four iteration (each iteration with a different train dataset size) were trained. The results of the four iterations are analyzed in section 5.2.1.

3.2.1   *Matterport Mask R-CNN architecture*

The first module of the Mask R-CNN architecture implemented by Matterport utilizes a standard CNN backbone (in this case ResNet101) that functions as a feature extractor. While the early layers detect trivial features such as edges and corners, the higher layers detect more complex features. The feature extraction is then improved using a feature pyramid network (FPN) which takes an arbitrarily sized input image and generates proportionally sized feature maps as output (Lin et al., 2017). The second module of the model, a region proposal network (RPN), is a lightweight neural network, that scans the feature maps in a window-like manner for regions, called anchors, that contain target objects. The top anchors (ROIs) are then passed to the next module that generates two outputs per ROI that is passed through: a classification (in this case PV or background) and a bounding box refinement which further specifies the location and shape of the object. Parallel to that, another CNN that takes the ROIs as input and generates one segmentation mask per object (He et al., 2017). Combined, the classifications, bounding boxes, and masks are the final outputs of the model (Waleed, 2018). Figure 6 illustrates the architecture of the Mask R-CNN implementation by Matterport as well as the loss functions detailed in section 3.2.3.
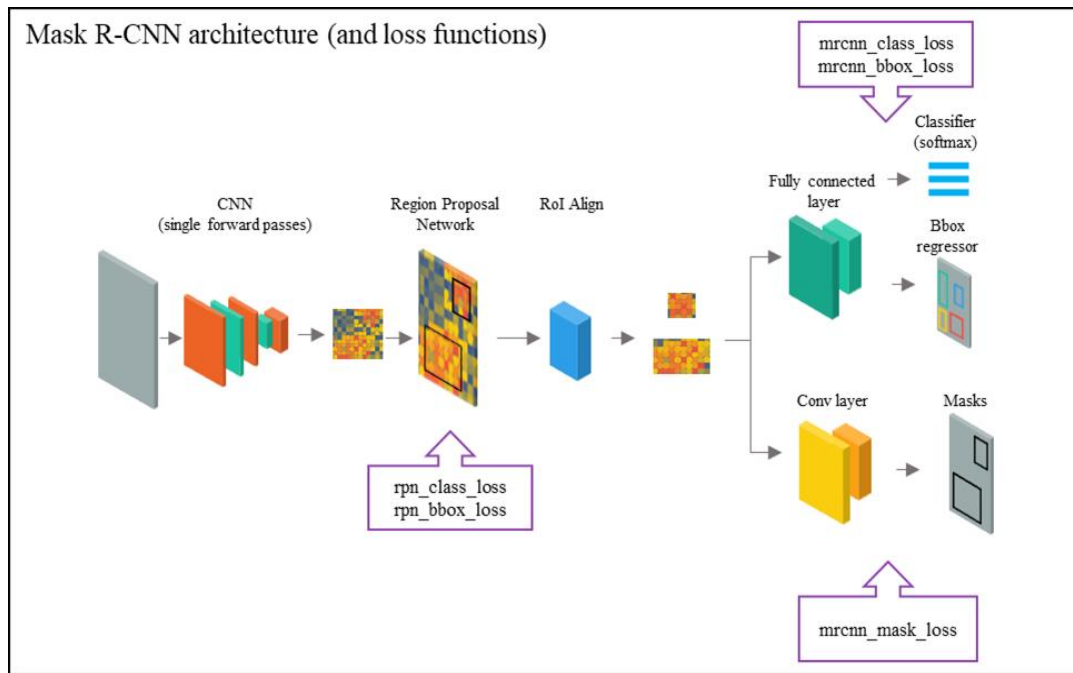
Figure 6: Matterport Mask R-CNN architecture and loss functions (adapted from Bobba, 2019)

3.2.2 Key hyperparameters

Alongside the commonly found hyperparameters that the Mask R-CNN shares with most other CNN algorithms, it also carries a set of hyperparameters that are unique to its functionality. The key hyperparameters specific to Mask R-CNN are listed below. An exhaustive list of all hyperparameters and their respective settings can be found in Appendix 10.3.

- *Backbone:* Defines the CNN that is used in the first step of the Mask R-CNN. Matterport supports three backbone options: ResNet50, Resnet101, and ResNext101. To include the necessary layer-depth but not suffer from excessive computational loads, a Resnet101 backbone with pretrained weights from the COCO model was chosen.

- *Train_ROIs_Per_Image:* Defines the maximum number of ROIs which are generated by the RPN for a given image.

- *Max_GT_Instances:* Defines the maximum number of instances (in this case PV panels) which can be segmented per image. Default value is 200.

- *Detection_Min_Confidence:* Defines the confidence threshold at which the detection is made by the model or not. Default value is 0.9.

- ***Image_Min_Dim and Image_Max_Dim:*** Defines the image size. As per the default settings, images are resized to squares of 1024 x 1024. This setting is crucially relevant as it determines the dimensional changes that are needed for comparing the masks to the SolkatDach dataset.

- ***Loss weights hyperparameter:*** The model optimizes its parameters based on five loss functions (see 3.2.3). The default weight of each loss function is set to 1, meaning all loss functions are weighted equally to optimize the model.

3.2.3   Loss functions

Matterport's Mask R-CNN implementation uses a weighted sum of different loss functions that are optimized at each state of the model (see Figure 6). The five individual loss functions are as follows:

- ***Rpn_class_loss:*** Defines the anchor box classification accuracy (absence/ presence of objects) or in other words, how well the backgrounds and objects are separated from each other by the RPN.

- ***Rpn_bbox_loss:*** Defines the RPN's bounding box localization accuracy or, in other words, how well the RPN can localize objects.

- ***Mrcnn_class_loss:*** Defines the classification accuracy or, in other words, how well the Mask R-CNN recognizes each class of objects.

- ***Mrcnn_bbox_loss:*** Defines the mask bounding box localization accuracy or, in other words, how well the model can localize objects through the masks.

- ***Mrcnn_mask_loss:*** Defines the loss attributed to pixel level masks or, in other words, how well the Mask R-CNN segments objects.

3.2.4   Image augmentation

Image augmentation was used to further extend the dataset. By using the imgaug library, 5/6 randomly selected images of the train datasets were augmented through either a horizonal or vertical flip, altering the rotation or scale. It is worth noting that the Mask R-CNN by Matterport categorizes only certain image augmentations techniques as safe (see mrcnn.model.py). The image augmentation was directly implemented to the train function found in the PVseg.py module as of iteration 3.
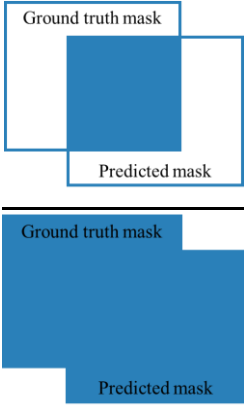
3.2.5   *Performance metrics*

Link to code

In their Mask R-CNN implementation, Matterport included some performance metrics functions within the mrcnn.utils module. However, these functions are solely based on the bounding boxes, which is suboptimal for PV segmentation. To analyze the model performance on a mask level, new performance metrics that function on a pixel level were developed. The performance of the model was analyzed (see 5.2.2) through the following performance metrics whereby the calculations were performed for each picture and then averaged to represent the entire validation dataset:

$$T_p = true\ positives,\ T_n = true\ negatives,$$

$$F_p = false\ positives,\ F_n = false\ negatives$$

- ***iou_score (intersection over union score):*** Pixel level comparison of the intersection and the union of ground truth masks (gt_masks) and predicted masks (pred_masks). This is the first and foremost important metric used to evaluate the Mask R-CNN model in this thesis.

$$IOU_{(pixel-level)} = \frac{area\ of\ intersection}{area\ of\ union} =$$

Equation 1: IOU (pixel level)

- ***mask_precision:*** Defines the ability of the model to only segment relevant object which is the percentage of correct positive predictions among all positive predictions (Padilla et al., 2021).

$$Precision_{(pixel-level)} = \frac{\sum_{i=1}^{n} T_p}{\sum_{i=1}^{n} T_p + F_p} = \frac{true\ positive}{total\ predicted\ positive}$$

Equation 2: Mask precision (pixel level)

- **_mask_recall:_** Defines the ability of the model to segment relevant cases which is the percentage of correct positive predictions among all actual positive (Padilla et al., 2021).

$$Recall_{(pixel-level)} = \frac{\sum_{i=1}^{n} T_p}{\sum_{i=1}^{n} T_p + F_n} = \frac{true\ positive}{total\ actual\ positive}$$

<center>*Equation 3: Mask recall (pixel level)*</center>

- **_mask_accuracy:_** Defines the number of correct predictions among the total predictions.

$$Accuracy_{(pixel-level)} = \frac{\sum_{i=1}^{n} T_p + T_n}{\sum_{i=1}^{n} T_p + T_n + F_p + F_n} = \frac{correct\ predictions}{total\ predictions}$$

<center>*Equation 4: Mask accuracy (pixel level)*</center>

- **_mask_f1:_** Defines the harmonic mean of precision and recall.

$$F1 - Score_{(pixel-level)} = 2 * \frac{precision * recall}{precisions + recall}$$

<center>*Equation 5: Mask F1 (pixel level)*</center>

3.3. Task 3: Comparison to SolkatDach suitability categorization

With the Mask R-CNN model functioning, the last step, the comparison to the SolkatDach suitability categorization can be undertaken. This step contains only two Subtasks; (1) the Mask R-CNN is utilized to generate PV segmentation masks which are then (2) compared to the SolkatDach model originating from the sonnendach.ch project.

3.3.1   Subtask 1: Inferencing and generation of PV segmentation masks

Link to code

This step is fairly straight forward. The final (iteration 4) Mask R-CNN model is loaded and used to generate binary segmentation masks. As the model creates a separate mask for every PV panel cluster it detects, all masks of a given picture were merged and then flattened to a one-dimensional array with binary values representing the presence or absence of PV panels for a given pixel. The arrays where then saved as ".jpg" images alongside their respective EGID. This was conducted for the pred_masks as well as the gt_masks.

3.3.2   Subtask 2: PV placement assessment

Link to code

The concluding step of this prototype methodology compares the pred_masks from the final model to the rooftop surface suitability categorization dataset SolkatDach. Unlike the model performance evaluation undertaken in Task 2 (see 3.2), this step requires dimensional changes to the categorization polygons from the SolkatDach. This is indispensable as the size and shape of the input images is altered inside the Mask R-CNN model resulting in dimensional differences between the masks (pred_mask and gt_mask) and the suitability categorization polygons from the SolkatDach. This challenge was solved by rasterizing the SolkatDach ".geojson" file and then converting it to a three-dimensional RGB image containing the suitability categorization number (1= low to 5= top) as its RGB code on a pixel level. The image could then resized following the logic of the resize_image function from the mrcnn.utils module. The image was then flattened to a one-dimensional array, thus, becoming dimensionally identical to the corresponding masks (see Appendix 10.6). By multiplying this one-dimensional array containing the suitability categorization with the pred_mask and gt_mask, the extent of rooftop exploitation could be estimated based on the pixel level frequency. The values of this pixel-level frequency of occurrence were then saved as a ".csv" file for further analysis. This concludes the prototype methodology developed in this thesis.
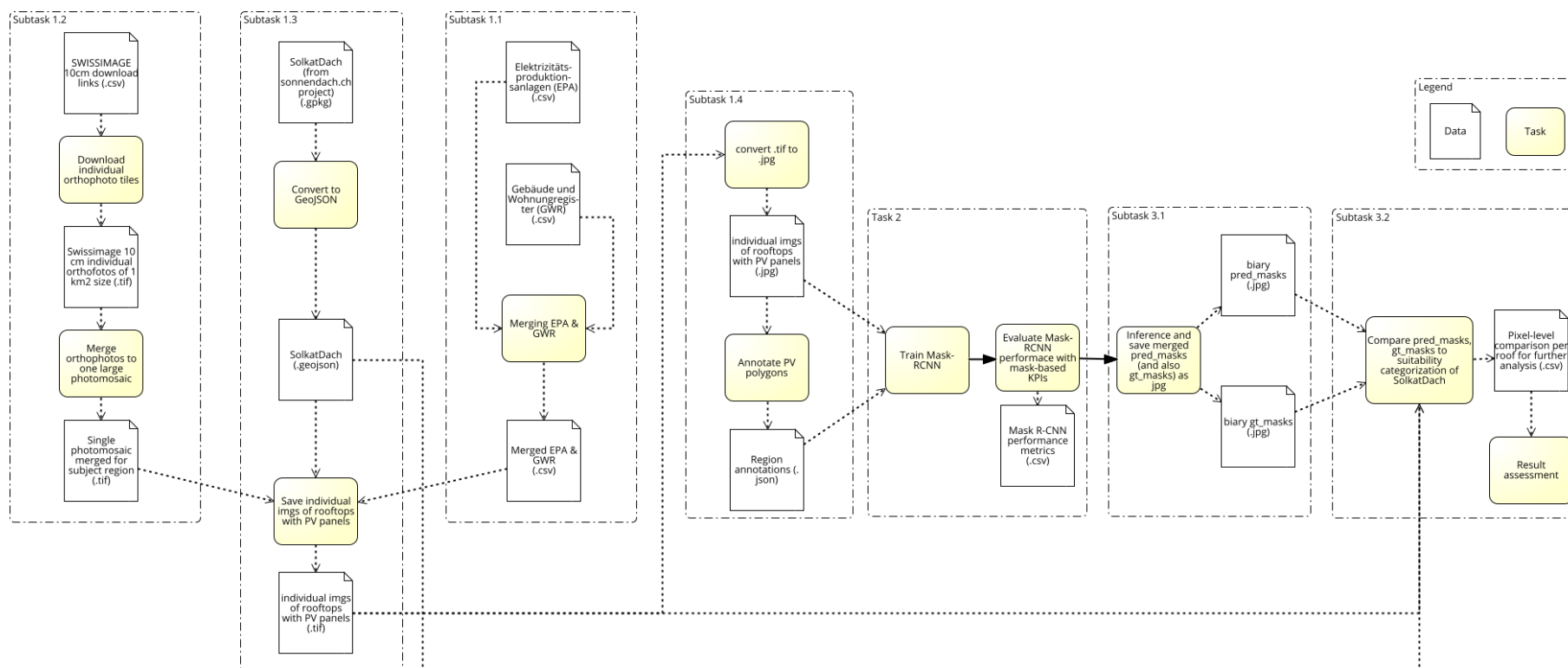
Figure 7: Overview of the prototype methodology developed (including data sources) visualized in Signavio

4. Data sources

All data sources used in the development of the prototype methodology outlined in the previous chapter are entirely open source and provided by the Swiss government. Table 2 provides an overview of the data sources detailed in this chapter.

Table 2: Overview of data sources (URL linked to source)

Name	EPA	GWR	SWISSIMAGE 10 cm	SolkatDach
Source	swisstopo	housing-stat	swisstopo	swisstopo
Format	.csv	.csv	.tif	.gpkg
Variables	XTF ID	EGID	areal image (10 cm resolution)	Plane polygons
	Post code	Post code		Suitability
	Municipality	Municipality		Rooftop area
	Canton	Canton		Orientation
	Beginning of operation	Class of building		Financial return
		Year of construction		Inclination
		Type of heating		
	Initial power	Area		
	Total power	Status		
	Main category	Number of levels		
	Subcategory	Renovations		
	Plant category	Longitude		
	Longitude	Latitude		
	Latitude	others		

4.1. ElectricityProductionPlants (EPA)

As described in section 3.1, the EPA dataset encompasses all electricity production plants registered in Switzerland. This includes large plants such as the nuclear power plants as well as small-scale privately owned PV installations that are of interest in this thesis. For each plant information such as the address, coordinates, beginning of operation or power output is available. Solely electricity production plants that are currently operational are included in the dataset. A supplementary dataset entitled PlantDetail contains additional information such as the orientation and inclination of PV installations (Bundesamt für Energie, 2022b). However, this does not contribute to the objectives of this thesis. All information can be freely accessed in the form of ".csv" files via the opendata.swiss portal (Bundesamt für Energie, 2021a). Figure 8 shows a visualized excerpt of the ElectricityProductionPlants dataset.

