

Remote Sensing based  
Renewable Energy System Recognition  
using Deep Learning

kumulative Dissertation  
zur Erlangung des Doktorgrades  
der Naturwissenschaften  
(Dr. rer. nat)

dem Fachbereich Geographie  
der Philipps-Universität Marburg  
vorgelegt von

Maximilian Kleebauer  
aus Bamberg

08. Oktober 2025

Am Fachbereich Geographie der Philipps-Universität Marburg  
(Hochschulkenziffer: 1180) am 08.10.2025 als Dissertation eingereicht.

Erstgutachter: Prof. Dr. Jörg Bendix  
Zweitgutachterin: Prof. Dr. Hanna Meyer

Tag der mündlichen Prüfung: 12.12.2025



## Preface

First of all, I would like to express my sincere gratitude to Jörg Bendix and Hanna Meyer, who not only provided me with numerous suggestions and advice, but also kindly undertook the formal examination of this thesis. My special thanks go to Christoph Reudenbach, who has accompanied me since my early studies and also supported me in many ways during this dissertation. I am also deeply grateful to Martin Braun, who, through his excellent advice and support, has greatly contributed to the success of the publications included in this dissertation. I would further like to thank my colleagues, in particular Jan Dobschinski and Doron Callies, for their invaluable guidance as well as their support in project management, administration, and funding acquisition. I am equally grateful to Daniel Horst, Carsten Pape, and Maximilian Pfennig for their great collaboration during various phases of this work and their support in developing the topic. Finally, I would like to thank my family, especially my parents. Without their invaluable support, this long journey would not have been possible. I would also like to express my deepest gratitude to my partner Marina, whose daily support and tireless encouragement were indispensable. Without her, this work would never have been completed.

During my studies, I was fascinated by how remote sensing and deep learning can complement each other to systematically collect and analyze information on a large scale. This allowed me to delve deeply into my areas of interest—remote sensing, deep learning, and energy systems—with the goal of contributing to these research topics. This was made possible primarily through several research projects, in particular the OASES project as part of the LEAP-RE program. Embedded in international collaborations with partners from South Africa, Algeria, Egypt, Finland and my German colleagues at Fraunhofer IEE and the University of Kassel, parts of the project results were published as research in peer-reviewed articles and thus form the basis of this dissertation, with additional academic support from the Department of Geography at Philipps University of Marburg.

## Abstract

The transition from fossil fuels to renewable energies is a key requirement for a sustainable energy future. Achieving this transformation requires not only technological innovation but also reliable, up-to-date, and spatially explicit data on existing infrastructures. For decentralized energy systems, precise information on the spatial distribution and technical characteristics of renewable energy systems is crucial. At the same time, the growing availability of satellite and aerial imagery provides unprecedented spatial and temporal resolutions, but extracting actionable insights at scale requires advanced deep-learning methods, particularly image-based recognition.

This cumulative dissertation is guided by three central research questions: (1) How can training data for deep-learning models be generated efficiently and reproducibly? (2) How robust are segmentation approaches across different spatial resolutions and geographic contexts? (3) How can the positional accuracy of existing datasets of renewable energy systems be improved?

The results show that registry and building data can be combined to automatically generate reproducible training datasets, reducing reliance on manual labeling. Multi-resolution segmentation improves robustness and transferability, enabling reliable recognition of photovoltaic systems across diverse regions and scales. Object-detection models systematically correct positional errors in wind-turbine datasets, enhancing their accuracy and reliability.

The dissertation makes several key contributions. Methodologically, it advances automated training-data generation, multi-resolution segmentation, and object detection-driven coordinate correction. Applied contributions include open, validated datasets on photovoltaic systems and wind turbines, strengthening the empirical foundation for modelling, planning, and scenario development. Together, these contributions demonstrate how remote sensing and deep learning can be combined to improve data quality and support evidence-based energy-transition strategies.

## Zusammenfassung

Der Übergang von fossilen Brennstoffen zu erneuerbaren Energien ist Voraussetzung für eine nachhaltige Energieversorgung. Dafür sind neben technologischen Innovationen vor allem verlässliche, aktuelle und räumlich detaillierte Daten notwendig. Gerade in dezentralen Energiesystemen sind präzise Informationen über Verteilung und technische Eigenschaften unverzichtbar. Parallel liefert die wachsende Verfügbarkeit von Erdbeobachtungsdaten Bildmaterial in bisher unerreichter Auflösung. Um zuverlässige Erkenntnisse in großem Maßstab zu gewinnen, bieten fortschrittliche Deep-Learning-Methoden enormes Potenzial, insbesondere bildbasierte Erkennungsmethoden.

Diese kumulative Dissertation konzentriert sich auf drei zentrale Forschungsfragen: (1) Wie können Trainingsdaten für Deep-Learning-Modelle effizient und reproduzierbar generiert werden? (2) Wie robust sind Segmentierungsansätze über verschiedene räumliche Auflösungen und geografische Kontexte hinweg? (3) Wie kann die Positionsgenauigkeit bestehender Datensätze zu erneuerbaren Energiesystemen verbessert werden?

Die Ergebnisse zeigen, dass sich aus Register- und Gebäudedaten automatisch reproduzierbare Trainingsdaten ableiten lassen, wodurch der Bedarf an manueller Arbeit deutlich sinkt. Segmentierung über mehrere Auflösungen hinweg erhöht Robustheit und Übertragbarkeit und ermöglicht eine genaue Erkennung von Photovoltaikanlagen über verschiedene Regionen und Skalen. Auf Objekterkennung basierende Modelle korrigieren systematisch Lagefehler in Windkraft-Datensätzen und steigern so Genauigkeit und Nutzbarkeit.

Die Arbeit leistet mehrere Hauptbeiträge. Methodisch werden Ansätze zur halb-automatischen Trainingsdatenerzeugung, zur Segmentierung über mehrere Auflösungen und zur erkenntungsgestützten Koordinatenkorrektur weiterentwickelt. Anwendungsseitig entstehen offene, validierte Datensätze zu Photovoltaikanlagen und Windturbinen, die eine empirische Grundlage für Modellierung, Planung und Szenarienbildung stärken. So zeigt die Dissertation, wie die Verbindung von Fernerkundung und Deep Learning die Datenqualität verbessert und evidenzbasierte Strategien der Energiewende unterstützt.

---

# Contents

|  |            |
|--|------------|
| <b>Preface</b>                                       | <b>i</b>   |
| <b>Abstract</b>                                      | <b>ii</b>  |
| <b>Zusammenfassung</b>                               | <b>iii</b> |
| <b>Contents</b>                                      | <b>iv</b>  |
| <b>List of Figures</b>                               | <b>vii</b> |
| <b>List of Tables</b>                                | <b>ix</b>  |
| <b>List of Abbreviations</b>                         | <b>xii</b> |
| <b>Introduction</b>                                  | <b>1</b>   |
| 1.1 Remote Sensing . . . . .                         | 3          |
| 1.2 Deep Learning . . . . .                          | 5          |
| 1.3 Renewable Energy System Recognition . . . . .    | 7          |
| 1.4 Central Research Hypotheses . . . . .            | 11         |
| 1.5 Core Structure of the Dissertation . . . . .     | 12         |
| 1.6 Research Context . . . . .                       | 14         |
| <b>Semi-Automatic Generation of Training Samples</b> | <b>17</b>  |
| 2.1 Abstract . . . . .                               | 19         |
| 2.2 Introduction . . . . .                           | 19         |
| 2.3 Materials . . . . .                              | 21         |
| 2.3.1 Georeferenced Address Data . . . . .           | 21         |
| 2.3.2 Official House Surroundings Germany . . . . .  | 21         |
| 2.3.3 German Plant Register . . . . .                | 22         |
| 2.3.4 Digital Orthophotos . . . . .                  | 22         |
| 2.4 Methodology . . . . .                            | 22         |

---

|  |   |           |
|--|---|-----------|
| 2.4.1  | Data Preprocessing . . . . .                          | 22        |
| 2.4.2  | Deep Learning Approach . . . . .                      | 24        |
| 2.5  | Results . . . . .                                     | 27        |
| 2.6  | Discussion . . . . .                                  | 30        |
| 2.7  | Conclusions . . . . .                                 | 32        |
|  | Bibliography . . . . .                                | 33        |
| <b>Multi-Resolution Segmentation of Solar Photovoltaic Systems</b> |   | <b>37</b> |
| 3.1  | Abstract . . . . .                                    | 39        |
| 3.2  | Introduction . . . . .                                | 39        |
| 3.3  | Materials . . . . .                                   | 43        |
| 3.4  | Methods . . . . .                                     | 44        |
| 3.4.1  | Data Preprocessing . . . . .                          | 44        |
| 3.4.2  | Model . . . . .                                       | 47        |
| 3.4.3  | Model Training . . . . .                              | 48        |
| 3.5  | Results . . . . .                                     | 55        |
| 3.5.1  | Hyperparameter Tuning . . . . .                       | 55        |
| 3.5.2  | Single-Resolution Training . . . . .                  | 57        |
| 3.5.3  | Multi-Resolution Training . . . . .                   | 58        |
| 3.5.4  | Final Model Configuration . . . . .                   | 60        |
| 3.6  | Discussion . . . . .                                  | 61        |
| 3.7  | Conclusions . . . . .                                 | 64        |
| 3.8  | Appendix . . . . .                                    | 65        |
|  | Bibliography . . . . .                                | 70        |
| <b>Globally Scalable, QGIS-Integrated Workflow</b>                 |   | <b>76</b> |
| 4.1  | Abstract . . . . .                                    | 78        |
| 4.2  | Introduction . . . . .                                | 78        |
| 4.3  | Study Area . . . . .                                  | 80        |
| 4.4  | Materials and Methods . . . . .                       | 81        |
| 4.4.1  | PV-Segmentation Model . . . . .                       | 81        |
| 4.4.2  | Deepness: Deep Neural Remote Sensing Plugin . . . . . | 81        |
| 4.4.3  | Google Satellite Basemap . . . . .                    | 82        |
| 4.4.4  | Validation and Solar Capacity Estimation . . . . .    | 83        |