Figure 8: ElectricityProductionPlants (EPA) dataset visualized

4.2. Federal Register of Buildings and Dwellings (GWR)

In order obtain the EGID, which is needed for the precise determination of the shape and location of each building rooftop (see 3.1.3), the Federal Register of Buildings and Dwellings was utilized. The GWR contains information on construction projects, buildings, apartments, building entrances and streets and can be freely accessed through the housing-stat.ch platform (Bundesamt für Statistik, 2021). The EGID plays a crucial role as it functions as the primary key that allows for interlinking all other data used in this thesis.


Figure 9: Federal Register of Buildings and Dwellings (GWR) dataset visualized

4.3. SWISSIMAGE 10 cm

The aerial images used in this thesis originate from the publicly available orthophoto mosaic SWISSIMAGE 10 cm. An orthophoto is an aerial image in which the inclination influences of the camera and the terrain were corrected. In the case of the SWISSIMAGE 10 cm dataset, the images were taken by plane while all image data was taken in the nadir (perfectly orthogonal relative the ground).

The images with a ground resolution of 10 cm in the plain areas and 25 cm over the alpine regions are updated on a triannual basis. Therefore, some PV installations included in the EPA dataset might not yet be shown in the aerial images from SWISSIMAGE 10 cm. Due to the actuality of the images, the canton of Aargau was chosen for image collection. SWISSIMAGE 10 cm is accessible in image tiles of 1 km^2 each. The tile numbering corresponds to the kilometer coordinates of the southwest corner point of a tile. The tiles can be downloaded freely via the swisstopo portal (Bundesamt für Landestopografie, 2022). Figure 10 displays the current state of the SWISSIMAGE 10 cm orthophoto mosaic which served as the criteria for identifying suitable regions.
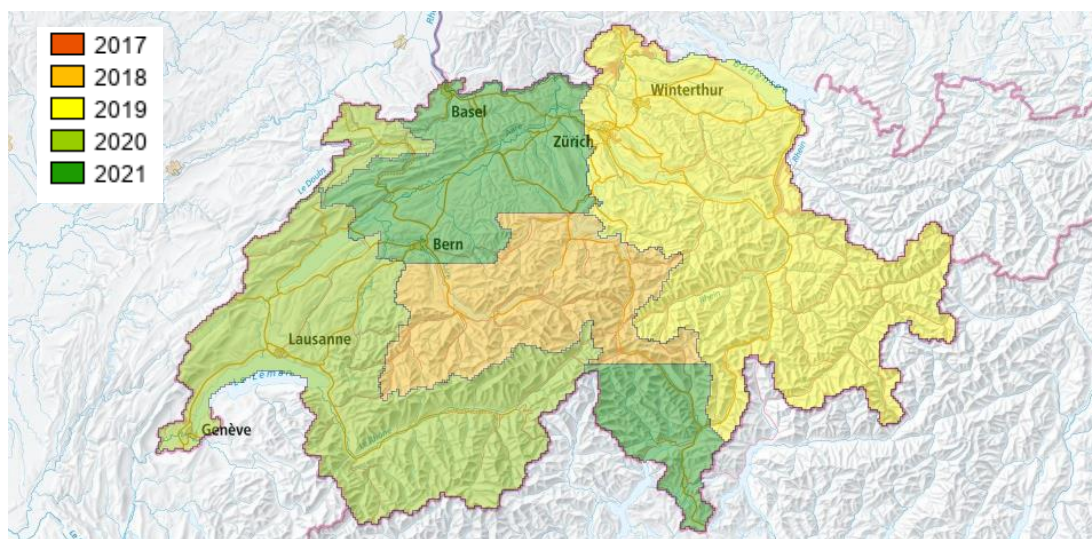

Figure 10: Actuality of SWISSIMAGE 10 cm orthophoto mosaic (as of April 2022)

4.4. SolkatDach

As aforementioned in section 2.1, the Federal Office of Energy, the Federal Office of Topology, and MeteoSwiss funded the large-scale project sonnendach.ch. The aim of this study was to analyze the total potential power production of PV installations as well as developing a model that provides a PV potential estimate of each building in Switzerland. The model is based on the swissBUILDINGS3D 2.0 vector dataset that represents buildings as 3D models with individual rooftop and facade surfaces including their overhangs. The 3D information was then converted into 2D rooftop surfaces (birds-perspective) and enriched with various additional information such as a terrain model, climate data as well as digital elevation models from the Shuttle Radar Topography Mission.

The project resulted in a dataset that contains the suitability (rated from 1= low to 5= top), the area in m^2, the orientation as well as the inclination of each individual rooftop surface (Klauser, 2016). This information serves two purposes: Firstly, the polygons of the rooftop surfaces are used to obtain the training images (see 3.1.3). Secondly, the rooftop suitability categorization is used to assess whether PV panels are placed optimally (see section 3.3).The dataset can be downloaded as a ".gpkg" file that can be freely accessed from the Federal Office of Energy through the geocat.ch portal (Bundesamt für Energie, 2022a). Figure 11 show the SolkatDach dataset as visualized on the geoadmin platform as well as in its raw form.
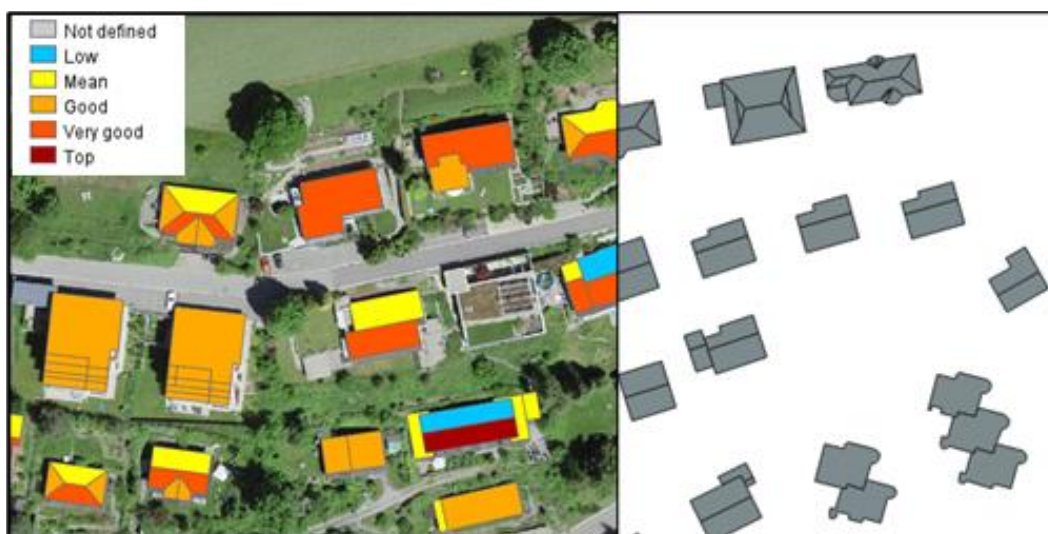
Figure 11: SolkatDach as visualized on the geoadmin platform and in raw form

5. Results

Analogous to the Methodology section, the Results section is split into three parts. Firstly, the dataset obtained through Task 1 is detailed (Sub-RQ1). Following this, the training and inferencing performance of the Mask R-CNN are discussed in detail (Sub-RQ2). Lastly, the outcomes from the placement assessment associated with answering Sub-RQ3 are elaborated.

5.1. Results of data collection and processing

Through the scripts outlined in the Methodology section 3.1, several hundred individual images of rooftops with PV panels were downloaded by specifying the dimensions of a desired region. The process is solely limited by the excessive computational load resulting from the merging of the individual image tiles to one single photomosaic. These merged photomosaics quickly reach a file size of >100 GB which exceeds the computational capacity of the available RAM (see chapter 3) by far. Overall, a total of 1130 individual images of rooftops that have PV panels installed were collected. Approximately 9% of all downloaded rooftop images contain no visible PV panels hand where thus deleted. This error can be attributed to placements of PV panels elsewhere than on the rooftop itself (e.g., in the garden, or self-standing buildings adjacent to the main roof). In very few cases, the error can also be explained by wrongful information in the EPA, GWR or SolkatDach datasets. Depending on the nature of the rooftop, a single image can contain one up to approximately 80 polygons (e.g., large industrial buildings), each representing individual PV panel cluster. On average, a single rooftop image contains 2.5 polygons. Relative to residential buildings, industrial buildings typically have a much more complex structure and, thus, far more polygons. At approximately 50 images per hour, the speed at which images could be annotated was roughly twice as fast compared to previous studies but still time consuming and can thus be considered another limitation (Meyer, 2020). Given the successful application of the data collection methodology developed in this thesis, Sub-RQ1 can be answered affirmatively. Table 3 provides an overview of the characteristics of each dataset per iteration.

Table 3: Train and validation dataset characteristics per iteration

Variable	Images	Polygons	Polygons / image	Train / test split
iteration_1				
train	131	345	2.6	90%
validation	15	26	1.7	10%
iteration_2				
train	271	781	2.9	87%
validation	42	120	2.9	13%
iteration_3				
train	501	1313	2.6	89%
validation	60	159	2.7	11%
iteration_4				
train	974	2512	2.6	86%
validation	156	367	2.4	14%
Test set				
Equal for all iterations	156	367	2.4	
Total annotated	1130	2879	2.5	

5.2. Results Mask R-CNN

As mentioned in the Methodology, the development and training of the Mask R-CNN model was conducted iteratively. The four iterations differ solely in terms of the training dataset size. While the size of the validation dataset used during model training was roughly proportional to the size of the train dataset (see Table 3), the model performance was evaluated using the same test dataset for all four iterations. This should facilitate an unbiased comparison between the performance of each iteration. All hyperparameters, remain constant throughout the iterations unless stated otherwise. Experiments with tweaking hyperparameters (see 10.3) were undertaken but showed little to no positive impact on model performance and where thus set to the specifications outlined in Appendix 10.3 for all iterations.

5.2.1 Mask R-CNN training

The first model implemented, iteration 1, indicated a typical overfit learning curve. This was caused by a vastly insufficient training dataset of merely 131 train images (see Figure 12). By doubling the size of the training dataset to 271 images for iteration 2, the strong overfit scenario could be slightly reduced, but not yet fully eliminated (see Figure 12).

With a dataset of just over 500 train images and image augmentation implemented, iteration 3 appeared to have reached the critical mass to overcome the risk of overfitting. By doubling the train dataset again to 974 images and increasing the epoch from 30 to 45, the model reached the loss plateau with the available data. The model eventually approximated a train as well as validation loss of ≈ 0.2 (all loss weights being equal see 3.2.2). Figure 12 shows the train and validation curves for each iteration.
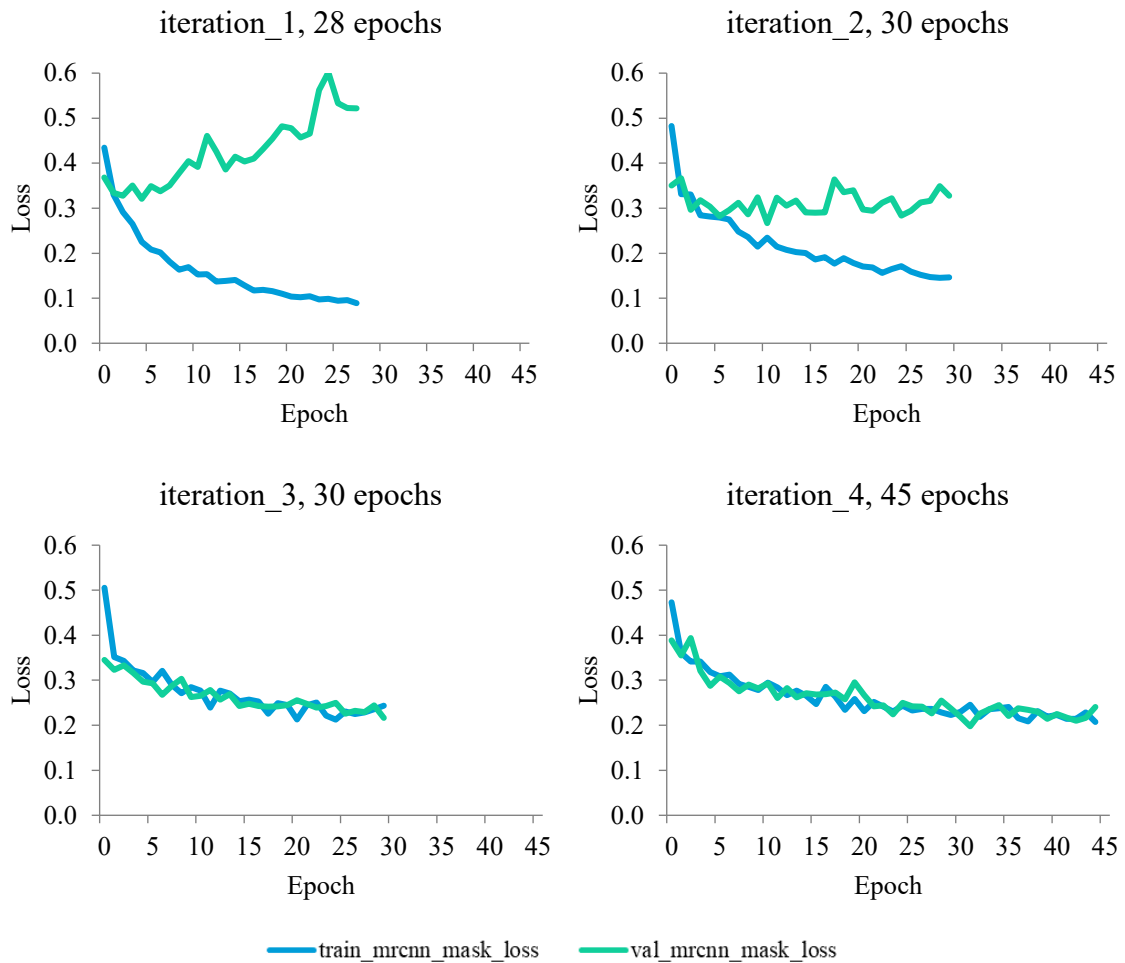

Figure 12: Train and validation curves per model iteration

### 5.2.2 Mask R-CNN model validation

As mentioned in section 3.2.5, the performance metrics included in Matterport's Mask R-CNN implementation are calculated based on the bounding boxes of every detection and not on a mask level. Due to this limitation the model's performance was analyzed using the performance metrics specified under 3.2.5. The model's inference performance progresses similarly to the train and validation losses. That is, improvement as a result of increased training dataset size.

Over all 156 images of the test dataset, the average iou_score (see 3.2.5) improves from 0.60 to 0.74 from the first to the fourth iteration respectively (see Figure 13). Interestingly, the standard deviation of the predicted iou_scores only decreases following iteration 3 (see Table 10). This observation aligns with the elimination of the overfit scenario detailed in section 5.2.1. We can also note that the model iou_scores range from 0.13 to 0.94, indicating that the model never fails to segment PV panels entirely but has difficulties to segment the entirety of the panels (see Table 10). This is likely linked to the weakness of Mask R-CNN to accurately segment rectangular shapes and corners as found by Liang et al. (2020).
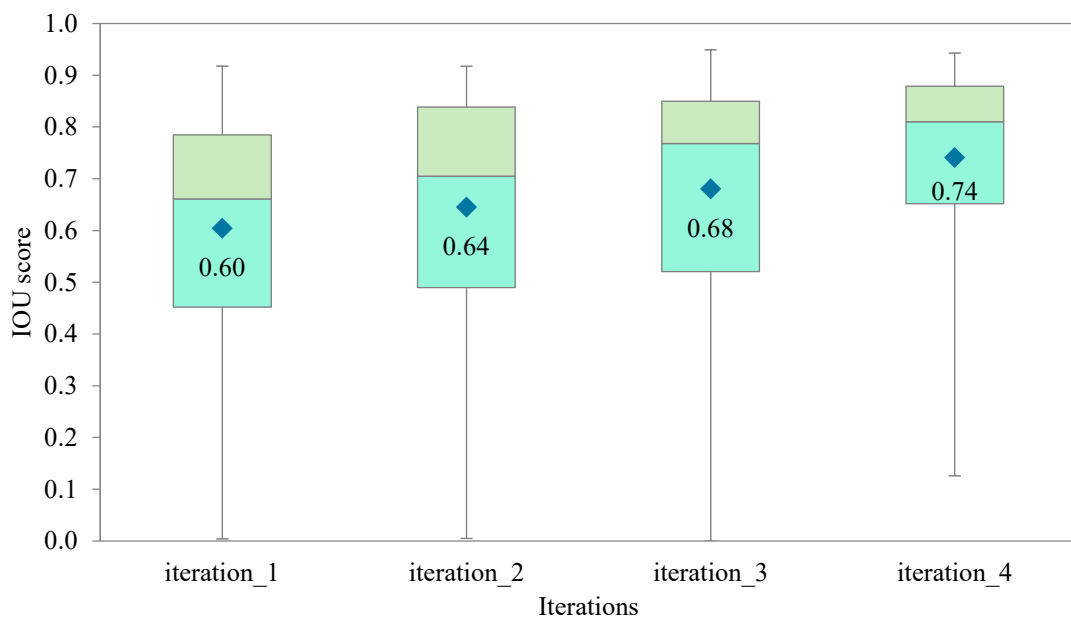
*Figure 13: Mask R-CNN per image IOU score distribution per iteration*

Likewise improve the metrics mask_precision, mask_recall and, subsequently mask_f1 with every iteration (see Figure 14). The mask_accuracy witnesses only light improvements as this metric also takes T_n pixels into account which dampens the results. The improvements over the four iterations become slightly more pronounced when looking at the median which improved from 0.66 to 0.82 from iteration 1 to iteration 4. In the final model, iteration 4, the average mask_precision and mask_recall are almost equal with 0.86 and 0.85 respectively (see Figure 14). This indicates that F_n, and F_p errors are represented equally frequent.
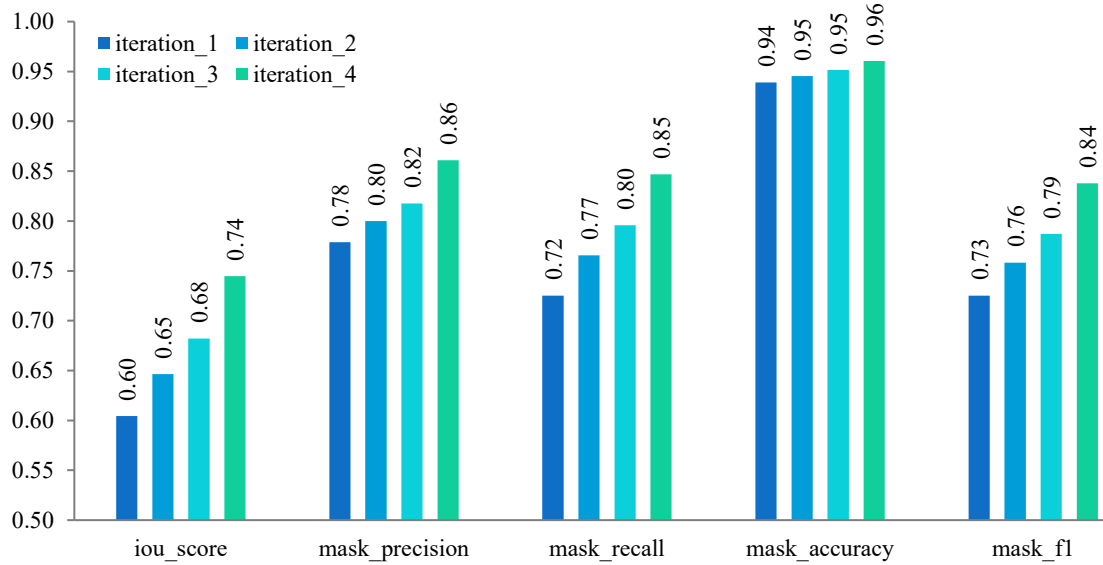
Figure 14: Mask R-CNN performance metrics averages per iteration

While the Mask R-CNN model still has room for improvement (see Discussion), accurate PV segmentation is possible for most simple rooftop structures. However, the segmentation performance weakens as the rooftop complexity increases (e.g., through roof windows, conservatories, or HVAC) (see Figure 30). As a result, Sub-RQ2 can be answered affirmatively in principle but not unconditionally. Depending on the use case, the segmentation performance might be insufficient (see Discussion). All performance metric evaluations can be found in Table 10 in the Appendixes section. A selection of randomly selected prediction masks for each iteration can be found in section 10.5.
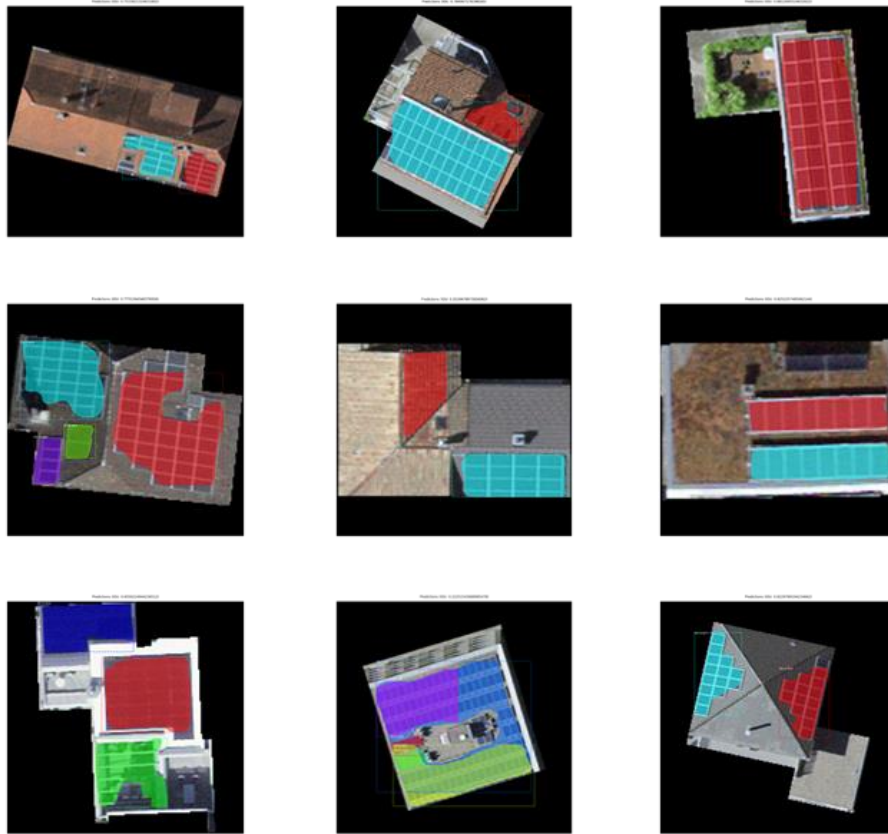
Figure 15: Mask R-CNN prediction masks (pred_mask) over input image from iteration 4 preview (randomly selected)

## 5.3. Results of comparison to SolkatDach suitability categorization

Following the procedure outlined under 3.3.1 and 3.3.2 both the pred_masks and gt_masks were compared to the rooftop suitability categorization of the SolkatDach dataset. The test dataset alongside the dataset used to train the Mask R-CNN were assessed which theoretically could contribute to biased results as the Mask R-CNN conducts inferencing on images already considered during training. However, the outcome differences between the test and train dataset were non-significant. Nevertheless, the data presented in this section originates solely from test dataset to ensure methodical correctness unless states otherwise. The combined outcomes from both the test and train dataset can be found in Appendix 10.7. Also, it must be noted that the obtained results exclusively stem from a very narrowly bounded region and are, thus, not representative.

The first and foremost observation is that from buildings with existing PV installations, only 29% of the total aggregated rooftop area is occupied by PV panels. When calculated on a per-roof-basis, the rooftop utilization ratio increases to 33%. Both numbers were calculated based on gt_masks exclusively and are therefore not affected by imperfect prediction by the Mask R-CNN model. Nevertheless, the distribution of available rooftop utilization does not differ between greatly between the pred_mask and gt_mask as can be derived from Figure 16 (or Figure 32) meaning that on an aggregated level, the model can be deemed accurate.
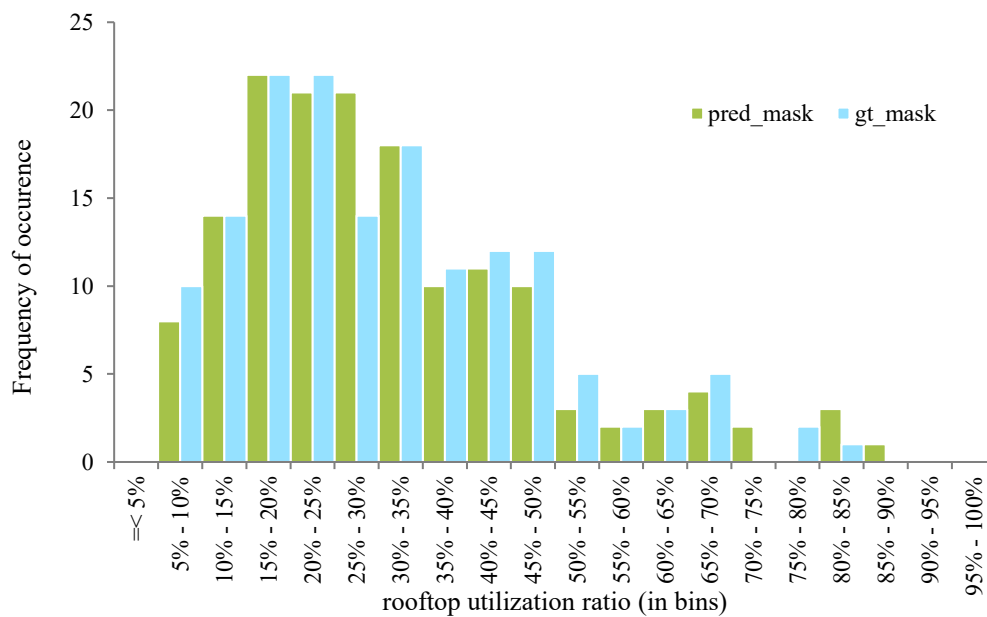
Figure 16: Distribution of available rooftop area utilization (comparison pred_mask and gt_mask, test dataset only)

From the average rooftop size of the test dataset of 334 m^2 only 105 m^2 are exploited for PV panels. The distribution of available rooftop area also follows a positively skewed distribution. When paired with the MAPE, calculated by comparing the rooftop utilization ratio from the pred_masks to the gt_masks, two things become evident: The model performs relatively constant for rooftops sized <1500 m^2 averaging a MAPE of 16% (excluding outliers >1.0). However, above this threshold, the performance decreases rapidly to >50% (unsmoothed) as illustrated in Figure 17.
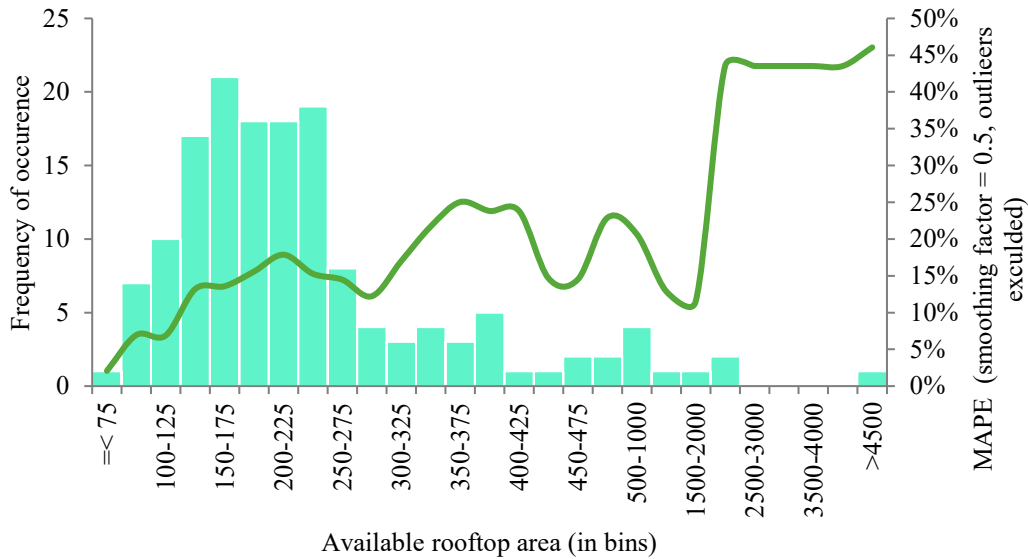
Figure 17: Distribution of available rooftop area in relation to rooftop area utilization MAPE between pred_mask and gt_mask (test dataset only)

The identification of buildings with rooftops that are equipped with PV panels which are not optimally placed, was determined by the following to criteria; (1) the given building does not have any PV panels installed in the highest available rooftop area category and (2) the highest available rooftop area category amounts to ≥10% (threshold) of the total available rooftop area. When both criteria are fulfilled, a building is categorized as "critical" meaning that the existing PV panels are not optimally placed to maximally exploit solar radiation. The sensitivity analysis in Table 4 shows that with an increasing threshold, the categorization performance decreases suggesting the use of a low threshold. In practice however, few homeowners will invest in PV installations to solely cover a marginal percentage of the available rooftop area. Taking this into account, the threshold should be adapted to the specific use case requirements. When using a threshold of 10% a total of 31 buildings are categorized as critical which amounts to 2.8% of total buildings assessed. The Mask R-CNN model manages to detect 22 of which, resulting in a recall of 0.71. Figure 18 shows an excerpt of rooftop buildings that have suboptimally placed PV panels and were detected by the model. The transparent colors thereby correspond to the categorization according to the suitability categorization from SolkatDach (blue= low to red= top). The location of existing PV panel is demarked with green boundaries. Additional images can be found in the Appendixes.

Table 4: Sensitivity analysis of detection of rooftops with misplaced PV panels ("critical") depending on thresholds

Threshold	Accuracy	Precision	Recall	F1	TP	FN	Total critical
5%	0.97	0.54	0.77	0.63	33	10	43
10%	0.96	0.36	0.71	0.48	22	9	31
15%	0.95	0.26	0.67	0.38	16	8	24
20%	0.95	0.18	0.61	0.28	11	7	18
25%	0.95	0.16	0.67	0.26	10	5	15
30%	0.95	0.11	0.64	0.19	7	4	11
35%	0.95	0.08	0.56	0.14	5	4	9

Figure 18: Excerpt of buildings with suboptimally placed PV panels detected by the model (categorized critical), green boundaries indicate existing PV panels

Given the successful comparison to the suitability categorization of the sonnendach.ch project, Sub-RQ3 can be answered affirmatively too. As all three Sub-RQs could be principally confirmed, the overarching RQ can also be positively answered in principle.