---

|     |  |            |
|-----|--|------------|
| 4.5 | Results . . . . .  | 83         |
| 4.6 | Discussion . . . . .   | 86         |
| 4.7 | Conclusions . . . . .  | 87         |
|     | Bibliography . . . . .   | 88         |
|     | <b>Enhancing Wind Turbine Location Accuracy</b>                    | <b>91</b>  |
| 5.1 | Abstract . . . . .   | 93         |
| 5.2 | Introduction . . . . .   | 93         |
| 5.3 | Materials . . . . .  | 95         |
| 5.4 | Methods . . . . .  | 96         |
|     | 5.4.1 Data preprocessing . . . . .                                 | 96         |
|     | 5.4.2 Deep learning approach . . . . .                             | 98         |
| 5.5 | Results . . . . .  | 99         |
| 5.6 | Discussion . . . . .   | 101        |
| 5.7 | Conclusions . . . . .  | 103        |
|     | Bibliography . . . . .   | 104        |
|     | <b>A Wind Turbines Dataset for South Africa</b>                    | <b>107</b> |
| 6.1 | Abstract . . . . .   | 109        |
| 6.2 | Introduction . . . . .   | 109        |
| 6.3 | Materials . . . . .  | 113        |
|     | 6.3.1 OpenStreetMap . . . . .                                      | 113        |
|     | 6.3.2 Google Satellite Data and Bing Image Data . . . . .          | 114        |
|     | 6.3.3 Core Energy Market Data Register . . . . .                   | 114        |
|     | 6.3.4 Digital Orthophotos . . . . .                                | 115        |
| 6.4 | Methods . . . . .  | 116        |
|     | 6.4.1 Training data preprocessing . . . . .                        | 116        |
|     | 6.4.2 Deep Learning Approach . . . . .                             | 117        |
|     | 6.4.3 South Africa Wind Turbine Pre-Dataset . . . . .              | 120        |
|     | 6.4.4 South Africa's wind turbines coordinate correction . . . . . | 121        |
|     | 6.4.5 Additional attribute Enrichment . . . . .                    | 121        |
| 6.5 | Results . . . . .  | 122        |
|     | 6.5.1 Performance and Results of Deep Learning Training . . . . .  | 122        |
|     | 6.5.2 OSM Data Extraction . . . . .                                | 124        |

|                     |   |            |
|---------------------|---|------------|
| 6.5.3               | Coordinate Correction . . . . .                           | 125        |
| 6.5.4               | Wind turbine dataset . . . . .                            | 127        |
| 6.5.5               | Validation against Official Capacity Figures . . . . .    | 135        |
| 6.6                 | Discussion . . . . .                                      | 136        |
| 6.7                 | Conclusions . . . . .                                     | 140        |
|                     | Bibliography . . . . .                                    | 141        |
| <b>Conclusions</b>  |   | <b>146</b> |
| 7.1                 | Synthesis of the Central Research Hypotheses . . . . .    | 147        |
| 7.2                 | Contribution to the Research Field . . . . .              | 150        |
| 7.3                 | Innovations in Remote Sensing and Deep Learning . . . . . | 151        |
| 7.4                 | Practical Applications . . . . .                          | 153        |
| 7.5                 | Limitations and Challenges . . . . .                      | 154        |
| 7.6                 | Outlook and Future Research . . . . .                     | 156        |
| 7.6.1               | Methodological Perspective . . . . .                      | 156        |
| 7.6.2               | Data Enrichment and Practical Applications . . . . .      | 157        |
| 7.6.3               | Long-term Perspectives . . . . .                          | 158        |
| 7.7                 | Overall Conclusions . . . . .                             | 159        |
| <b>Bibliography</b> |   | <b>168</b> |
| <b>Appendix</b>     |   | <b>169</b> |
| 8.1                 | Curriculum Vitæ . . . . .                                 | 169        |
| 8.2                 | Publications . . . . .                                    | 171        |
| 8.3                 | Eidesstattliche Erklärung . . . . .                       | 173        |

---

## List of Figures

|     |  |     |
|-----|--|-----|
| 2.1 | The figure shows the workflow schematically. . . . .                     | 21  |
| 2.2 | Loss values during first training . . . . .                              | 28  |
| 2.3 | AP values during first training . . . . .                                | 28  |
| 2.4 | Regression boxes within building geometries . . . . .                    | 29  |
| 2.5 | Loss values during second training . . . . .                             | 30  |
| 2.6 | AP values during second training . . . . .                               | 30  |
| 2.7 | Applications True Positive examples . . . . .                            | 30  |
| 2.8 | Applications False Positive examples . . . . .                           | 31  |
| 3.1 | Training images and resolution levels . . . . .                          | 46  |
| 3.2 | Predictions by resolution and training dataset . . . . .                 | 59  |
| 3.3 | Comparison of single- and multi-resolution predictions . . . . .         | 62  |
| 3.4 | Illustration of the different hyperparameter tuning results . . . . .    | 68  |
| 4.1 | Workflow for segmentation and capacity estimation . . . . .              | 79  |
| 4.2 | Segmentation results for Reggane PV at multiple resolutions . . . . .    | 84  |
| 4.3 | Segmentation comparison: Reggane PV at 20–320 cm resolution . . . . .    | 84  |
| 4.4 | Segmented PV systems in Piat region . . . . .                            | 85  |
| 5.1 | Comparison of static and centroid-based image samples cropping . . . . . | 97  |
| 5.2 | Samples based on their suitability for training . . . . .                | 98  |
| 5.3 | Loss curves and average precision . . . . .                              | 99  |
| 5.4 | Examples of False Positives and Negatives . . . . .                      | 100 |
| 5.5 | Examples of True Positive Detections . . . . .                           | 101 |
| 6.1 | Workflow for Wind Dataset Creation . . . . .                             | 112 |

|      |   |     |
|------|---|-----|
| 6.2  | Static Cutting Method for Training Images . . . . .                                     | 117 |
| 6.3  | Samples based on their suitability for training. . . . .                                | 118 |
| 6.4  | Average Precision and different Losses during training . . . . .                        | 123 |
| 6.5  | Examples of False Positives and False Negatives . . . . .                               | 123 |
| 6.6  | True Positives from the DOP-based Application . . . . .                                 | 124 |
| 6.7  | True Positives in Bing Imagery during Wind Detection . . . . .                          | 126 |
| 6.8  | True Positives in Google Imagery during Wind Detection . . . . .                        | 126 |
| 6.9  | Histogram of wind turbine location errors based on Bing and<br>Google imagery . . . . . | 128 |
| 6.10 | Summary statistics of key parameters of South African wind<br>farms . . . . .           | 130 |
| 6.11 | Development of wind power capacity in South Africa by year . . . . .                    | 130 |
| 6.12 | Installed Capacity by Wind Farm . . . . .   | 131 |
| 6.13 | Number of Wind Turbines by Wind Farm . . . . .  | 132 |
| 6.14 | Capacity per Wind Turbine by Wind Farm . . . . .  | 133 |
| 6.15 | Spatial distribution of all existing wind turbines in South Africa                      | 135 |
| 6.16 | Spatial distribution of wind energy infrastructure by munici-<br>pality . . . . .       | 136 |

---

# List of Tables

|      |  |     |
|------|--|-----|
| 2.1  | Formulas of the evaluation, precision, recall, and the average precision. . . . .  | 26  |
| 2.2  | Number of images, the area covered by the images, and the number of buildings used for validation. . . . .                       | 27  |
| 2.3  | True Positive, False Positive, False Negative, and True Negative as well as the precision and the recall for both test datasets. | 29  |
| 3.1  | Summary of the information on the various datasets. . . . .  | 44  |
| 3.2  | Ratio of pixels containing solar PV systems per dataset. . . . .   | 45  |
| 3.3  | Hyperparameter search space metrics. . . . .   | 48  |
| 3.4  | Brief overview of the best hyperparameters tuning results. . . . .   | 55  |
| 3.5  | Metrics from batch size optimizations. . . . .   | 56  |
| 3.6  | Metrics for different stride values. . . . .   | 56  |
| 3.7  | Single-resolution models F1-Score) . . . . .   | 57  |
| 3.8  | Multi-resolution models F1-Score) . . . . .  | 58  |
| 3.9  | Side-by-side F1-Score from validation between single-resolution trained networks and the final multi-resolution trained network. | 60  |
| 3.10 | Hyperparameter tuning results. . . . .   | 66  |
| 3.11 | <i>Cont.</i> . . . . .   | 67  |
| 3.12 | Single-resolution models IoU performance . . . . .   | 67  |
| 3.13 | Multi-resolution models IoU performance) . . . . .   | 69  |
| 4.1  | Comparison of segmented and official PV capacities . . . . .   | 85  |
| 5.1  | Distances between MaStR and validated coordinates. . . . .   | 102 |

|     |  |     |
|-----|--|-----|
| 6.1 | Comparison of Wind Turbine Confidence Scores (Bing vs. Google)             | 125 |
| 6.2 | Comparison of Wind Turbine Coordinate Deviations in South Africa . . . . . | 127 |
| 6.3 | Summary of Wind Turbines in South Africa . . . . .                         | 129 |
| 6.4 | Overview of wind farms in South Africa by province. . . . .                | 134 |