When looking at the average categorization of all 1108 images investigated, one can note that roughly two thirds of all rooftop areas are categorized as good (3) or very good (4). More interestingly however, is that the rooftop utilization ratio grows almost linearly over the suitability categories, meaning that the more suitable a rooftop area is, the greater its extents of exploitation (see Figure 19 and Table 5).
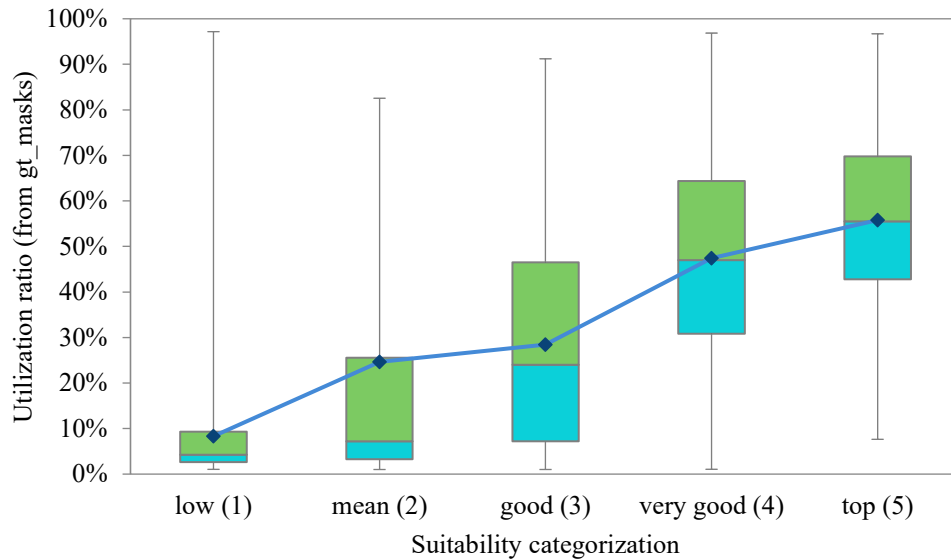
Figure 19: Growth of utilization ratio (from gt_masks) over suitability categorization

This finding motivated an extrapolation of PV potential estimate for all of Switzerland, despite this not being the primary focus of this work. The potential estimate is based on the average rooftop categorization as well as the average utilization derived from the gt_masks. Given the electric output defined by the SolkatDach categories as well as the total available rooftop surface of 252 km^2 indicated by Assouline et al. (2018), the estimated total annual PV potential in Switzerland ranges from 35.4 to 41.4 TWh. This estimate is brought into context in the Discussion.

Table 5: Extrapolation of total annual PV potential in Switzerland based on 252 km^2 total available rooftop area as found by Assouline et al. (2018)

Variable	Low (1)	Mean (2)	Good (2)	Very good (4)	Top (5)
PV placement assessment					
average rooftop categorization	11%	20%	32%	34%	3 %
average utilization from gt_masks	8%	25%	29%	47%	56%
SolkatDach					
min KWh / year / m^2	-	800	1000	1200	1400
max KWh / year / m^2	800	1000	1200	1400	-
Extrapolation					
available rooftop area (in m km^2)	28.2	50.4	80.5	85.7	7.3
utilized rooftop area (in m km^2)	2.3	12.4	23.2	40.6	4.0
min energy potential (in TWh)	0.0	2.4	6.7	23.1	3.2
max energy potential (in TWh)	0.2	3.1	8.0	27.0	3.2
average energy potential (in TWh)	0.1	2.8	7.4	25.0	3.2

6. Discussion

With the prototype methodology developed in this thesis, the feasibility of a large-scale detection and comparison of existing PV panels to the suitability categorization of the sonnendach.ch project was technically proven. As such, several hundred images from building rooftops with existing PV panels installed can be downloaded for a given area with only approximately 9% of the images not containing any PV panels. The reasons therefore are placements outside the main rooftop boundaries or wrongful information in the underlying datasets. Nevertheless, it must be noted that the merging of the orthophotos into a single photomosaic as required by the data collection process results in file sizes that rapidly lead to unmanageable computational loads. This may be considered a potential bottleneck of the process but could be solved by employing additional computational power comparable to the FHNW project (Meyer, 2020). Moreover, despite being roughly twice as fast as previous studies, manually annotating the PV panels on rooftop images is time consuming. Overall, the annotation of all 1130 images consumed approximately 25 hours. Through that, the Mask R-CNN which was trained on a total of 974 images improved its performance progressively over the four iterations conducted and achieved a final iou_score of 0.74. The predicted masks were then successfully compared to the rooftop area suitability categorization from the sonnendach.ch project. Thus, all three Sub-RQs defined in section 1.4 were answered affirmatively in principle. However, whether or not the segmentation performance of the Mask R-CNN is sufficient depends on the comparison use case. Is the goal to identify buildings that have their PV panels placed suboptimally for commercial reasons, the performance of the final model can be considered acceptable. This is the case as from the entire dataset 22 out 31 critically misplaced PV panels could be identified (recall of 0.71 (see 5.3)) which might be considered sufficient for commercial use cases such as the launching of after-sales initiatives. When the use case requires more than a binary categorization, the performance of the Mask R-CNN model might not be sufficient. This is particularly the case when assessing industrial regions as the model has difficulties in segmenting PV panels on more complex rooftop structures that are frequently found in non-residential areas (see 5.2).

As mentioned in section 5.3, the average rooftop utilization ratio of 29% did not differ greatly on an aggregate level between the pred_masks and gt_masks.

As the distributions also follow similar patterns, we can conclude that the Mask R-CNN model is sufficiently accurate for aggregate estimates such as the average rooftop utilization ratio (see Figure 16 vs Figure 32). However, when looking at the averages of each suitability category individually, the results differ vastly which indicates that the model's performance is non-satisfactory for detailed analyses. This can mostly be explained by outlier values. Also, it must be noted that the obtained results are based on a limited number of rooftops in a relatively narrowly bounded region in canton of Aarau (see Figure 21). Subsequently, the results of this assessment are not representative for the entirety of Switzerland. The results should be seen as an indication to what the developed prototype methodology is capable and considered under the circumstances outlined in this thesis. While many error sources exist, it is vital to emphasize that the quality of the comparisons exceedingly depends on the segmentation performance of the Mask R-CNN. Table 6 provides a critical overview of the challenges experienced in the development of the prototype methodology and weights them according to their estimated severity.

Table 6: Challenges experienced in the development of the prototype methodology

Origin	Challenges	Est. severity (1-5)
Datasets	Imprecise mapping between SolkatDach and SWISSIMAGE 10 cm	5
Datasets	Rooftop inclination affecting m^2 calculation	2
Datasets	Incorrect information in EPA / GWR dataset	2
Datasets	SWISSIMAGE 10 cm not taken from nadir	1
Datasets	Reflections from sun	1
Datasets	Warped imaging	1
Logic/others	PV panels placed outside rooftop boundaries (e.g., in garden)	4
Logic/others	Hardware limitations	5
Logic/others	Noise associated with the resizing of raster data	1
Mask R-CNN	Complex rooftop structures	5
Mask R-CNN	Segmentation of rectangular corners	5
Mask R-CNN	Visual similarity to solar heating systems	3
Mask R-CNN	Squared rooftop patterns (often resulting from tiling)	2

When bringing the findings of this thesis into the context of existing PV potential studies, it becomes evident that critical underlying assumptions for country-wide extrapolations might be conflicting. This thesis found that the average rooftop utilization ratio equates to 29% while previous studies estimated the ratio to be between 50% and 70% (see Table 1). The rooftop utilization ratio is thus likely to be systematically overestimated by all studies analyzed as part of the literature review (see section 2.1). However, the reducing effects this has on the total PV potential estimates are partly compensated by the non-

uniform distribution of rooftop utilization ratios across different suitability categorizations. In other words, while previous studies assume the same rooftop utilization ratio for all buildings, this thesis found that the rooftop areas with a higher suitability categorization are exploited to a greater extent than those with a low suitability. Thereby, the rooftop utilization ratio grows almost linearly across the suitability categorizations (see Table 5). Subsequently, this thesis estimates the total PV potential of Switzerland to be between 35.4 and 42.4 TWh annually, which is greatly below the potential estimate conducted by the sonnendach.ch project (Portmann et al., 2016) but in line with the most recent findings by Moro et al. (2021) (see Table 1). Other sources of differences between the studies include the disregard of small rooftop surfaces or such with a low suitability categorization, varying estimates of the total available rooftop area or the efficiency of the PV cells.

The reason for PV installation owners merely exploiting roughly one third of the available rooftop areas is tightly linked to the pricing dynamics behind private solar electricity exchange. When the PV electricity generation exceeds the maximal electricity consumption of a building, excess electricity can be sold back to the grid at a regionally specified feed-in tariff. The problem thereby is that energy providers are buying the electricity at a discount so substantial that producing excess electricity becomes uneconomical (see Appendix 10.9) (VESE, 2021). Subsequently, the prototype methodology developed in this thesis could be used to design a dynamic incentive system that aims at financially motivating PV installation owners that still have relatively high solar potential unexploited to fully utilize their rooftop area. On the other hand, the methodology can aid in assessing regional differences in terms of rooftop exploitation by assessing the consequences of federally planned incentive systems in place (e.g., rooftop utilization compared to feed-in-tariffs or electricity prices etc.) currently. Today, the incentive systems maximally subsidize new PV installations with 30% of the acquisition cost (Swissolar, 2022).

The prototype methodology developed in this thesis lays the foundation to scrutinize Switzerland's rooftop utilization on a national scale. Doing so would allow to draw invaluable conclusions that support the decision-making regarding energy distribution or storage. The recommendations suggested in Table 7 should serve as a starting point to improve the model quality and extend the prototype methodology to a national scale.

Table 7: Recommendations for further research to improve the prototype methodology

Origin	Recommendations for further research
Data collection	Expand rooftop polygon before cropping images so the entire rooftop is saved (problem as SWISSIMAGE 10 cm and SolkatDach are not perfectly aligned)
Data collection	Combine Python scripts to facilitate data collection
Data collection	Eliminate the need to merge orthophotos into single photomosaic so that the file sizes stay within manageable levels
Mask R-CNN	Enlarge training data set (to approx. 10'000 images as Meyer (2020)) particularly with: • More images of complex rooftop structures • More images with solar heating • More images with shadows • More images of other similar looking object (e.g., conservatories)
Mask R-CNN	Perform grid search on Mask R-CNN hyperparameters, particularly: • Image_Max_Dim (eliminate as performance decreased for large images) • Train_ROIs_Per_Image (to lower computational load) • Max_GT_Instances (to lower computational load) • Loss weights hyperparameter (give more weight to mrcnn_mask_loss to favor masks over other loss functions)
Mask R-CNN	Improve rectangular / corner segmentation with right-angle polygon fit algorithm as implemented by Liang et al. (2020)
Mask R-CNN	Expand Mask R-CNN to segment multiple classes (incorporation of rooftop structures such as chimneys and roof windows or solar heating panels)
Comparison to SolkatDach	Combine Python scripts to incorporate both inferencing and comparison to SolkatDach
Comparison to SolkatDach	Extract all available information from the SolkatDach dataset and enrich with EPA dataset to extract further insights
Comparison to SolkatDach	Conduct geo-referenced polygon-based comparison rather than pixel level to improve accuracy and enable further fitting options
Others	Acquire suitable hardware (e.g., Nvidia Tesla V100 SXM2 with a HPE Apollo 6500 Deep Learning GPU server or Amazon SageMaker instance with comparable power)

7. Conclusion

As part of this thesis, a prototype methodology that can detect and compare the placement of PV panels with the rooftop area suitability categorization according to the sonnendach.ch project was developed. This was achieved by leveraging publicly available data in conjunction with a Mask R-CNN architecture facilitating accurate PV segmentation. With the methodology, 1130 images of individual rooftops were automatically collected and then manually annotated. The Mask R-CNN was trained with 974 images over four iterations of which the final achieved an iou_score of 0.74. The segmentation performance of the Mask R-CNN is sufficient for commercial use cases requiring binary classification such as the identification of suboptimally placed PV panels. However, for detailed quantitative analyses, the performance achieved with the resources of this thesis is insufficient. The final Mask R-CNN model performs best when rooftop structures are simple and uniform.

However, when complexity in form of roof windows, conservatories, HVAC, or other obstacles arise, the model experiences prediction difficulties. Given the functionality of the developed prototype methodology, all three Sub-RQs, and thus the overarching RQ can be answered affirmatively in principle but further research with the aim of improving the segmentation accuracy of the Mask R-CNN is required to allow for detailed analyses.

The findings from the comparison to the suitability categorization of the sonnendach.ch project propose that the average rooftop utilization ratio is 29%, suggesting that previously conducted studies overestimate the extent of rooftop utilization systematically. Thus, the assumptions used by the PV potential studies are likely not displaying a realistic picture. The low utilization ratio might be explained as the production of excess electricity is uneconomical since the locally defined feed-in-tariffs are vastly below the cost of electricity. The exact reasons for PV installation owners not maximally exploiting the available rooftop area were not investigated as part of this thesis. Hence, qualitative research to understand the motives behind this phenomenon should be conducted alongside quantitative assessments such as developed in this thesis. Moreover, while previous studies assumed the same utilization ratio across all suitability categorizations, the findings of this thesis suggest that the more suitable a rooftop area is, the greater its extent of utilization. This non-uniform distribution of rooftop utilization ratios across the suitability categories has not yet been considered by previous PV potential studies. Overall, this thesis proved that a large-scale assessment of PV placement efficiency is technically feasible and that previous PV potential estimates might be extrapolated on inaccurate assumptions. By implementing the improvements to the prototype methodology as outlined in Table 7, valuable insight could be obtained on a country wide level. On one hand, this would allow to display a realistic picture of the current placement efficiency in Switzerland but also allow assessing the effectiveness of regionally differing incentive systems. Hence, further research that compares the placement efficiency to incentive systems or electricity feed-in-tariffs on a regional level should be carried out. However, as the quality of the results excessively depends on the accuracy of PV segmentation masks, further studies should be conducted after improving the segmentation performance of the Mask R-CNN. It is therefore suggested that the Mask R-CNN improvements recommended in Table 7 serve as a starting point for another research project aiming at the improvement of the Mask R-CNN performance.

8. Bibliography

Assouline, D., Mohajeri, N., & Scartezzini, J.-L. (2017). Quantifying rooftop photovoltaic solar energy potential: A machinelearning approach. *Journal of Solar Energy*, *141*.

Assouline, D., Mohajeri, N., & Scartezzini, J.-L. (2018). Large-scale rooftop solar photovoltaic technical potential estimation using Random Forests. *Journal of Applied Energy*, *217*, 23. https://doi.org/doi.org/10.1016/j.apenergy.2018.02.118

Bobba, R. (2019, December 14). *Taming the Hyper-Parameters of Mask RCNN*. https://medium.com/analytics-vidhya/taming-the-hyper-parameters-of-mask-rcnn-3742cb3f0e1b

Buffat, R., Grassi, S., & Raubal, M. (2018). A scalable method for estimating rooftop solar irradiation potential over large regions. *Journal of Applied Energy*, *216*, 389–40. https://doi.org/doi.org/10.1016/j.apenergy.2018.02.008.

Bundesamt für Energie. (2021a). *Dataset: Eelektrizitatsproduktionsanlagen (EPA)*. Bundesamt für Energie. https://opendata.swiss/de/dataset/elektrizitatsproduktionsanlagen

Bundesamt für Energie. (2022a). *Dataset: SolkatDach*. Bundesamt für Energie. https://www.geocat.ch/geonetwork/srv/ger/md.viewer#/full_view/b614de5c-2f12-4355-b2c9-7aef2c363ad6/tab/complete

Bundesamt für Energie. (2022b). *Dokumentation «minimales Geodatenmodell» Elektrizitätsproduktionsanlagen*. https://pubdb.bfe.admin.ch/de/publication/download/10449

Bundesamt für Energie, B. für E. (2021b). *Analyse des schweizerischen Energieverbrauchs 2000–2020 nach Verwendungszwecken*. Bundesamt für Energie. https://www.bfe.admin.ch/bfe/de/home/versorgung/statistik-und-geodaten/energiestatistiken/energieverbrauch-nach-verwendungszweck.html

Bundesamt für Energie, B. für E. (2021c). *Statistik Sonnenenergie—Referenzjahr 2020*. Bundesamt für Energie. https://www.bfe.admin.ch/bfe/de/home/versorgung/statistik-und-geodaten/energiestatistiken/teilstatistiken.exturl.html/aHR0cHM6Ly9wdWJkYY 5iZmUuYWRtaW4uY2gvZGUvcHVibGljYX/Rpb24vZG93bmxvYWQvMTA1 Mzk=.html

Bundesamt für Landestopografie. (2022). *Dataset: SWISSIMAGE 10 cm*.

> https://www.swisstopo.admin.ch/de/geodata/images/ortho/swissimage10.html

Bundesamt für Statistik. (2021). *Dataset: Gebäude- und Wohnungsregister (GWR)*.

> https://www.housing-stat.ch/de/madd/file.html

Ghosh, S., Das, N., Das, I., & Maulik, U. (2020). Understanding Deep Learning

> Techniques for Image Segmentation. *ACM Computing Surveys*, *52*(4), 1–35.
> https://doi.org/10.1145/3329784

Girshick, R. (2015). *Fast R-CNN*. 1440–1448.

> https://openaccess.thecvf.com/content_iccv_2015/html/Girshick_Fast_R-
> CNN_ICCV_2015_paper.html

Golovko, V., Bezobrazov, S., Kroshchanka, A., Sachenko, A., Komar, M., & Karachka,

> A. (2017). *Convolutional Neural Network Based Solar Photovoltaic Panel*
> *Detection in Satellite Photos*. 6. https://doi.org/10.1109/IDAACS.2017.8094501

Gupta, R., Sossan, F., & Paolone, M. (2021). Countrywide PV hosting capacity and

> energy storage requirements for distribution networks: The case of Switzerland.
> *Journal of Applied Energy*, *281*.
> https://doi.org/doi.org/10.1016/j.apenergy.2020.116010.

He, K., Gkioxari, G., Dollar, P., & Girshick, R. (2017). *Mask R-CNN*. 2961–2969.

> https://openaccess.thecvf.com/content_iccv_2017/html/He_Mask_R-
> CNN_ICCV_2017_paper.html

International Energy Agency (IEA). (2002). *Potential for Building Integrated*

> *Photovoltaics*. International Energy Agency (IEA). https://iea-pvps.org/wp-
> content/uploads/2020/01/rep7_04.pdf

Jiang, H., Yao, L., Lu, N., Qin, J., Liu, T., Liu, Y., & Zhou, C. (2021). Multi-resolution

> dataset for photovoltaic panel segmentation from satellite and aerial imagery.
> *Earth System Science Data*, *13*(11), 5389–5401. https://doi.org/10.5194/essd-13-
> 5389-2021

K. He & L. Zhang. (2020). Automatic Detection and Mapping of Solar Photovoltaic

> Arrays with Deep Convolutional Neural Networks in High Resolution Satellite
> Images. *2020 IEEE 4th Conference on Energy Internet and Energy System*
> *Integration (EI2)*, 3068–3073. https://doi.org/10.1109/EI250167.2020.9347211

Klauser, D. (2016). *Solarpotentialanalyse für Sonnendach.ch*. Bundesamt für Energie.

> https://pubdb.bfe.admin.ch/de/publication/download/8196

Li, Q., Feng, Y., Leng, Y., & Chen, D. (2020, June 9). *SolarFinder: Automatic Detection of Solar Photovoltaic Arrays*. International Conference on Information Processing in Sensor Networks (IPSN), Sydney, Australia.

Liang, S., Qi, F., Ding, Y., Cao, R., Yang, Q., & Yan, W. (2020). Mask R-CNN based segmentation method for satellite imagery of photovoltaics generation systems. *2020 39th Chinese Control Conference (CCC)*, 5343–5348. https://doi.org/10.23919/CCC50068.2020.9189474

Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature Pyramid Networks for Object Detection. *ArXiv:1612.03144 [Cs]*. http://arxiv.org/abs/1612.03144

Lonergan, K. E., & Sansavini, G. (2022). Business structure of electricity distribution system operator and effect on solar photovoltaic uptake: An empirical case study for Switzerland. *Journal of Energy Policy*, *160*. https://doi.org/doi.org/10.1016/j.enpol.2021.112683

Malof, J. M., Bradbury, B., Collins, L. M., & Newell, R. G. (2016). Automatic detection of solar photovoltaic arrays in high resolution aerial imagery. *Journal of Applied Energy*, *183*, 12. https://doi.org/10.1016/j.seta.2020.100651

Meyer, A. (2020, July 23). *Detecting and Analyzing Solar Panels in Switzerland using Aerial Imagery (SolAI)*. europython, Online. https://ep2020.europython.eu/media/conference/slides/detecting-and-analyzing-solar-panels-switzerland-using-aerial-imagery.pdf

Moro, N., Sauter, D., Strebel, S., & Rohrer, J. (2021). *Das Schweizer Solarstrompotenzial auf Dächern* [22,application/pdf]. https://doi.org/10.21256/ZHAW-2652

N. Wang, Z. -L. Sun, Z. Zeng, & K. -M. Lam. (2021). Effective Segmentation Approach for Solar Photovoltaic Panels in Uneven Illuminated Color Infrared Images. *IEEE Journal of Photovoltaics*, *11*(2), 478–484. https://doi.org/10.1109/JPHOTOV.2020.3041189

Padilla, R., Passos, W. L., Dias, T. L. B., Netto, S. L., & da Silva, E. A. B. (2021). A Comparative Analysis of Object Detection Metrics with a Companion Open-Source Toolkit. *Electronics*, *10*(3), 279. https://doi.org/10.3390/electronics10030279

Portmann, M., Galvagno-Erny, D., Lorenz, P., & Schacher, D. (2016). *Sonnendach.ch: Berechnung von Potenzialen in Gemeinden*. Bundesamt für Energie. https://pubdb.bfe.admin.ch/de/publication/download/8593

Remund, J., Albrecht, S., & Stickelberger, D. (2019). *Das Schweizer PV-Potenzial basierend auf jedem Gebäude*. 6.

Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(6), 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031

Souffer, I., Sghiouar, M., Sebari, I., Zefri, Y., Hajji, H., & Aniba, G. (2022). Automatic Extraction of Photovoltaic Panels from UAV Imagery with Object-Based Image Analysis and Machine Learning. In S. Bennani, Y. Lakhrissi, G. Khaissidi, A. Mansouri, & Y. Khamlichi (Eds.), *WITS 2020* (pp. 699–709). Springer Singapore.

Sutton, T., & Dassau, O. (2022). *QGIS Desktop*. QGIS Project. https://www.qgis.org/de/site/

Swissolar. (2022). *PV-Förderung*. https://www.swissolar.ch/topthemen/pv-foerderung/

UEVK, G. E. D. für U., Verkehr, Energie und Kommunikation UVEK. (2017). *Energiestrategie 2050*. https://www.uvek.admin.ch/uvek/de/home/energie/energiestrategie-2050.html

VESE. (2021). *PV Tarif*. Verband unabhängiger Energieversorger Schweiz. https://www.vese.ch/pvtarif/

Walch, A., Castello, R., Mohajeri, N., & Scartezzini, J.-L. (2020). Big data mining for the estimation of hourly rooftop photovoltaic potential and its uncertainty. *Journal of Applied Energy*, *262*. https://doi.org/10.1016/j.apenergy.2019.114404

Walch, A., Mohajeri, N., & Scartezzini, S. (2019). A critical comparison of methods to estimate solar rooftop photovoltaic potential in Switzerland. *Journal of Physics*. CISBAT 2019 | Climate Resilient Cities – Energy Efficiency & Renewables in the Digital Era, Lausanne, Switzerland.

Waleed, A. (2018). *Mask R-CNN for Object Detection and Instance Segmentation on Keras and TensorFlow*. Matterport. https://github.com/matterport/Mask_RCNN

9. Statement of truth

I hereby declare that I have written this thesis independently, without the assistance of third parties and using only the sources indicated, and that I will not hand over copies of this thesis to third parties without the written consent of the course director.

At the same time, all rights to the work are assigned to the Zurich University of Applied Sciences (ZHAW). The right to citation of authorship remains unaffected.
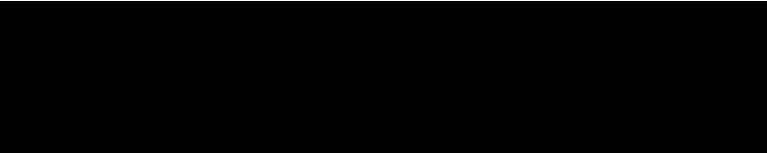
Tim Claude Bänziger, Degersheim 26.5.2022

10. Appendixes

10.1. Overview of deep learning based image segmentation algorithms

Method	Year	Supervision					Learning			Type			Modules		Description
		S	W	U	I	P	SO	MO	AD	SM	CL	IN	RNN	E-D	
Global Average Pooling	2013	✓					✓				✓				Object-specific soft segmentation
DenseCRF	2014					✓	✓			✓					Using CRF to boost segmentation
FCN	2015	✓					✓			✓					Fully convolutional layers
DeepMask	2015	✓						✓							Simultaneous learning for segmentation and classification
U-Net	2015	✓					✓			✓					Encoder-decoder with multi-scale feature concatenation
SegNet	2015	✓					✓			✓					Encoder-decoder with forwarding pooling indices
CRFasRNN	2015	✓						✓		✓			✓		Simulating CRFs as trainable RNN modules
Deep Parsing Network	2015	✓						✓							Using unshared kernels to incorporate higher-order dependency
BoxSup	2015		✓							✓					Using bounding box for weak supervision
SharpMask	2016	✓						✓			✓				Refined Deep Mask with multi-layer feature fusion
Attention to Scale	2016	✓					✓			✓					Fusing features from multi-scale inputs
Semantic Segmentation	2016	✓							✓	✓					Adversarial training for image segmentation
Conv LSTM and Spatial Inhibition	2016	✓						✓				✓	✓		Using spatial inhibition for instance segmentation
JULE	2016			✓				✓		✓			✓		Joint unsupervised learning for segmentation
ENet	2016	✓					✓			✓					Compact network for real-time segmentation
Instance-Aware Segmentation	2016	✓						✓				✓			Multi-task approach for instance segmentation
Mask R-CNN	2017	✓						✓		✓					Using region proposal network for segmentation
Large Kernel Matters	2017	✓					✓			✓					Using larger kernels for learning complex features
RefineNet	2017	✓					✓			✓					Multi-path refinement module for fine segmentation
PSPNet	2017	✓					✓			✓					Multi-scale pooling for scale-agnostic segmentation
Tiramisu	2017	✓					✓			✓					Dense Net 121-feature extractor
Image-to-Image Translation	2017	✓							✓	✓					Conditional GAN for translation image to segment maps
Instance Segmentation with Attention	2017	✓						✓				✓	✓		Attention modules for image segmentation
W-Net	2017			✓				✓		✓					Unsupervised segmentation using normalized cut loss
Polygon RNN	2017				✓		✓			✓			✓		Generating contours by RNN
Deep Layer Cascade	2017	✓					✓			✓					Multi-level approach to handle pixels of different complexity
Spatial Propagation Network	2017	✓					✓			✓					Refinement using linear label propagation
DeepLab	2018	✓					✓			✓					Atrous convolution ,spatial pooling pyramid, dense CRF
SegCaps	2018	✓						✓							Capsule networks for segmentation
Adversarial Collaboration	2018							✓							Adversarial collaboration between multiple networks
Superpixel Supervision	2018							✓							Using super-pixel refinement as supervisory signals
Deep Extreme Cut	2018				✓		✓			✓					Using extreme points for interactive segmentation
Two Stream Fusion	2019				✓					✓					Using image stream and interaction stream simultaneously
SegFast	2019	✓					✓			✓					Using depthwise separable convolution in SqueezeNet encoder

S, supervised; W, weakly supervised; U, unsupervised; I, interactive; P, partially supervised; SO, single-objective optimization; MO, Multi-objective optimization; AD, adversarial learning; SM, semantic segmentation; CL, class-specific segmentation; IN, instance segmentation; RNN, recurrent neural network modules; E-D, encoder-decoder architecture; GAN, generative adversarial network; CRF, conditional random field.

Figure 20: Chronological overview of the development of deep learning based image segmentation algorithms adapted from Ghosh et. al (2020)
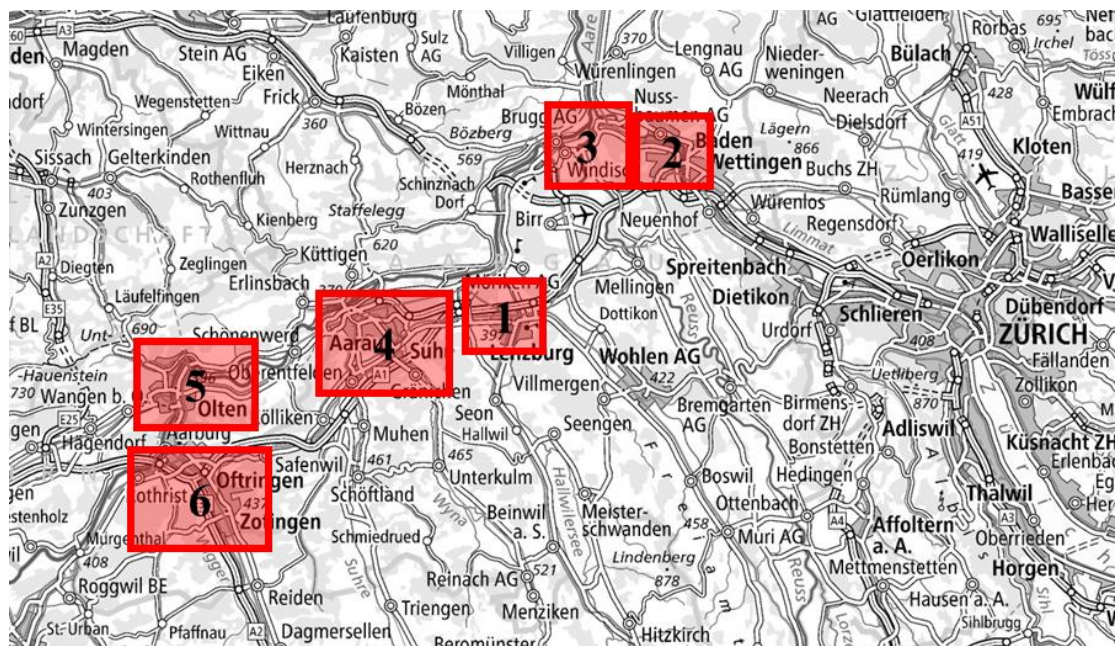
10.2.   Areas of image collection