# List of Abbreviations

The following abbreviations are used throughout this thesis:

|      |  |
|------|--|
| AI   | Artificial Intelligence  |
| AP   | Average Precision  |
| ASPP | Atrous Spatial Pyramid Pooling                                   |
| AU   | African Union  |
| BCE  | Binary Cross-Entropy   |
| BKG  | Federal Agency for Cartography and Geodesy (Germany)             |
| CARE | Collective Benefit, Authority to Control, Responsibility, Ethics |
| CDSE | Copernicus Data Space Ecosystem                                  |
| CE   | Cross-Entropy  |
| CIR  | Colour Infrared  |
| CNN  | Convolutional Neural Network                                     |
| COCO | Common Objects in Context  |
| CV   | Computer Vision  |
| DL   | Deep Learning  |
| DOP  | Digital Orthophotos  |
| EO   | Earth Observation  |
| ESA  | European Space Agency  |
| EU   | European Union   |
| F1   | F1 Score   |
| FAIR | Findable, Accessible, Interoperable, Reusable                    |
| FCN  | Fully Convolutional Network                                      |
| FN   | False Negative   |
| FP   | False Positive   |

|        |   |
|--------|---|
| FPN    | Feature Pyramid Network   |
| GAN    | Generative Adversarial Network  |
| GA     | Georeferenced Address Data  |
| GADM   | Global Administrative Areas   |
| GEE    | Google Earth Engine   |
| GFLOPS | Giga Floating Point Operations Per Second   |
| GPU    | Graphics Processing Unit  |
| HPC    | High-Performance Computing  |
| HU-DE  | Official House Surroundings Germany   |
| IGN    | National Institute of Geographic and Forest Information (France)  |
| IoU    | Intersection over Union   |
| IPP    | Independent Power Producer  |
| IPPPP  | Independent Power Producer Procurement Programme  |
| IRENA  | International Renewable Energy Agency   |
| IRP    | Integrated Resource Plan  |
| LiDAR  | Light Detection and Ranging   |
| LLM    | Large Language Model  |
| LSTM   | Long Short-Term Memory  |
| MaStR  | Marktstammdatenregister (German Core Energy Market Data Register)   |
| mAP    | mean Average Precision  |
| ML     | Machine Learning  |
| MLC    | Maximum Likelihood Classification   |
| MLP    | Multilayer Perceptron   |
| MPC    | Microsoft Planetary Computer  |
| MSE    | Mean Squared Error  |
| NDVI   | Normalized Difference Vegetation Index  |
| NIR    | Near-Infrared   |
| OASES  | Development and Demonstration of a Sustainable Open Access<br>AU-EU Ecosystem for Energy System Modelling |
| OSM    | OpenStreetMap   |
| PGC    | Provincial Geomatics Center   |
| PV     | Photovoltaic  |
| QGIS   | Quantum Geographic Information System   |

|         |   |
|---------|---|
| RE      | Renewable Energy                            |
| RES     | Renewable Energy System(s)                  |
| RF      | Random Forest                               |
| RMSProp | Root Mean Square Propagation                |
| RNN     | Recurrent Neural Network                    |
| RS      | Remote Sensing                              |
| SAM     | Segment Anything Model                      |
| SDG     | Sustainable Development Goal                |
| SGD     | Stochastic Gradient Descent                 |
| SPP     | Spatial Pyramid Pooling                     |
| SVM     | Support Vector Machine                      |
| TN      | True Negative                               |
| TP      | True Positive                               |
| TPU     | Tensor Processing Unit                      |
| UAV     | Unmanned Aerial Vehicle                     |
| UN      | United Nations                              |
| VGG     | Visual Geometry Group                       |
| ViT     | Vision Transformer                          |
| VHR     | Very High Resolution                        |
| VOC     | Visual Object Classes                       |
| WGS84   | World Geodetic System 1984                  |
| Zenodo  | Open-access repository for research outputs |

# Chapter 1

## Introduction

The expansion of renewable energies is a key component for a sustainable and clean energy future [1]. Achieving global climate goals while ensuring security of supply requires not only technical advances in energy generation, but also precise and up-to-date data on existing infrastructures [2, 3]. In the context of a decentralized energy transition – as is being pursued in Germany and many other countries – detailed information on the locations, capacities and spatial distribution of renewable energy systems is particularly crucial. Such comprehensive and spatially explicit data are a prerequisite for accurate planning, analysis and modelling of future energy systems [4]. While meteorological and geophysical fundamentals such as solar radiation and wind resources are already being intensively studied and increasingly made available worldwide [5], comparable structured, quality-assured information on the locations, types and technical characteristics of renewable energy systems is still lacking in many regions. In many cases, such data remain only partially accessible.

In recent years, the number of satellites orbiting the earth has increased rapidly. They provide data for a wide range of applications - from global communication and navigation to earth observation [6, 7]. Technological progress and falling launch costs have led to numerous countries and private sector players now operating their own satellites. As a result, large quantities of high-resolution earth observation data are now available, often at low cost or even free of charge [8]. At the same time, the amount of data creates new challenges for processing and analysis. This results in enormous potential for automated processes such as deep learning.

Deep Learning (DL) methods are particularly suitable for the automated processing of large, complex amounts of data. Deep learning methods have already proven themselves many times over in data-driven domains such as science, business, and government [9]. In recent years, they have also emerged as a key technology in remote sensing, where they are increasingly applied to address large-scale societal challenges such as climate change and urbanization. Such techniques can be used to identify recurring patterns and extract information that would be difficult or impossible to access by conventional methods [10]. The combination of remote sensing and deep learning opens up

a wide range of possibilities for the systematic recognition of renewable energy systems. The recognition of photovoltaic systems and wind turbines on aerial and satellite images for spatial localization has already been the subject of numerous studies using different image data, resolutions, and techniques [11]. A more detailed review can be found in the following sections remote sensing 1.1, deep learning 1.2, and renewable energy system recognition 1.3.

## 1.1 Remote Sensing

Remote sensing refers to the acquisition of information about the Earth's surface without direct physical contact, typically by recording reflected or emitted electromagnetic radiation [12]. This can be achieved by passive sensors, which measure naturally reflected or emitted radiation [13], or by active sensors, which emit their own signal and record the backscattered response such as radar systems [14]. Data resulting from remote sensing is characterized by three key resolution dimensions: spectral, spatial, and temporal. The spectral resolution defines the number and width of spectral bands recorded by a sensor and allows the differentiation of materials based on their reflectance properties [15]. The spatial resolution refers to the ground size of a pixel and determines the level of detail that can be captured in an image [16]. The temporal resolution describes the revisit time, i.e. the interval between successive acquisitions of the same location [17].

In contrast to the technical fundamentals described above, remote sensing can be understood more generally as a multi-stage process that includes the radiation source, its interaction with the atmosphere and the surface, sensor detection, data transmission, and subsequent processing, interpretation, and application [18]. In practice, however, the term remote sensing is often used simply in connection with image analysis and interpretation. Much of the practical work in this field involves the digital processing of image data from which information is extracted to support scientific and applied contexts [16]. In addition to the term remote sensing, the more comprehensive concept of earth observation has emerged, which focuses on programmatic and application-oriented contexts. Earth observation refers not only to the

collection of data, but also to its systematic reuse for analysis, monitoring, and decision support, thus underscoring its role in linking technological capabilities with a wide range of applications [19]. This distinction has become increasingly important in the historical development of the field, where early sensor-oriented approaches have evolved into comprehensive observation programs such as Landsat and Copernicus.

The historical development of earth observation is closely linked to the launch of various satellite programs, with Sputnik 1 in 1957 marking the beginning of the satellite age and heralding the era of space-based remote sensing [20]. A milestone was the introduction of the Landsat program in 1972, which has since provided the longest continuous record of space-based observations of the Earth's land surface [21]. Over the past five decades, Landsat has become a cornerstone of global environmental monitoring, providing important data for science, resource management, and policy decisions worldwide [22]. This has been further expanded since the launch of the first Copernicus Sentinel missions with Sentinel-1 in 2014 [23]. The observation capabilities of the various Sentinel constellations have been specifically developed for the systematic monitoring of land, ocean, and cryosphere variables and provide long-term, spatially explicit data sets that meet a wide range of scientific and applied requirements [24]. Since the 2000s, the commercialization of earth observation has accelerated significantly, reflected in a rapid increase in privately operated satellite constellations [25]. Planet Labs is among the most prominent commercial providers, operating large satellite systems organized into constellations. These include RapidEye, which provides high-resolution optical imagery at 3–5 meters [26], and PlanetScope—combining the Flock, Dove, and SuperDove—to deliver very-high-resolution, near-daily multispectral coverage [26, 7]. In addition, Airbus Defence & Space has established itself as another important provider, operating long-standing systems such as SPOT [27] and, more recently, the Pléiades constellation, which has been delivering very high-resolution optical images since 2011 [28].

Parallel to the growth in earth observation constellations, new cloud-based infrastructures have emerged to ensure that this big amounts of data can be accessed and analyzed efficiently. The Copernicus Data Space Ecosys-

tem (CDSE), the official platform of the European Union, offers comprehensive access to a wide range of earth observation data, with a particular focus on the Sentinel constellations [29]. In addition, systems such as Microsoft Planetary Computer (MPC) [30] and Google Earth Engine (GEE) [31] provide large-scale access to earth observation archives such as Sentinel, Landsat, and MODIS, as well as curated environmental datasets. These platforms are designed for scalability and enable the integration of advanced analytical methods, supporting the processing and analysis of large amounts of data.

## 1.2 Deep Learning

Deep learning has emerged in recent years as a key technology in remote sensing, playing a crucial role in the automated analysis of ever-growing data volumes [10]. Its strength lies particularly in image processing and pattern recognition, enabling the systematic extraction of information from large-scale datasets. Foundational work dates back several decades, such as the application of backpropagation to handwritten digit recognition [32]. A major breakthrough for image recognition was achieved with the introduction of convolutional neural networks (CNNs), which demonstrated their power in large-scale image classification tasks [33]. Today, the processing and interpretation of image data is generally subsumed in the field of computer vision under the definition of image recognition. Within this field, deep learning methods are applied to different but related tasks. These include image classification, in which entire images or image sections are assigned to predefined categories; object detection, which combines classification with regression to identify and localize objects within an image; and semantic segmentation, which assigns a class label to each pixel to produce detailed spatial representations [34].

Deep learning methods in image processing are commonly based on a number of well-established methodologies. CNNs form an essential basis, with architectures such as Visual Geometry Group (VGG) [35] and Residual Network (ResNet) improving depth and representation performance [36]. For object detection, region-based approaches such as R-CNN [37], Fast R-

CNN [38] and Faster R-CNN [39] introduced the combination of classification and regression for bounding box prediction, while architectures such as RetinaNet additionally addressed the problem of class imbalances through focal loss [40]. Semantic segmentation tasks rely on encoder-decoder networks such as U-Net, which was originally developed for biomedical imaging [41], as well as advanced designs such as DeepLab, which introduced atrous convolutions for multiscale context [42]. A further step is Mask R-CNN, which extends object detection frameworks to simultaneously generate instance-level segmentation [43].

The introduction of transformer-based models also represents progress in this area, with the groundbreaking publication “Attention is All You Need”, which laid the foundation for Vision Transformers (ViTs) by demonstrating the power of self-attention mechanisms in capturing long-range dependencies [44]. Foundation models have also gained significantly in importance in this area. They are characterized by large-scale pre-training with diverse datasets, which allows for flexible adaptation to a variety of downstream tasks [45]. A particularly prominent example within the computer vision community is the Segment Anything Model (SAM), which provides zero-shot segmentation capabilities, generating masks for new categories without task-specific training, and has also been transferred to remote sensing applications [46].