Figure 21: Areas of image collection visualized (map from geoadmin.ch)

Table 8: Areas of image collection detailed information

Iteration	Saved in	Coordinates				Area
		UB_N	LB_N	UB_E	LB_E	
1	data_col_folder_1	1'251	1'247	2'657	2'650	Lenzburg
2	data_col_folder_2	1'261	1'256	2'668	2'663	Nussbaumen, Baden, Wettingen
3	data_col_folder_3	1'263	1'257	2'663	2'653	Brugg AG, Untersiggenthal,
4	data_col_folder_4	1'252	1'243	2'650	2'642	Aarau, Suhr
4	data_col_folder_5	1'247	1'243	2'643	2'633	Olten, Trimbach, Dulliken
4	data_col_folder_6	1'243	1'235	2'640	2'632	Oftringen, Rothrist, Zofingen

UB = upper bound, LB = lower bound, N = north, E = east

10.3. Mask R-CNN hyperparameters

Table 9: Mask R-CNN hyperparameter configuration

Mask R-CNN hyperparameter configuration	
BACKBONE	resnet101
BACKBONE_STRIDES	[4, 8, 16, 32, 64]
BATCH_SIZE	2
BBOX_STD_DEV	[0.1 0.1 0.2 0.2]
COMPUTE_BACKBONE_SHAPE	None
DETECTION_MAX_INSTANCES	100
DETECTION_MIN_CONFIDENCE	0.90
DETECTION_NMS_THRESHOLD	0.30
FPN_CLASSIF_FC_LAYERS_SIZE	1024
GPU_COUNT	1
GRADIENT_CLIP_NORM	5
IMAGES_PER_GPU	2
IMAGE_CHANNEL_COUNT	3
IMAGE_MAX_DIM	1024
IMAGE_META_SIZE	14
IMAGE_MIN_DIM	800
IMAGE_MIN_SCALE	0%
IMAGE_RESIZE_MODE	square
IMAGE_SHAPE	[1024 1024 3]
LEARNING_MOMENTUM	0.9
LEARNING_RATE	0.001
LOSS_WEIGHTS	{'rpn_class_loss': 1.0, 'rpn_bbox_loss': 1.0, 'mrcnn_class_loss': 1.0, 'mrcnn_bbox_loss': 1.0, 'mrcnn_mask_loss': 1.0}
MASK_POOL_SIZE	14
MASK_SHAPE	[28, 28]
MAX_GT_INSTANCES	100
MEAN_PIXEL	[123.7 116.8 103.9]
MINI_MASK_SHAPE	(56, 56)
NAME	PV
NUM_CLASSES	2
POOL_SIZE	7
POST_NMS_ROIS_INFERENCE	1000
POST_NMS_ROIS_TRAINING	2000
PRE_NMS_LIMIT	6000
ROI_POSITIVE_RATIO	0.33
RPN_ANCHOR_RATIOS	[0.5, 1, 2]
RPN_ANCHOR_SCALES	(32, 64, 128, 256, 512)
RPN_ANCHOR_STRIDE	1
RPN_BBOX_STD_DEV	[0.1 0.1 0.2 0.2]
RPN_NMS_THRESHOLD	0.7
RPN_TRAIN_ANCHORS_PER_IMAGE	256
STEPS_PER_EPOCH	100
TOP_DOWN_PYRAMID_SIZE	256
TRAIN_BN	False
TRAIN_ROIS_PER_IMAGE	200
USE_MINI_MASK	True
USE_RPN_ROIS	True
VALIDATION_STEPS	50
WEIGHT_DECAY	0.0001

10.4.   Mask R-CNN detailed loss curves
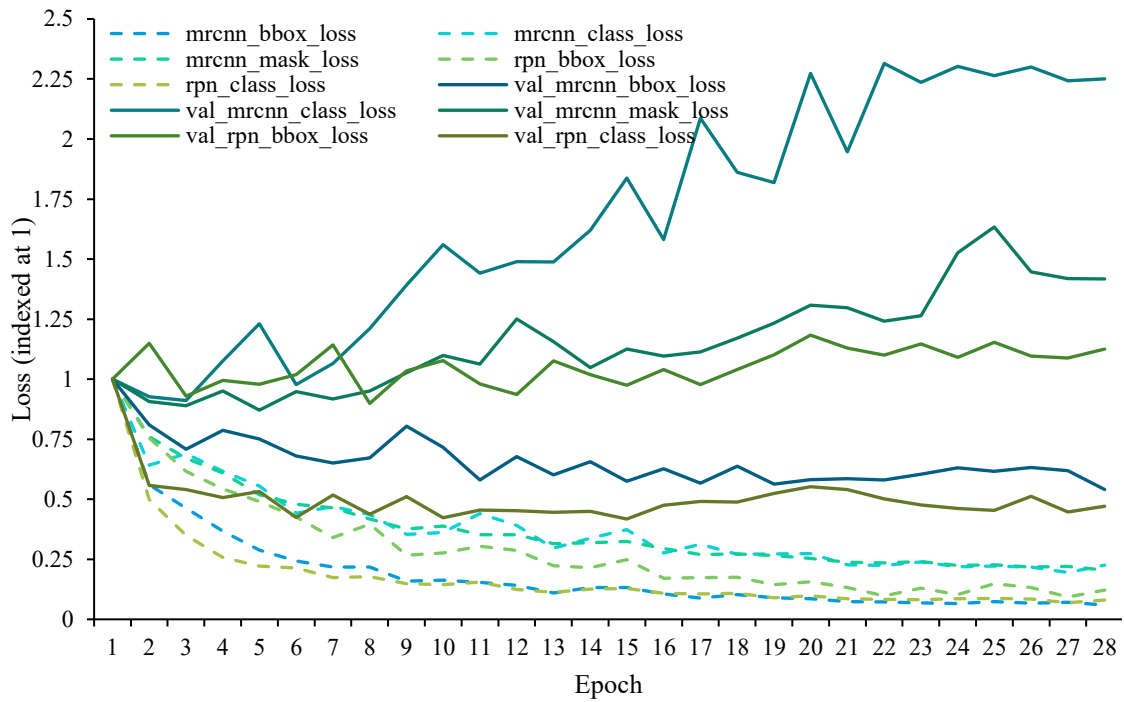
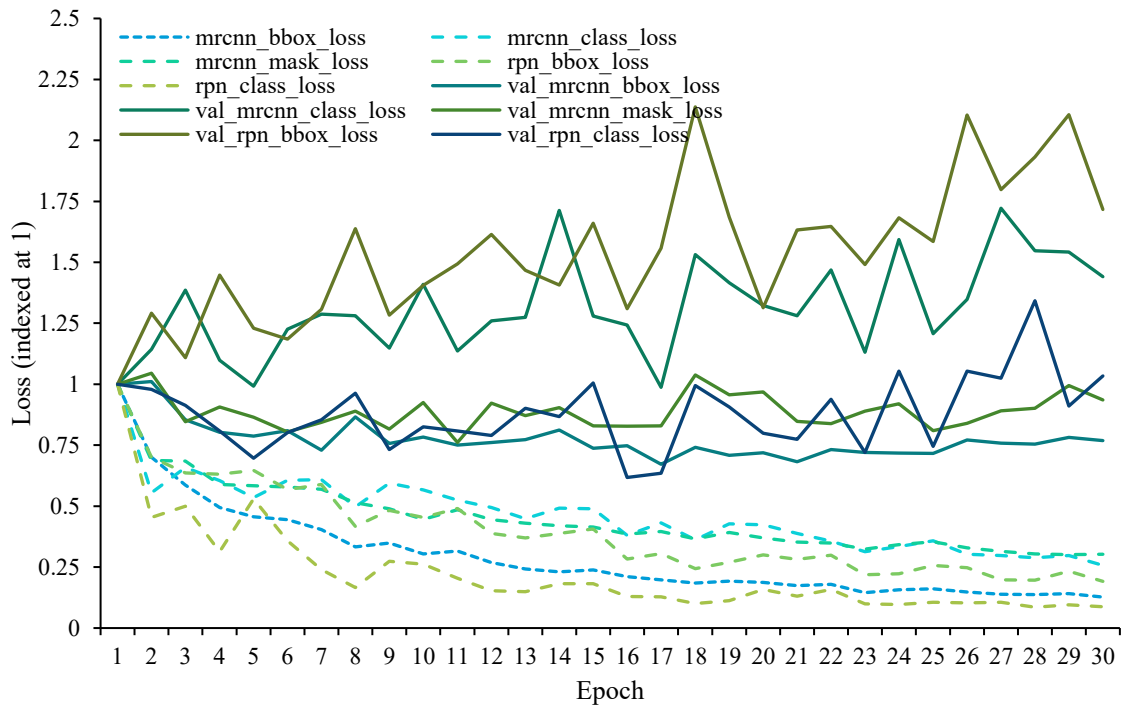*Figure 22: Mask R-CNN iteration_1 training and validation loss indexed at 1.0*

Figure 23: Mask R-CNN iteration_2 training and validation loss indexed at 1.0
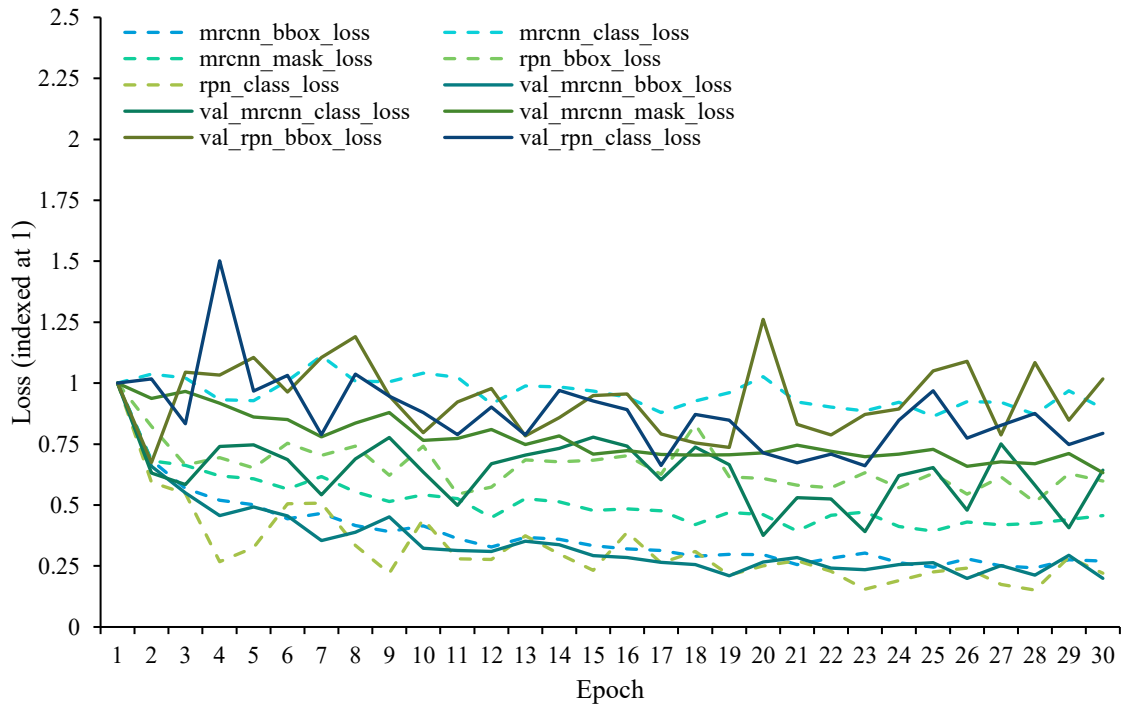
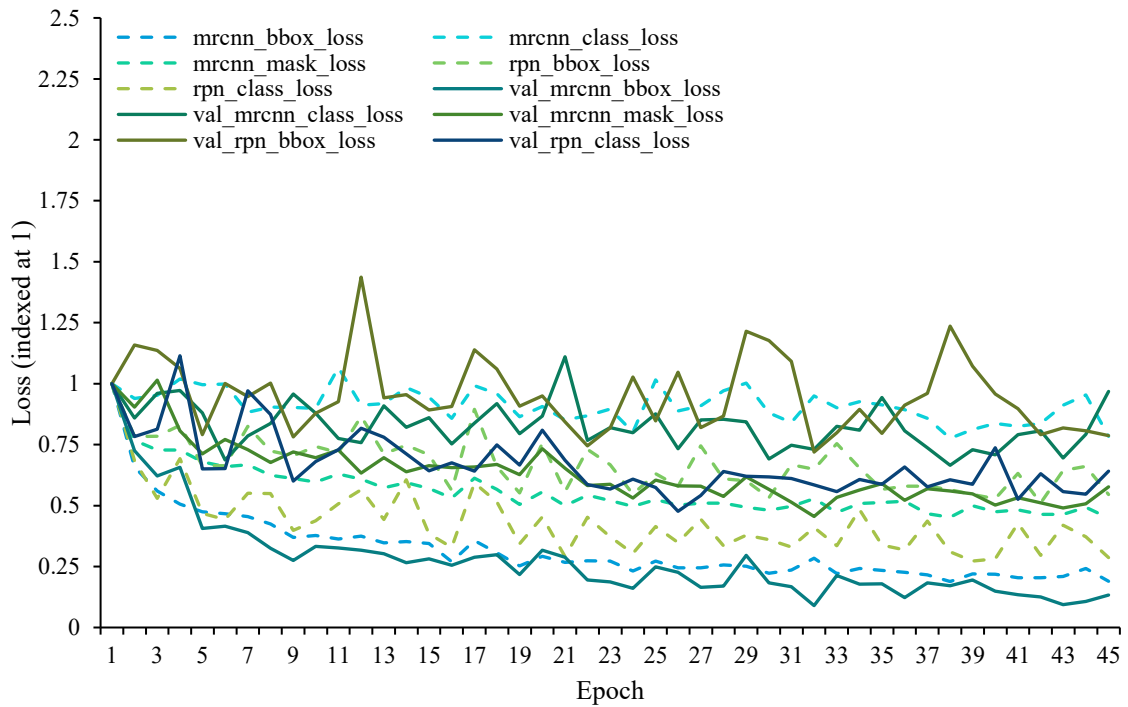*Figure 24: Mask R-CNN iteration_3 training and validation loss indexed at 1.0*