When using deep learning models, the first step usually involves a training phase. Training data is used to supervise the learning process. This data is often created manually and is therefore time-consuming and costly, especially when detailed information at the pixel or object level is required. Various strategies have been developed to address these limitations. Semi-supervised approaches aim to reduce dependence on manually labeled data by combining a small amount of labeled data with larger unlabeled datasets or by generating labels from additional geodata sources such as existing inventories or maps [47]. In contrast, self-supervised learning uses the data itself to generate supervision signals without requiring external annotations. Typical pretext tasks include rotation prediction or contrastive learning, which enable models to learn general feature representations from large amounts of unannotated

images before being fine-tuned for specific downstream applications [48, 49].

Validation of a model can be performed using standardized metrics. Accuracy, defined as the proportion of correctly classified samples, provides a basic measure, but is often not meaningful in cases of class imbalance. Therefore, additional metrics such as precision, recall, and the F1-score are used as examples for classification tasks. For object detection, Average Precision (AP) is often used, which summarizes precision-recall curves. In the field of segmentation, Intersection over Union (IoU) is the standard metric for the pixel-wise overlap between predicted and actual regions [50, 51].

Key challenges arise when applying deep learning models. One of the most common problems is overfitting, where models adapt too strongly to the training data and therefore cannot be generalized well to new data. To reduce this effect, regularization techniques such as dropout [52], weight decay [53], or early stopping [54] are widely used. In addition, bias in data sets should be avoided, as imbalances in the training data are transferred to the model and can limit its applicability in different contexts [55].

Finally, it should be considered that deep learning approaches often have high computing power requirements, as training and deploying complex architectures can require powerful Graphics Processing Units (GPUs) or Tensor Processing Units (TPUs) and large storage capacities. This limitation is particularly evident in modern transformer-based approaches [44] and in large language models, where inference alone for models with hundreds of billions of parameters already demands massive energy and hardware resources [56].

### 1.3 Renewable Energy System Recognition

Combining the two fields of remote sensing and deep learning opens up numerous possibilities for the recognition of renewable energy systems. The following section examines the recognition of renewable energy systems in two main areas: firstly, the recognition of PV systems and secondly, the recognition of wind turbines.

The recognition of solar photovoltaic systems from remote sensing imagery has progressed significantly over the past decade. Early studies applied

classical computer vision techniques to identify PV systems in high-resolution aerial and satellite images, demonstrating the feasibility of automatic recognition in small test areas in the United States [57, 58]. A major step toward scalability was introduced with the DeepSolar framework, which constructed a large-scale satellite image dataset using the Google Static Map API and around 50,000 manually labeled training images combined with a deep learning framework to detect PV systems across large regions in the United States [59]. The framework was later adapted to new regional contexts, such as Germany, where retraining on locally specific data demonstrated the transferability of the approach [60]. The effort to expand from regional to global applications resulted in the first worldwide inventory of large-scale PV systems. By combining Sentinel-2 and SPOT-7 data with deep learning, more than 68,000 systems were identified and mapped, resulting in a globally consistent database of large-scale PV [61].

Until then, known approaches relied on manually annotated training data. While these datasets have enabled important advances, their creation can be extremely labor-intensive, costly, and difficult to reproduce at scale. This dependence on manual labeling poses a significant limitation for the broader application and transferability of deep learning-based PV detection methods.

Further methodological investigations addressed the challenge of transferring models across spatial resolutions. In addition to presenting a dataset with multiple resolutions that combines satellite and UAV images, it was shown that targeted fine-tuning can lead to robust transferability between different image data [62]. Subsequently, new architectures such as GenPV demonstrated that models trained on RGB images with different resolutions can achieve generalizable segmentation performance when optimized appropriately [63]. In addition to methodological contributions, comprehensive reviews have structured the field. An overview of more than 100 studies on applications in context highlighted the rapid growth of data-driven approaches and their diverse use cases [64]. Another review systematically evaluated available image sources, public datasets, and identification methods and concluded that segmentation models based on U-Net and DeepLab currently achieve the highest accuracy and that existing datasets can be

reused extensively for training purposes [65].

Another central challenge lies in the limited robustness of existing segmentation models when applied across different spatial resolutions and geographic contexts. Most approaches are optimized for a single data source or resolution, which often restricts their transferability. This lack of generalization hampers the development of scalable methods that could be reliably applied in diverse regional or global contexts.

In contrast to the numerous studies addressing the recognition of PV systems, the literature on the recognition of wind turbines is still comparatively limited. However, existing work has presented both optical and radar-based approaches that represent important methodological advances. A early study used high-resolution aerial and satellite images derived from Google Earth Engine platform and employed visual salience detection, which involved coarse identification followed by fine extraction. This method showed that wind turbines can be reliably extracted from high-resolution images with a spatial resolution of 1 m, underscoring the suitability of optical data for object-level detection tasks [66].

In addition, the use of high-resolution oblique UAV imagery combined with deep learning has proven to be a highly promising approach for the detection of wind turbines [67], demonstrating that a combination of feature descriptors and pre-trained deep neural networks can achieve high accuracy in detection. Using Sentinel-1 data and Google Earth, a global inventory of offshore wind turbines was first created through a systematic data processing pipeline. The extraction workflow relied on a structured sequence from the removal of temporary objects to the identification of high-backscatter features and the postprocessing of object records, ultimately mapping the global distribution of offshore systems [68]. The DeepOWT framework introduced the first global dataset of offshore wind turbines derived from deep learning techniques. Based on Sentinel-1 imagery, nearly 10,000 wind turbine locations were identified along with their stage of deployment, establishing a consistent large-scale inventory of offshore wind turbines [69]. In addition, the SyntEO approach introduced the first framework for the automated generation of synthetic training data for offshore wind farms using Sentinel-1

imagery for robust detection in complex environments [70]. At the national level, Sentinel-2 imagery were used to map onshore and offshore wind turbines with high precision in China, demonstrating the feasibility of large-scale turbine inventories based on medium-resolution optical imagery [71].

Recent advances aim to achieve consistent global monitoring over time. PlanetScope images analyzed at quarterly intervals between 2017 and 2024 were used to create a global dataset on solar and wind energy systems, identifying over 375,000 wind turbines and 86,000 PV systems with the aim of providing a permanent using deep learning-based monitoring platform for time-resolved, inventory of renewable energy systems worldwide [72].

Beyond the detection of wind turbines in imagery, a related challenge lies in the completeness and reliability of the resulting datasets. While large-scale inventories can provide valuable coverage and consistent location data, they typically contain spatial information but do not provide crucial technical details such as turbine type, hub height, or rotor diameter—attributes that are essential for energy system research. At the same time, existing registers or community-sourced datasets can suffer from positional inaccuracies. Such shortcomings reduce their reliability for subsequent analyses and underline the need for methods that can systematically correct these datasets.

This section shows that, despite numerous advances in the field of remote sensing based renewable energy system recognition using deep learning, significant challenges remain. First, most deep learning-based studies rely on labor-intensive, manually annotated training data, which constrains scalability and reproducibility. Second, the robustness of segmentation models across different spatial resolutions and geographic contexts is still insufficiently understood, although such transferability is crucial for global applications. Third, existing infrastructure datasets often contain positional inaccuracies and incomplete attribute information, raising the question of how deep learning can be integrated into workflows that correct and enrich these datasets for reliable use. In this context, this dissertation addresses three overarching research questions:

1. How can training data for deep learning models be generated efficiently and reproducibly?
2. How robust are segmentation approaches across different spatial resolutions and geographic contexts?
3. How can deep learning methods improve the positional accuracy of existing renewable energy system datasets?

These questions are derived from the identified gaps in the current research landscape and form the foundation for the hypotheses outlined in the following section.

## 1.4 Central Research Hypotheses

Based on the identified research gaps, the following hypotheses were formulated. These hypotheses are well suited for straightforward testing within the dissertation, enabling systematic processing, verification, and quantitative evaluation.

**Hypothesis 1:** *Semi-automated methods for generating training samples can provide reliable inputs for deep learning-based recognition of renewable energy systems and offer a possible way to reduce dependence on time-consuming manual labeling.*

Chapter 2 investigates this by combining publicly available address-based information from official register data with national building geometries to derive building footprints. These footprints are used to generate large-scale training datasets for rooftop photovoltaic systems, which then serve as input for deep learning-based object detection.

**Hypothesis 2:** *Deep learning-based segmentation methods can enable precise and robust identification of solar photovoltaic systems across multiple spatial resolutions and can be transferred to new geographic contexts without region-specific retraining.*

Chapter 3 and 4 address this by evaluating multi-resolution segmentation models in different geographic environments. The analyses cover satellite, aerial, and UAV imagery at spatial resolutions from 0.1, 0.2, 0.3, 0.8, 1.6, to 3.2 meters, focusing on cross-scale robustness. Integration into open-source environments such as QGIS is further presented as a way to make the methods accessible to non-experts and applicable in diverse planning contexts.

**Hypothesis 3:** *Object detection methods can systematically reduce positional errors in renewable energy system datasets, thereby enhancing their spatial accuracy and overall reliability for energy system research.*

This third hypothesis is mainly investigated in Chapter 5 and 6, which apply a RetinaNet-based object detection approach to wind turbine datasets in Germany and South Africa. The workflows combine high-resolution aerial and satellite imagery with register and crowdsourced data to correct turbine coordinates and enrich attribute information, resulting in improved datasets for subsequent use in energy system research and planning.

## 1.5 Core Structure of the Dissertation

The three hypotheses outlined above provide the conceptual framework of this dissertation. Each of the five core Chapters (2–6) is based on a scientific publication that addresses one or more aspects of these hypotheses. Chapters 2, 3, and 6 are derived from peer-reviewed journal articles, while Chapters 4 and 5 build on conference papers that underwent only a more limited review process. Together, these contributions follow a logical sequence from data preparation, to segmentation and scaling, to dataset correction and enrichment. In this way, the structure of the dissertation reflects both a methodological progression and the thematic breadth across remote sensing, deep learning, and renewable energy system recognition.