*Figure 25: Mask R-CNN iteration_4 training and validation loss indexed at 1.0*

10.5.   Mask R-CNN detailed performance analysis

Table 10: Mask R-CNN key performance metrics per iteration

Variable	iou_ score	mask_ precision	mask_ recall	mask_ accuracy	mask_ f1
average					
iteration_1	0.604	0.779	0.725	0.939	0.725
iteration_2	0.647	0.800	0.766	0.945	0.758
iteration_3	0.682	0.818	0.796	0.951	0.787
iteration_4	0.745	0.861	0.847	0.961	0.838
median					
iteration_1	0.661	0.872	0.819	0.957	0.796
iteration_2	0.705	0.893	0.860	0.966	0.827
iteration_3	0.769	0.911	0.865	0.970	0.870
iteration_4	0.816	0.931	0.896	0.974	0.899
standard deviation					
iteration_1	0.225	0.224	0.223	0.049	0.207
iteration_2	0.227	0.211	0.225	0.049	0.203
iteration_3	0.218	0.214	0.190	0.047	0.189
iteration_4	0.187	0.163	0.155	0.042	0.149
range					
iteration_1	0.004-0.918	0.075-1.000	0.004-0.989	0.768-0.996	0.008-0.957
iteration_2	0.005-0.919	0.008-1.000	0.012-0.983	0.742-0.996	0.010-0.958
iteration_3	0.000-0.949	0.000-1.000	0.000-0.971	0.742-0.997	0.000-0.974
iteration_4	0.126-0.943	0.128-1.000	0.205-0.981	0.738-0.998	0.224-0.971
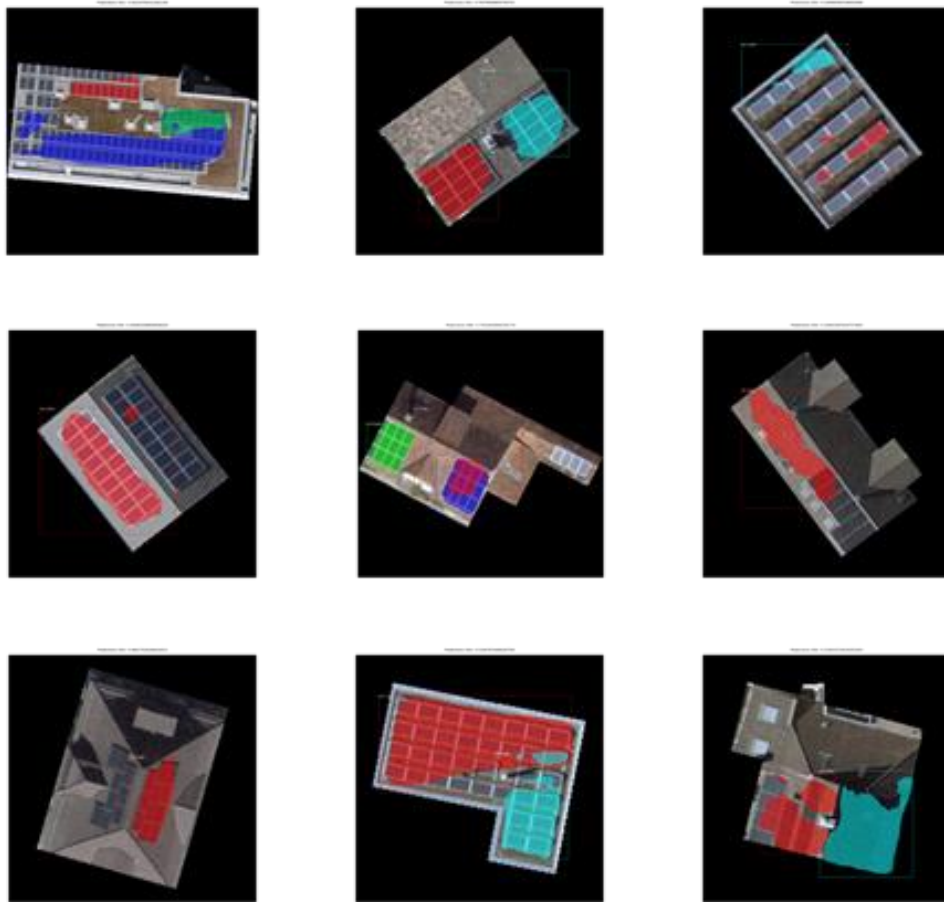
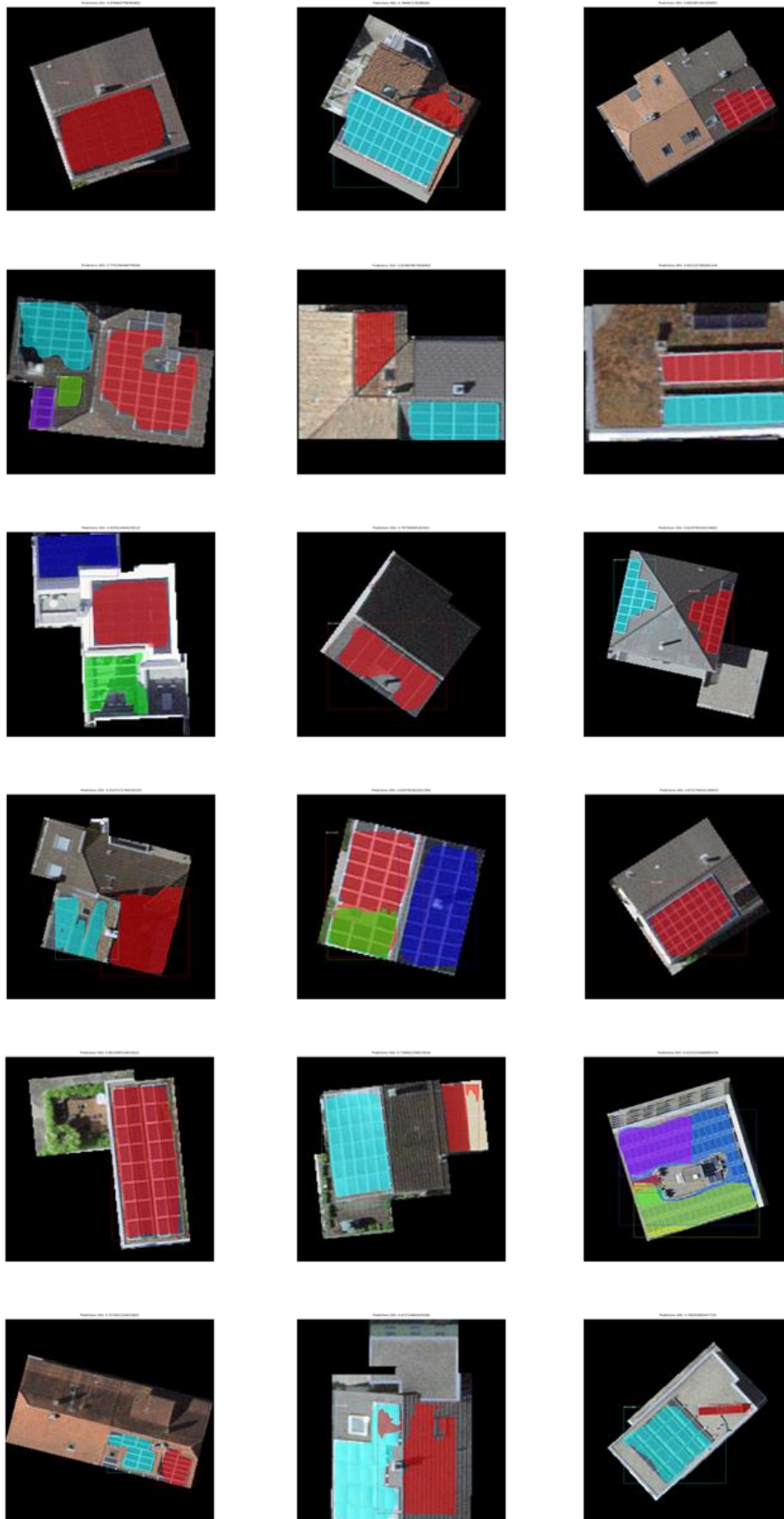*Figure 26: Mask R-CNN prediction masks (mask_pred) over input image from iteration 1 (randomly selected)*
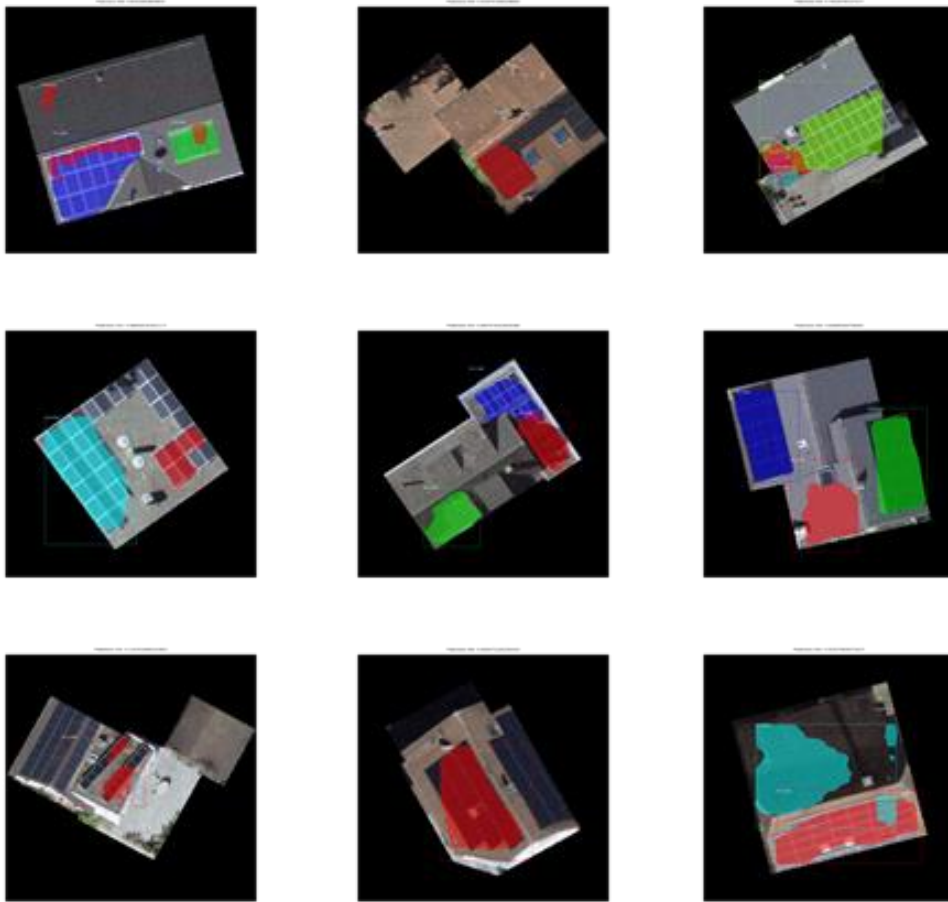
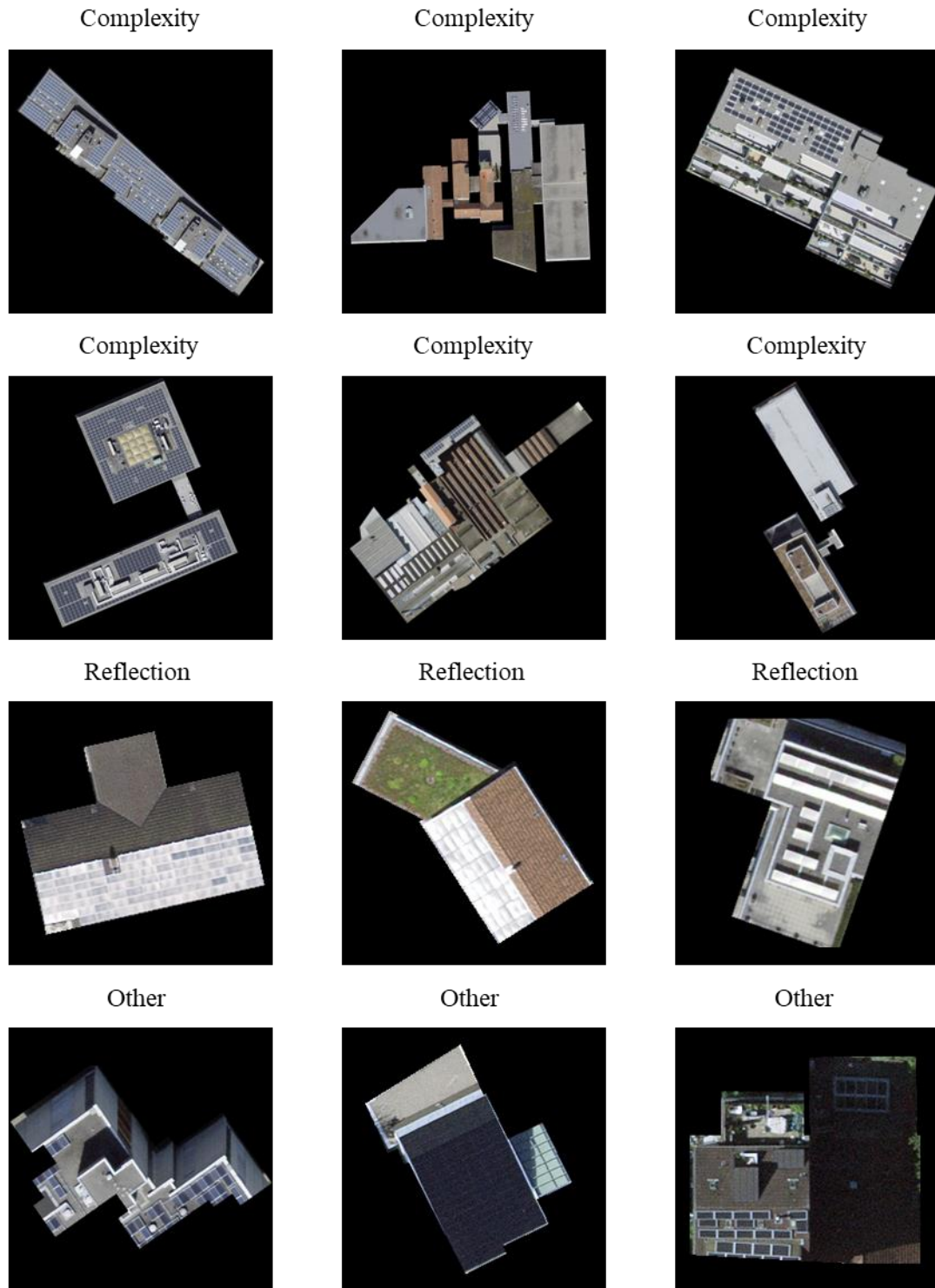*Figure 27: Mask R-CNN prediction masks (mask_pred) over input image from iteration 2 (randomly selected)*
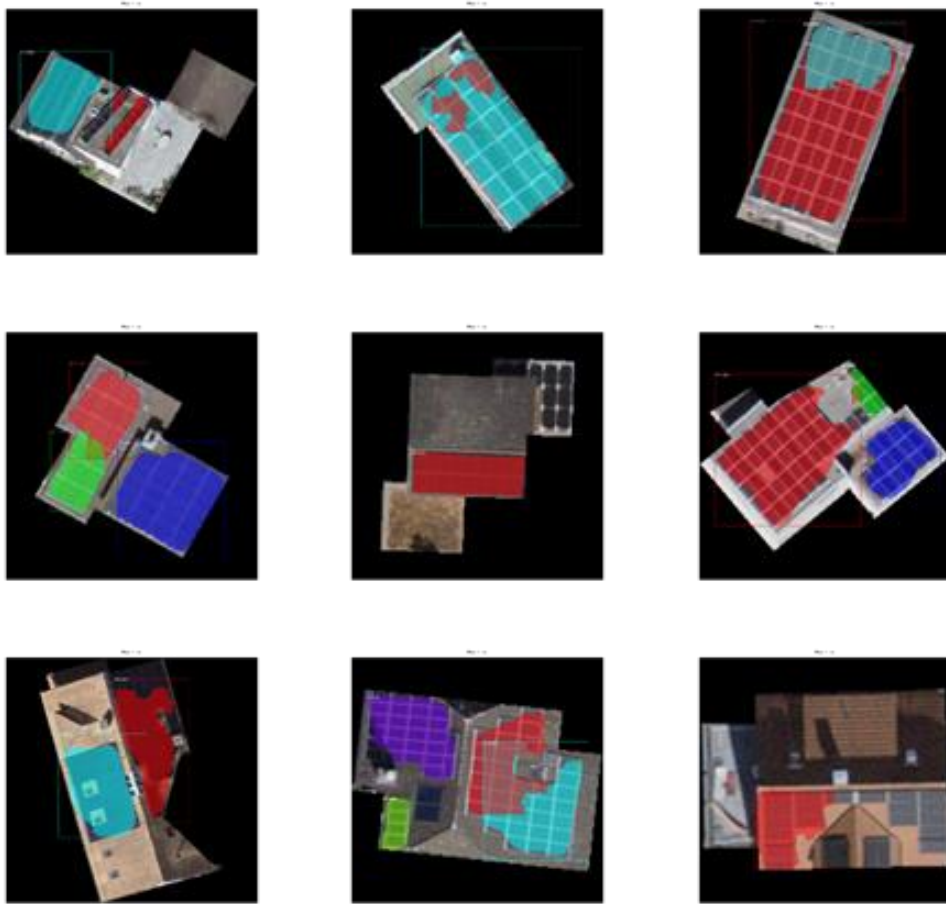
Figure 28: Mask R-CNN prediction masks (mask_pred) over input image from iteration 3 (randomly selected)

Figure 29: Mask R-CNN prediction masks (mask_pred) over input image from iteration 4 (randomly selected)

Figure 30: Rooftops where the Mask R-CNN experiences difficulties (randomly selected)

10.6. Dimension tests mask_gt / SolkatDach
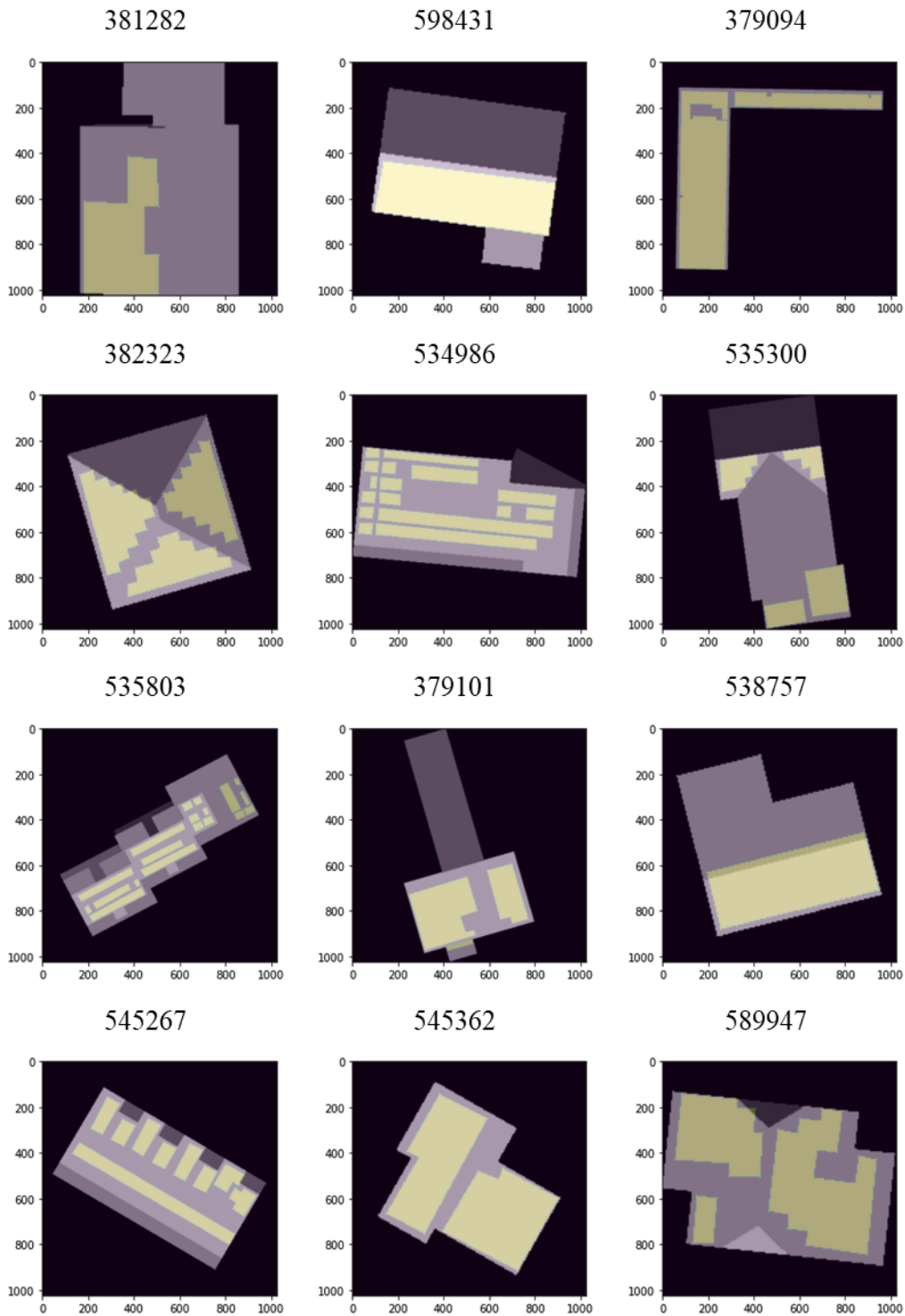
Figure 31: gt_masks over SolkatDach rasterized (with EGID) to inspect dimensional fit

10.7.   Comparison to SolkatDach
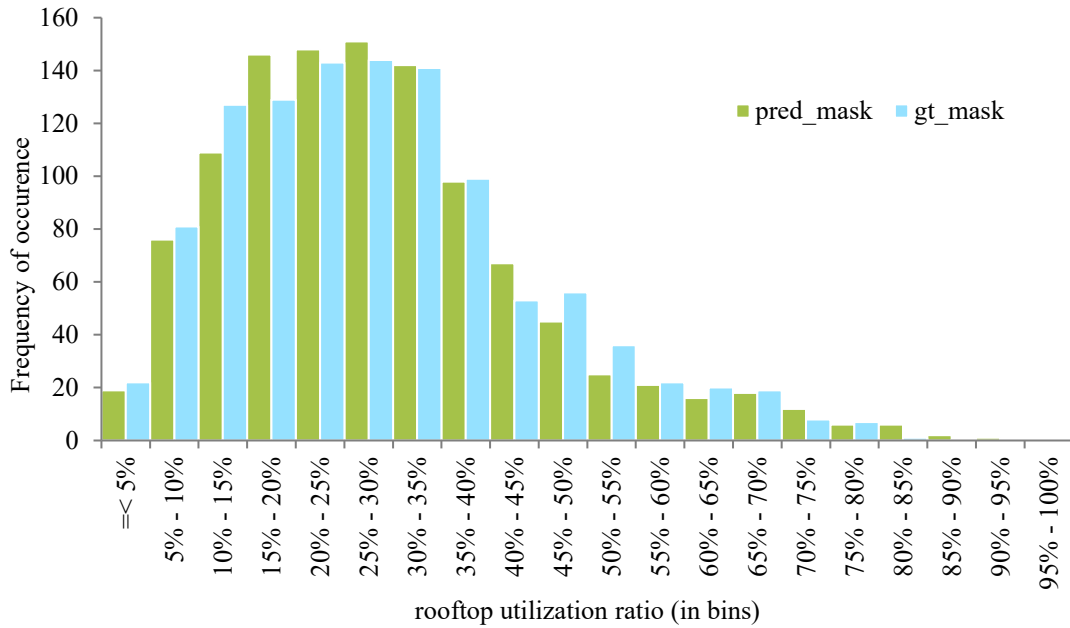
Figure 32: Distribution of available rooftop area utilization (comparison pred_mask inferenced on train and test dataset and gt_mask)
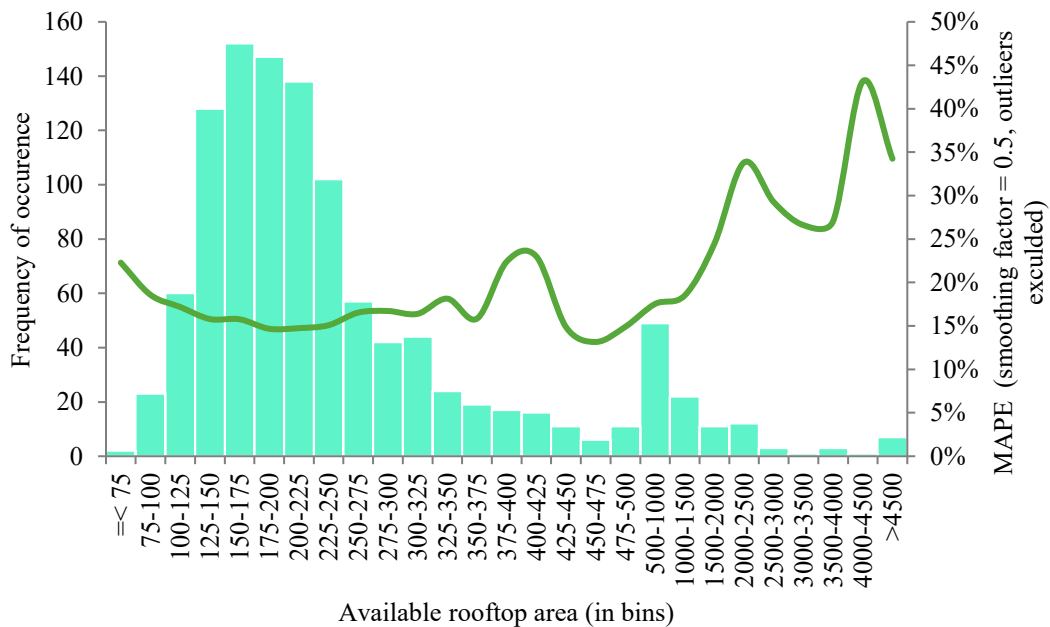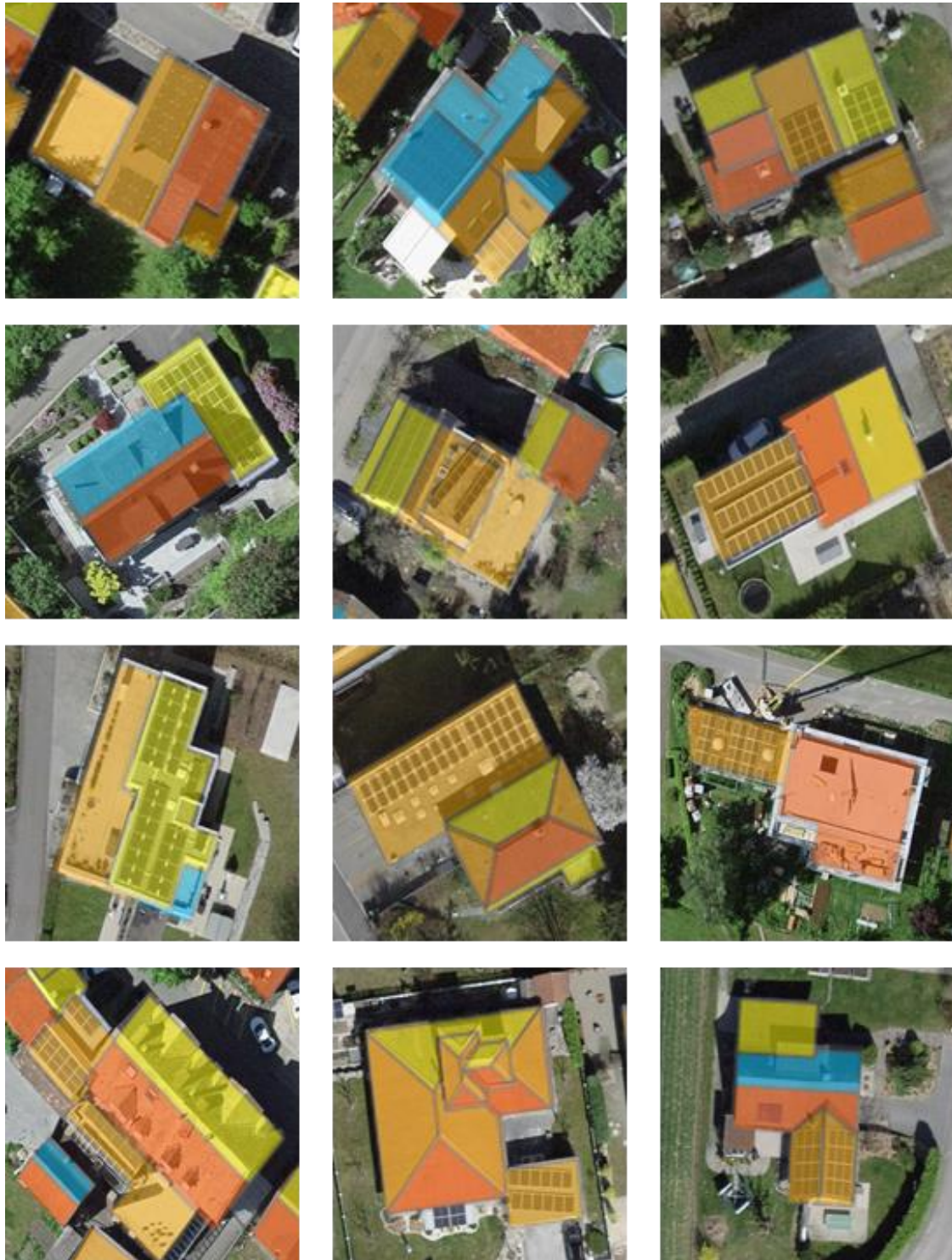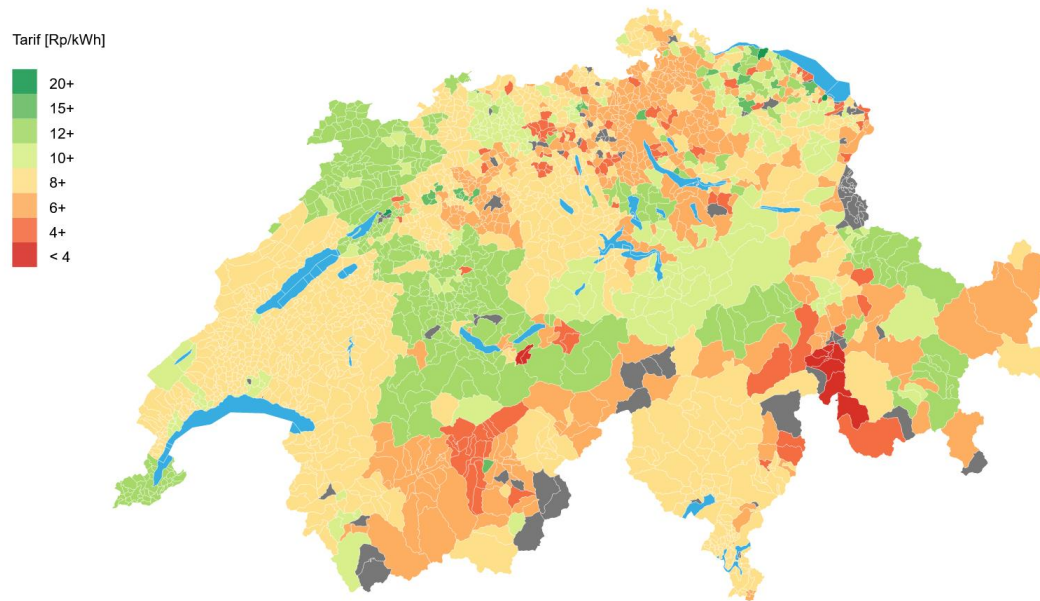

Figure 33: Distribution of available rooftop area in relation to rooftop area utilization MAPE between pred_mask inferenced on train and test dataset and gt_mask

10.8. Misplaced PV panels

Figure 34: Buildings with suboptimally placed PV panels detected by the model (categorized critical)

10.9.   Feed-in Tariffs for PV electricity

Figure 35: The feed-in tariffs for PV electricity across Switzerland in cents / KWh (communal level) (VESE, 2021)