- Chapter 2, *"Semi-Automatic Generation of Training Samples for Detecting Renewable Energy Plants in High-Resolution Aerial Images"*, investigates a semi-automated approach for generating training samples

for rooftop PV detection in aerial imagery. Unlike earlier approaches that relied exclusively on manual annotation, it combines registry-based information with building geometries to generate reproducible datasets, thereby contributing to Hypothesis 1.

- Chapter 3, "*Multi-Resolution Segmentation of Solar Photovoltaic Systems Using Deep Learning*", evaluates the application of a popular deep learning architecture across different image sources and experimental settings, thereby assessing its robustness under varying conditions. While previous segmentation models were typically restricted to a single resolution or data source, this study systematically combines and compares multiple resolutions across geographic contexts, thereby supporting Hypothesis 2.
- Chapter 4, "*Globally Scalable, QGIS-Integrated Workflow for Solar Photovoltaic System Segmentation and Capacity Estimation: A Case Study in Algeria*", applies the segmentation-based workflow from the previous study to large-scale PV systems in Algeria's Piat region. This study demonstrates how the workflow can be scaled up and integrated into open-source environments for international use without region-specific retraining, further supporting Hypothesis 2.
- Chapter 5, "*Enhancing Wind Turbine Location Accuracy: A Deep Learning-Based Object Regression Approach for Validating Wind Turbine Geo-Coordinates*", presents an object detection-based approach for correcting turbine coordinates in Germany. For the first time, existing turbine registers with often inaccurate location data can thus be systematically improved using RetinaNet and high-resolution aerial imagery, directly contributing to Hypothesis 3.
- Chapter 6, "*A Wind Turbines Dataset for South Africa: OpenStreetMap Data, Deep Learning Based Geo-Coordinate Correction and Capacity Analysis*", transfers the workflow from the previous chapter to South Africa. This study provides the first openly available, corrected, and

enriched dataset of wind turbines for the country, thereby extending the investigation of Hypothesis 3.

Taken together, these studies illustrate how image-based methods can advance renewable energy system recognition, providing methodological contributions and practical workflows that form a transferable basis for future applications in remote sensing and deep learning.

## 1.6 Research Context

The research presented in this dissertation was made possible through the synergy of several research projects. The following sections provide a brief overview of the most relevant projects that supported and shaped the individual studies included in this thesis.

### OASES

Several parts of the dissertation were carried out as part of the OASES project “Development and demonstration of a sustainable open access ecosystem for modelling energy systems between Africa and Europe” at the University of Kassel, Department of Electrical Engineering and Computer Science. OASES is funded under the “LEAP-RE, Europe-Africa Research and Innovation Call on Renewable Energy” program of the European Union’s “Horizon 2020” research and innovation program (grant agreement 963530) and the German Federal Ministry of Education and Research (03SF067). The aim of the project is to develop and demonstrate a sustainable AU-EU ecosystem for energy system modelling based on open-source software and freely accessible data. A key focus is the creation of input data for energy system models, including the spatial distribution of renewable energy systems, the evaluation of potential areas and resources, and power generation time series for solar and wind energy generation. The University of Kassel investigates the recognition of renewable energy systems from satellite data. The aim is to create an open, georeferenced dataset of installed renewable energy systems in selected regions and make it freely available on suitable platforms. The

methodology is based on deep learning techniques and open-source earth observation data for the automated recognition of wind turbines and PV systems. Where possible, relevant technological parameters such as installed capacity were added and published as open source. To ensure sustainable usability, the source code was documented and published together with the datasets. Model accuracy was evaluated through validation with manually verified data. The project funded the Publication 3 “Multi-resolution segmentation of solar photovoltaic systems using deep learning.” In one use case, the developed methods were applied to large-scale PV systems in Algeria, resulting in a globally scalable, QGIS-integrated workflow that was published as 4 “Globally Scalable, QGIS-Integrated Workflow for Solar Photovoltaic System Segmentation and Capacity Estimation: A Case Study in Algeria.” Another use case extended the methodology to South Africa and combined OpenStreetMap data and satellite imagery to create a complete dataset of wind turbines. The results are presented in the Publication 6 “A Wind Turbine Dataset for South Africa: OpenStreetMap Data, Deep Learning-Based Geocoordinate Correction, and Capacity Analysis” and provide a freely accessible dataset for the research of energy systems in South Africa.

## **RobustPlan**

Robust Transmission Grid Planning - RobustPlan project focuses on determining resilient transformation paths for the German electricity transmission grid under profound uncertainties, such as changes in energy policy, scenario assumptions or technological developments. The aim is to ensure security of supply and cost efficiency until 2050 despite these uncertainties. This requires comprehensive scenarios, probabilistic approaches and the automation of grid planning workflows. The University of Kassel, in collaboration with Fraunhofer IEE, was responsible for the spatial mapping of renewable energy systems and conventional loads. This included the regionalization of existing and planned renewable energy systems, the use of geostatistical data on population and infrastructure and the creation of high-resolution time series based on historical weather data for wind and solar energy. An important

---

innovation in this context was the application of deep learning to identify PV systems in aerial images. This work led to the Publication 2, which was supported by the associated Cognitive Energy Systems project.

### **Cognitive Energy Systems**

The project Cognitive Energy Systems is essential for the funding of the Publication "Semi-Automatic Generation of Training Samples for Detecting Renewable Energy Plants in High-Resolution Aerial Images". Since its foundation in mid-2020, Cognitive Energy Systems has been dedicated to the research and development of cognitive processes. The aim is to research innovative solutions for energy system technology, the energy industry and the design and optimization of energy grids. The idea and motivation behind the project is to support the continuous development of a simulation environment for the future energy supply with spatial and temporal mapping of the expansion dynamics of renewable energy systems using remote sensing data and machine learning methods. Existing data bases are usually maintained manually. By using DL methods, the method developed is intended to provide a way of spatially determining very high-resolution location data of systems.

## Chapter 2

# Semi-Automatic Generation of Training Samples for Detecting Renewable Energy Plants in High-Resolution Aerial Images

## Bibliographic Information

Kleebauer, M., Horst, D., Reudenbach, C. (2021). Semi-automatic generation of training samples for detecting renewable energy plants in high-resolution aerial images. *Remote Sensing*, 13(23), 4793.

## Author's contribution

In the following paragraph, the individual contributions of the authors are briefly broken down with regard to the publication. The authors are abbreviated as follows: Maximilian Kleebauer (M.K.), Daniel Horst (D.H.) and Christoph Reudenbach (C.R.). Conceptualization, M.K., D.H., and C.R.; methodology, M.K. and D.H.; software, M.K. and D.H. ; validation, M.K.; formal analysis, M.K.; investigation, M.K., D.H., and C.R.; resources, D.H.; data curation, M.K.; writing - original draft preparation, M.K.; writing—review and editing, M.K., D.H., and C.R.; visualization, M.K.; supervision, M.K.; project administration, D.H.; funding acquisition, D.H. All authors have read and agreed to the published version of the manuscript.

## Copyright Notice

© 2021 by the authors. Published by MDPI. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) license (<https://creativecommons.org/licenses/by/4.0/>).

Original article: Semi-automatic generation of training samples for detecting renewable energy plants in high-resolution aerial images. *Remote Sensing*, 2021, 13(23), 4793.

<https://doi.org/10.3390/rs13234793>

## 2.1 Abstract

Deep learning (DL)—in particular convolutional neural networks (CNN)—methods are widely spread in object detection and recognition of remote sensing images. In the domain of DL, there is a need for large numbers of training samples. These samples are mostly generated based on manual identification. Identifying and labelling these objects is very time-consuming. The developed approach proposes a partially automated procedure for the sample creation and avoids manual labelling of rooftop photovoltaic (PV) systems. By combining address data of existing rooftop PV systems from the German Plant Register, the Georeferenced Address Data and the Official House Surroundings Germany, a partially automated generation of training samples is achieved. Using a selection of 100,000 automatically generated samples, a network using a RetinaNet-based architecture combining ResNet101, a feature pyramid network, a classification and a regression network is trained, applied on a large area and post-filtered by intersection with additional automatically identified locations of existing rooftop PV systems. Based on a proof-of-concept application, a second network is trained with the filtered selection of approximately 51,000 training samples. In two independent test applications using high-resolution aerial images of Saarland in Germany, buildings with PV systems are detected with a precision of at least 92.77 and a recall of 84.47.

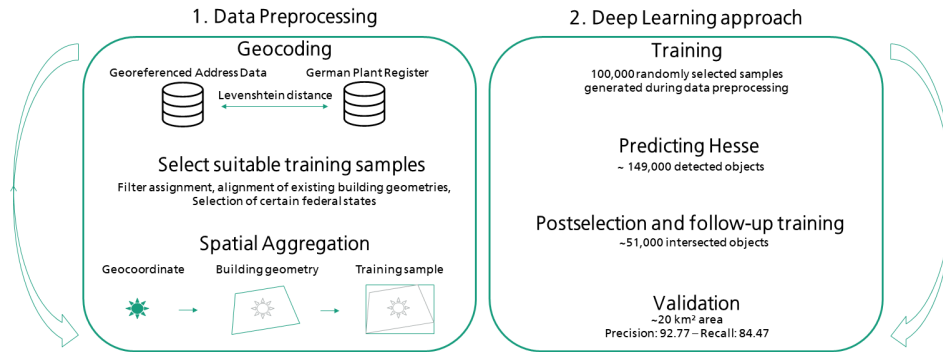
**Keywords** deep learning; remote sensing; object detection; photovoltaic

## 2.2 Introduction

Machine learning (ML) methods have already proven to be helpful for localizing objects in the past. This offers the possibility to recognize objects in aerial images and thus to identify the locations of renewable power generators such as PV roof-mounted systems for energy system analysis. Several papers on the detection of PV systems have already been published over the last years [1, 2, 3, 6, 5, 6, 7].

In general, all authors used supervised learning methods, where datasets are created initially by labelling a huge number of objects of the given class. The number of training samples used for the detection of PV systems increased steadily with the development of the methodologies. For example, [3] uses over 2700 images containing PV systems, [5] more than 50,000 and [6] up to 70,673 labelled images. The creation of such large amounts of training data is very time- and cost-intensive [8, 9]. To avoid the tedious preparation of data, only a few approaches have been presented. An example of the automated creation of training data using different data sources such as OpenStreetMap (OSM) and automatic information indexes referring to buildings, shadow, vegetation, and water has already shown good results in the task of image classification using maximum likelihood classification (MLC), multilayer perceptron (MLP), support vector machine (SVM), and random forest (RF) classifier. Not included in that study are approaches using modern CNN networks [10]. During the research of the present work, no approaches of an application of CNN classifiers for object detection using automatically generated training data could be found. Therefore, the present work ties in with the idea of automatic generation of training samples. Since deep learning networks such as CNN networks in particular are capable of producing reliable results on a growing base of training data, this paper relies on them. Due to the enormous costs and the unpredictable systematic bias of expert training, the semi-automatic acquisition of labelled training samples seems to be a suitable basis for efficient CNN classification.

In this study, an operational semi-automatic workflow as shown in Figure 2.1 is developed by using an automated matching of location data from the German PV Plant Register and the Georeferenced Address Data first generate geocoordinates of existing plants. After selecting suitable training samples, the coordinates are aggregated by area using the Official House Surroundings Germany dataset. Subsequently, these marked areas are used as training samples for a CNN-based object detector. Through a pre-application to the territory of the German state of Hesse and an automated matching with geocoded plants, the objects are postfiltered and used for a follow-up



**Figure 2.1:** The figure shows the workflow schematically.

training. Finally, the object detector is validated in a proof-of-concept study.

## 2.3 Materials

### 2.3.1 Georeferenced Address Data

The Georeferenced Address Data (GA) dataset provided by the Federal Agency for Cartography and Geodesy (BKG) contains completed and cleaned addresses, georeferencing coordinates, and a quality key for georeferenced building addresses at a uniform level of actuality for the federal territory of Germany. It is a summary of the Official House Coordinates of Germany of the land surveying administrations of the federal states and is supplemented at the BKG by addresses from the address dataset of Deutsche Post Direkt GmbH. The GA comprises about 22.6 million data records, proportionately about 730,000 data records from Deutsche Post [11].

### 2.3.2 Official House Surroundings Germany

The dataset published by the BKG summarizes the Official House Surroundings Germany (HU-DE) of the Central Office House Coordinates and House Surroundings of the surveying and cadastral administrations of the federal states and contains georeferenced house surround polygons of building ground plans of the Automated Real Estate Map [12].

### 2.3.3 German Plant Register

The German Plant Register was maintained by the Federal Network Agency for the German electricity and gas markets until 2019. The register contains information on players and installations in the grid-based energy supply market. In addition to registered power generation units, addresses, and power values for PV systems installed in Germany are listed, among other variables. In this work, these form the basis for determining the locations of existing PV systems.

### 2.3.4 Digital Orthophotos

The product of the Digital Orthophotos (DOP) of Germany [13] consists of georeferenced, differentially rectified aerial images of the surveying administrations of the federal states. They are true-to-scale raster data of photographic images of the earth's surface, limited to the territory of the Federal Republic of Germany. In the present work the DOP with a ground resolution of 20 cm are used. They are available as tiles with a resolution of  $5000 \times 5000$  pixels and a positional accuracy of  $\pm 0.4$  m standard deviation. Thus, an area of  $1000 \times 1000$  m is displayed per image. The product includes colour images (RGB) as well as infrared images and colour infrared (CIR) images. The present work uses the false colour composite CIR images, combining the infrared channel with the two visible colour or channels red and green.

## 2.4 Methodology

### 2.4.1 Data Preprocessing

To generate training samples, the address data of existing PV rooftop systems from the Plant Register were first geocoded by mapping these address data to geocoordinates from the GA. A method called Levenshtein distance was used for this purpose [14]. In contrast to exact mapping algorithms, this technique allows an approximation to two strings [15]. The minimum distance  $D_{i,j}$  was calculated in a  $m \times n$  matrix where each original address

string from the Plant Register  $X = x_1, x_2, \dots, x_m$  was adapted sign-by-sign to match the target address string from the GA addresses  $Y = y_1, y_2, \dots, y_n$ . Three different possible operations, such as insert  $D_{i,j-1}$ , delete  $D_{i-1,j}$ , and substitute  $D_{i-1,j-1}$  can be used to convert each cell  $(i, j)$ , which represents the distance between the original substring  $X_i^1 = x_1, x_2, \dots, x_i$  and the target substring  $Y_j^1 = y_1, y_2, \dots, y_j$  [16].

$$D_{i,j} = \min \begin{cases} D_{i-1,j-1} & + 0 \quad \text{if } X_{i-1} = Y_{j-1} \\ D_{i,j-1} & + 1 \quad (\text{insert}) \\ D_{i-1,j} & + 4 \quad (\text{delete}) \\ D_{i-1,j-1} & + 10 \quad (\text{substitute}) \end{cases} \quad (2.1)$$

The minimum number of operations was added up according to their cost (insert +1, delete +4, replace +10), so that the result was a weighted sum of operations that have to be performed to find the best match between the respective strings. The weighted sum could be used as a quality measure for the match and was exploited and filtered so that only geocoded PV addresses with a difference of 0, i.e., ideal matches, were used for further labelling. The assigned geocoordinates of the existing PV rooftop systems could then be provided with polygons of the HU-DE using a spatial intersection. In addition, the plants were filtered by a quality key of HU-DE, which ensured that the coordinates used lay safely within the recorded building geometries. Since the CNN architecture provides rectangular bounding boxes, the polygons were abstracted to rectangles in a further step. In order not to exceed the maximum tile size of the selected Backbone ResNet101 (768–1024 pixels), the original DOP tiles were divided into subareas of  $1000 \times 1000$  pixels, each  $200 \times 200$  m. It should be emphasized that the buildings with PV systems used as training samples always represented only a very small proportion of the image sections. Image sections without assignment to a class were interpreted as background class.

## 2.4.2 Deep Learning Approach

### CNN Architecture

The CNN network was based on the architecture called RetinaNet [17], combining the deep residual neural network ResNet101 [18], a feature pyramid network (FPN) [19] following previous object detectors such as “faster R-CNN” [20], and two task-specific classification and regression subnetworks. RetinaNet was used because, when compared by the COCO benchmark, it outperformed all previous one- and two-stage detectors, including the winners of the COCO 2016 competition, in terms of prediction accuracy relative to speed. The classification subnetwork performs object classification at the output of the backbone network based on focal loss ( $FL$ ). The focal loss is designed to train extremely unevenly distributed foreground and background classes. Based on the cross entropy loss for binary classification it adds a weighting factor  $\alpha \in [0, 1]$  for class 1 and  $1 - \alpha$  for class  $-1$  as well as a modulation factor  $(1 - p_t)^\gamma$  containing a tunable focusing parameter  $\gamma \geq 0$ , as shown in Equation (2) [17].

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t) \quad (2.2)$$

The parameters were set to  $\alpha = 0.25$ ,  $\gamma = 2.0$ ,  $p = 0.5$ , as ablation experiments achieved good results with this parameter combination. The regression subnetwork was implemented for regressive delineation of objects. The regression loss ( $L_{loc}$ ) is based on the smooth L1 loss ( $smooth_{L_1}$ ) approach, originally designed as part of the Fast R-CNN network [20]. The regression loss was used to target the bounding box regression and was defined over a tuple for the ground-truth class  $u$  and  $v$  for the bounding box regression target  $u, v = (v_x, v_y, v_w, v_h)$  and a predicted tuple for the class  $u$  with  $t^u = (t_x^u, t_y^u, t_w^u, t_h^u)$  [20], adapted [17] and implemented in [21] as shown in Equations (3) and (4).

$$L_{loc}(t^u, v) = \sum_{i \in \{x, y, w, h\}} smooth_{L_1}(t_i^u - v_i) \quad (2.3)$$

in which

$$\text{smooth}_{L_1}(x) = \begin{cases} 0.5 * (\sigma * x)^2 & \text{if } |x| < 1/\sigma^2 \\ |x| - 0.5/\sigma^2 & \text{otherwise} \end{cases} \quad (2.4)$$

The parameter  $\sigma$  was set to 3.0. The regression targets were output as rectangles that were entirely within the images shown and could take given aspect ratios of 1:2, 1:1, and 2:1. The result in predictions can be at multiple levels of the network, since there are multiple output layers. Based on good performance from previous object detection tasks, ResNet101 with an image size of 768–1024 pixels was chosen as the backbone network [18, 22, 23]. Since it has been repeatedly shown that pretrained networks achieve good generalization more quickly, an already implemented, freely available ResNet101 was used, which had already been pretrained with 500 classes from the Open Image Dataset [21, 24]. The learning rate started with 0.00001 and was reduced by factor 0.1 after 2 epochs during training without improvement, as measured by the values of total loss, which was calculated by summing up the  $FL$  and the  $L_{loc}$ .

## Training

The first training used 100,000 randomly selected, labelled images from the German states of Berlin, North Rhine-Westphalia and Thuringia as they were all generated automatically during the data preprocessing. The batch size was set to 100 and a termination criterion was selected so that the training was terminated if no progress was achieved as measured by the area under the precision–recall curve called average precision (AP) after 5 epochs. The validation dataset was prepared manually and contained images with 280 located PV systems. For this purpose, the selection of images was screened and the existing bounding boxes were supplemented or adjusted so that all validation images were fully equipped with bounding boxes of existing PV systems.

The Intersection over union (IoU) was first calculated by dividing the area overlap of detected objects and labelled objects by the total area of the two boxes. By classifying the results based on the threshold value of 0.5, we de-

terminated whether the detection matched the labelled object. If the IoU was higher than 0.5, the detection was considered as a true positive (TP). If it was lower, the detection was considered a false positive (FP). False negatives (FNs) referred to existing objects that were not recognized during prediction. True negatives (TNs) could also be considered a background class as they represented the image sections that did not belong to any class. As a result of the number of TP as well as FP and FN examples, the precision, recall, and the AP were calculated and are shown in Table 2.1.

**Table 2.1:** The table summarizes the formulas of the evaluation. In addition to the precision and the recall, the average precision is presented.

|                   | Formula                   |
|-------------------|---------------------------|
| precision         | $prec = \frac{TP}{TP+FP}$ |
| recall            | $rec = \frac{TP}{TP+FN}$  |
| average precision | $AP = \int_0^1 Pr(Re)dRe$ |

### Predicting Hesse

The trained model was applied to the area of the state of Hesse to ensure that none of the training data was reused for this initial test application. All detected plants with a threshold classification score of more than 0.3 were prepared as training examples for a follow-up training session. Doubly detected PV systems were reduced in a further step while the number of overlapping regression boxes was reduced based on the values of the score, so that only the boxes with the highest score remained.

### Post Selection and Follow-Up Training

In a further step, the geocoded addresses with a PV system and objects resulting from the detection were compared. Since the automated generation of the training samples also generated erroneous bounding boxes, all objects in Hesse were marked where the centre of the detected objects lay within the building geometry of the geocoded addresses with a PV system. Furthermore, the existing bounding boxes were adjusted by adopting the boxes resulting from the detection as new bounding boxes. These marked objects

were subsequently used for a follow-up training with the same architecture and parameters.

### Validation

To assess the accuracy of the model, an independent test application was carried out using a selection of DOP from the federal state of Saarland in Germany. Here, 385 randomly drawn images used in a first test represented the overall coverage as far as possible. In addition, an artificial enrichment of 121 images was used in a second test leading to a disproportionate occurrence of PV systems. The following Table 2.2 shows the numbers of images and the area covered by the images in km<sup>2</sup>. Furthermore, the number of buildings recorded in the HU-DE within the test images are noted.

**Table 2.2:** The table shows the number of images, the area covered by the images, and the number of buildings used for validation.

|                                 | Random Images | Images with PV | Overall |
|---------------------------------|---------------|----------------|---------|
| Images                          | 385           | 121            | 506     |
| Covered Area in km <sup>2</sup> | 15.4          | 4.84           | 20.24   |
| Buildings                       | 2870          | 6475           | 9345    |

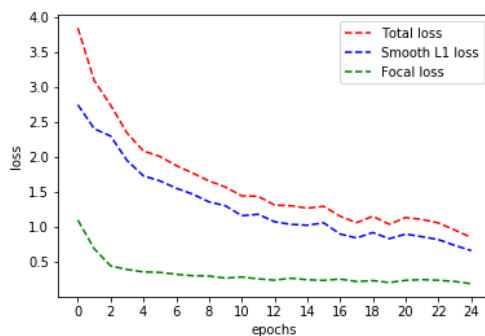
According to the HU-DE, there were a total of 9345 buildings in the test application on a total area of 20.24 km<sup>2</sup>. Buildings without a PV system were subsequently summarized as background class TN. To get an overview of the output of the trained network, all images including the predictions of the test were displayed. A manual check of all outputs was then carried out to examine the extent to which repeated patterns occurred in the detection of the objects. The precision and recall were then calculated.

## 2.5 Results

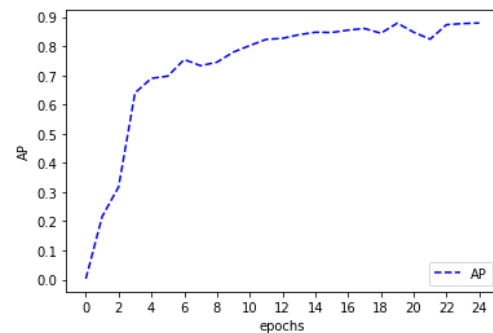
This section presents the results of the proof-of-concept. First, the results of the training with automatically generated samples from the federal states of Berlin, North Rhine-Westphalia, and Thuringia are shown. The subsequent filtering from the prediction and geocoded systems of Hesse as well as the

follow-up training are presented in the second part. Finally, the results of the independent validation in the federal state of Saarland are presented.

During the first training, the regression loss reaches a minimum of 0.19. The classification loss drops to a minimum of 0.66 (Figure 2.2). The calculated AP is 87.95% after the 24 epoch Figure 2.3.



**Figure 2.2:** Loss values during first training



**Figure 2.3:** AP values during first training

A total of 148,898 objects were detected during the prediction of the entire area of Hesse. The comparison of the overlap of the detected objects and the automatically generated geometries of the preprocessing resulted in 50,875 PV system locations. The existing bounding boxes (Figure 2.4, red) were adjusted by adopting the boxes resulting from the detection (Figure 2.4, blue) as new bounding boxes. These marked objects were subsequently used for a follow-up training with the same architecture and parameters.

During training with the postfiltered training samples, the classification loss reached the minimum value of 0.17 after the 16th epoch, with a regression loss of 0.49 and a total loss of 0.66 (Figure 2.5). The calculated AP was 91.76 after training in epoch 16 (Figure 2.6). Training was stopped because no progress was made after five epochs, as measured by the AP.

The different test applications and the resulting metric are summarized in Table 2.3.

From the validation with randomly selected images, 72 correctly detected plants resulted, 8 objects were incorrectly detected as PV systems, and 10 PV systems could not be detected. A total of 2780 buildings without PV

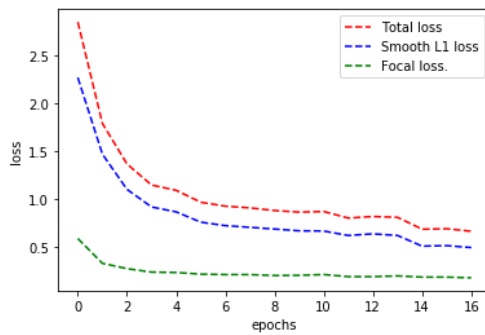


**Figure 2.4:** One red bounding box and two blue regression boxes are shown. Since the centers of the regression boxes lie within the building geometry of the geocoded facilities (drawn in black), they are subsequently used for the new training. Image/building geometry:© GeoBasis-DE/BKG (2019, 2020)

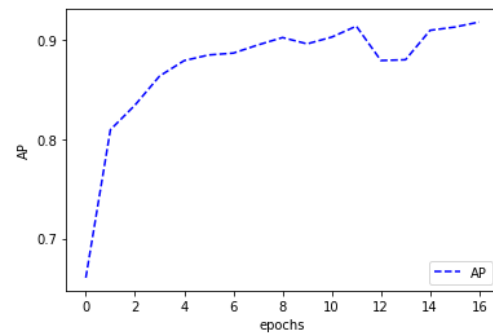
systems were correctly assigned to the background class. This resulted in a precision of 90.00 and a recall of 87.80. In the dataset additionally enriched with PV systems, 249 correctly detected PV systems, 17 objects incorrectly detected as PV systems, and 49 nondetected PV systems result. A total of 6160 buildings were included in the background class. The precision achieved was 93.60 and the recall was 83.56. The overall metrics summing up both test applications resulted in a precision of 92.77 and a recall of 84.47. Figure 2.7 shows a selection of correctly detected objects.

**Table 2.3:** The table represents the TP, FP, FN, and TN as well as the precision and the recall for both test datasets.

|               | Random Images | Images with PV | Overall |
|---------------|---------------|----------------|---------|
| TP            | 72            | 249            | 321     |
| FP            | 8             | 17             | 25      |
| FN            | 10            | 49             | 59      |
| TN            | 2780          | 6160           | 8940    |
| Precision (%) | 90.00         | 93.61          | 92.77   |
| Recall (%)    | 87.80         | 83.56          | 84.47   |



**Figure 2.5:** Loss values during second training



**Figure 2.6:** AP values during second training

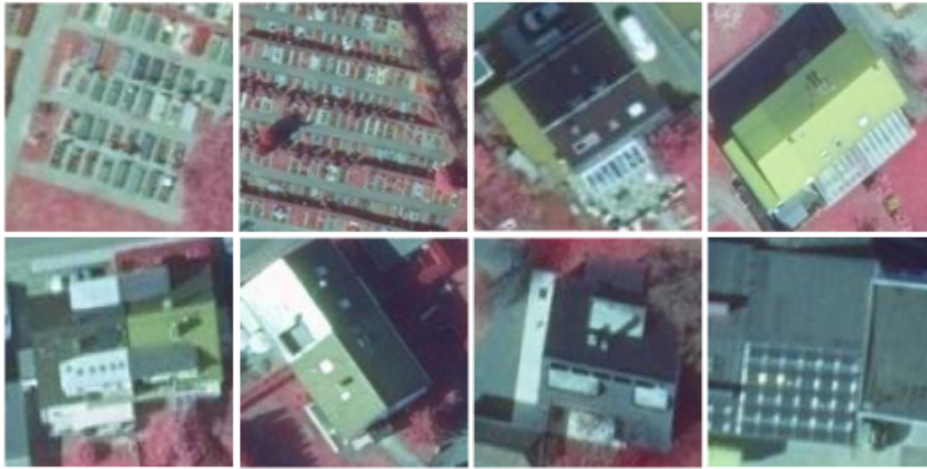


**Figure 2.7:** True positive examples. Images: © GeoBasis-DE / BKG (2019)

Figure 2.8 shows a selection of false positives. The objects in the first two pictures are not buildings. Furthermore, wrongly classified objects include canopies, patio roofs, and conservatories, as shown in the third and fourth picture. The second row shows objects where it is not possible to see whether they are PV systems or not.

## 2.6 Discussion

Using the two cases of the test application, the accuracy of the model was tested by selecting randomly drawn images, and in a dataset enriched with PV systems. It can be clearly observed that the enrichment of TP exam-



**Figure 2.8:** False positive examples. Images: © GeoBasis-DE / BKG (2019)

ples leads to a higher precision, whereby the recall decreases. Looking at the FP examples, repetitive patterns become visible. Particularly canopies, patio roofs, and conservatories are detected. The examples classified as FPs, which are not buildings, could be filtered in a real application based on the overlaps of the HU-DE, which could increase the quality of the application.

The FP examples that cannot be clearly assigned can be used as good examples of the challenge of remote-supported procedures in dealing with the creation of training areas. In some cases, the mere evaluation of the available image data is not sufficient for correct classification. The large number of TNs compared to the TPs shows the enormously unevenly distributed class ratio, as the number of TNs represents all buildings without PV systems contained in the test images. Since the TN value is not included in the calculation of precision and recall, the number of TN does not affect the metric calculation.

The automated generation of large amounts of training data can, of course, also lead to some of the training samples being incorrect. A comparison of the first and second training shows that despite halving the number of samples from the first to the second training, better results were achieved. Thus, all loss values decreased further, the AP increased from 87.95 to 91.76. This effect is probably due to the reduction of incorrectly generated samples,

which results in faster and better generalization during training.

Based on the proof-of-concept application, it can be shown that the developed workflow provides good results, with the semi-automatic processing of the training samples having the advantage of saving time by avoiding the time-consuming manual preparation of a large number of samples.

## 2.7 Conclusions

We developed a method for automated detection of buildings with PV systems. The generation of training data without manual selection allows the use of a large amount of training data. However, these may be inaccurate or erroneous. Comparisons of different tests suggested that postfiltering of the training data reduced the number of erroneous locations in a fully automated way. Based on the test application, the precision of the object detector was up to 92.77 and the recall was 84.47.

In future work, we will extend the procedure of semi-automated generation of object classes. By transferring the methodology to other classes of existing renewable energy plants such as wind turbines and biogas plants, the application can be expanded.

## Bibliography

- [1] Jiangye Yuan, Hsiu-Han Lexie Yang, Olufemi A Omitaomu, and Budhendra L Bhaduri. Large-scale solar panel mapping from aerial images using deep convolutional networks. In *2016 IEEE International Conference on Big Data (Big Data)*, pages 2703–2708. IEEE, 2016.
- [2] Jordan M. Malof, Rui Hou, Leslie M. Collins, Kyle Bradbury, and Richard Newell. Automatic solar photovoltaic panel detection in satellite imagery. In *2015 International Conference on Renewable Energy Research and Applications (ICRERA)*, pages 1428–1431. IEEE, 22.11.2015 - 25.11.2015.
- [3] Jordan M. Malof, Leslie M. Collins, Kyle Bradbury, and Richard G. Newell. A deep convolutional neural network and a random forest classifier for solar photovoltaic array detection in aerial imagery. In *2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA)*, pages 650–654, 2016.
- [4] Jordan M. Malof, Leslie M. Collins, and Kyle Bradbury. A deep convolutional neural network, with pre-training, for solar photovoltaic array detection in aerial imagery. In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 874–877, 2017.
- [5] Jiafan Yu, Zhecheng Wang, Arun Majumdar, and Ram Rajagopal. Deep-solar: A machine learning framework to efficiently construct a solar deployment database in the united states. *Joule*, 2(12):2605–2617, 2018.
- [6] Kevin Mayer, Zhecheng Wang, Marie-Louise Arlt, Dirk Neumann, and Ram Rajagopal. Deepsolar for germany: A deep learning framework for pv system mapping from aerial imagery. In *2020 International Conference on Smart Energy Systems and Technologies (SEST)*, pages 1–6. IEEE, 2020.

- 
- [7] Matthias Zech and Joseph Ranalli. Predicting pv areas in aerial images with deep learning. In *2020 47th IEEE Photovoltaic Specialists Conference (PVSC)*, pages 0767–0774. IEEE, 15.06.2020 - 21.08.2020.
- [8] Devis Tuia, Michele Volpi, Loris Copa, Mikhail Kanevski, and Jordi Munoz-Mari. A survey of active learning algorithms for supervised remote sensing image classification. *IEEE Journal of Selected Topics in Signal Processing*, 5(3):606–617, 2011.
- [9] Li Liu, Wanli Ouyang, Xiaogang Wang, Paul Fieguth, Jie Chen, Xinwang Liu, and Matti Pietikäinen. Deep learning for generic object detection: A survey. *International Journal of Computer Vision*, 128(2):261–318, 2020.
- [10] Xin Huang, Chunlei Weng, Qikai Lu, Tiantian Feng, and Liangpei Zhang. Automatic labelling and selection of training samples for high-resolution remote sensing image classification over urban areas. *Remote Sensing*, 7(12):16024–16044, 2015.
- [11] Bundesamt für Kartographie und Geodäsie. Georeferenzierte adressdaten - ga. 2020.
- [12] Bundesamt für Kartographie und Geodäsie. Dokumentation hausumringe deutschland: Hu-de. 2020.
- [13] Bundesamt für Kartographie und Geodäsie. Dokumentation digitale orthophotos. 2019.
- [14] VI Levenshtein. Binary Codes Capable of Correcting Deletions, Insertions and Reversals. *Soviet Physics Doklady*, 10:707, 1966.
- [15] Nimisha Singla and Deepak Garg. String matching algorithms and their applicability in various applications. *International journal of soft computing and engineering*, 1(6):218–222, 2012.
- [16] Aouragh Si Lhoussain, Gueddah Hicham, and YOUSFI Abdellah. Adapting the levenshtein distance to contextual spelling correction. *Inter-*

- national Journal of Computer Science and Applications*, 12(1):127–133, 2015.
- [17] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollar. Focal loss for dense object detection. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2999–3007. IEEE, 22.10.2017 - 29.10.2017.
- [18] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition.
- [19] Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 936–944. IEEE, 21.07.2017 - 26.07.2017.
- [20] Ross Girshick. Fast r-cnn. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 1440–1448. IEEE, 07.12.2015 - 13.12.2015.
- [21] Hans Gaiser, Maarten de Vries, Valeriu Lacatusu, vcarpani, Ashley Williamson, Enrico Liscio, András, Yann Henon, jjiun, Cristian Gratie, Mihai Morariu, Charles Ye, Martin Zlocha, Ben Weinstein, Rodrigo Meira de Andrade, Pedro Conceição, Alexander Pacha, hannedsvartsen, Daniyal Shahrokhian, Wudi Fang, Mike Clark, meagerYak, Iver Jordal, Max Van Sande, Jin, Etienne-Meunier, Andrew Grigorev, Guillaume Erhard, Eduardo Ramos, and Denis Dowling. *fizyr/keras-retinanet 0.5.1*. 2019.
- [22] Stefan Wolf, Lars Sommer, and Arne Schumann. Fastaer det: Fast aerial embedded real-time detection. *Remote Sensing*, 13(16):3088, 2021.
- [23] Zhifeng Xiao, Kai Wang, Qiao Wan, Xiaowei Tan, Chuan Xu, and Fanfan Xia. A2s-det: Efficiency anchor matching in aerial image oriented object detection. *Remote Sensing*, 13(1):73, 2021.

- [24] GitHub. Github - zfturbo/keras-retinanet-for-open-images-challenge-2018: Code for 15th place in kaggle google ai open images - object detection track, 2021.

## Chapter 3

# Multi-Resolution Segmentation of Solar Photovoltaic Systems Using Deep Learning

## Bibliographic Information

Kleebauer, M., Marz, C., Reudenbach, C., Braun, M. (2023). Multi-resolution segmentation of solar photovoltaic systems using deep learning. *Remote Sensing*, 15(24), 5687.

## Author's contribution

In the following paragraph, the individual contributions of the authors are briefly broken down with regard to the publication. The authors are abbreviated as follows: Maximilian Kleebauer (M.K.), Christopher Marz (C.M.), Christoph Reudenbach (C.R.) and Martin Braun (M.B.). Conceptualization, M.K.; methodology, M.K.; software, M.K. and C.M.; validation, M.K.; formal analysis, M.K., C.M., and M.B.; investigation, M.K., C.M., C.R., and M.B.; resources, M.K.; curation, M.K.; writing—original draft preparation, M.K., C.M., C.R., and M.B.; writing—review and editing, M.K. and C.M.; visualization, M.K. and C.M.; supervision, M.K.; project administration, M.K.; funding acquisition, M.K. and M.B. All authors have read and agreed to the published version of the manuscript.

## Copyright Notice

© 2023 by the authors. Published by MDPI. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) license (<https://creativecommons.org/licenses/by/4.0/>).

Original article: Multi-resolution segmentation of solar photovoltaic systems using deep learning. *Remote Sensing*, 2023, 15(24), 5687.

<https://doi.org/10.3390/rs15245687>

## 3.1 Abstract

In the realm of solar photovoltaic system image segmentation, existing deep learning networks focus almost exclusively on single image sources both in terms of sensors used and image resolution. This often prevents the wide deployment of such networks. Our research introduces a novel approach to train a network on a diverse range of image data, spanning UAV, aerial, and satellite imagery at both native and aggregated resolutions of 0.1 m, 0.2 m, 0.3 m, 0.8 m, 1.6 m, and 3.2 m. Using extensive hyperparameter tuning, we first determined the best possible parameter combinations for the network based on the DeepLabV3 ResNet101 architecture. We then trained a model using the wide range of different image sources. The final network offers several advantages. It outperforms networks trained with single image sources in multiple test applications as measured by the F1-Score (95.27%) and IoU (91.04%). The network is also able to work with a variety of target imagery due to the fact that a diverse range of image data was used to train it. The model is made freely available for further applications.

**Keywords** solar photovoltaic systems; photovoltaic plants; remote sensing; machine learning; deep learning; object detection; image segmentation

## 3.2 Introduction

The expansion of renewable energies is of key importance for the future of energy supply. Numerous reasons speak to the need for this transformation and highlight the need for a transition from fossil fuels to sustainable energy sources. For example, access to clean energy plays a central role in improving the quality of life and protecting our environment [1]. In order to provide a solid basis for the management of the development and integration of renewable energies on a small local, national, and continental scale, the modeling of energy systems can be seen as a practical tool [2, 3, 4]. A large set of reliable input data is required to run such models. The data required depend on the specific energy management problem. In general, however,

energy system models always require information on energy demand and energy production. If detailed information about the existing energy supply is not available for study areas, then machine learning (ML)-based methods of image recognition as well as earth-observation-based data offer a perspective for the generation of supplementary datasets.

The detection of existing renewable energy (RE) systems for mapping has been the goal of numerous studies [5]. While some early work on this topic deal with local spatial areas and individual PV systems [6], others generate datasets for entire countries [7]. Currently, from detailed local to global observations, various studies using remote sensing data such as satellite data [8, 9], aerial images from aircraft [10, 5], and drone images [12] exist to obtain detection at different scales in the respective study areas.

From published research in the field of PV system detection from images, it can be seen that there are numerous different ML methods used to perform the detection of PV systems. The meta-study “Advances and prospects on estimating solar photovoltaic (PV) installation capacity and potential based on satellite and aerial images” [13], for example, lists 17 different studies on the segmentation of PV systems in which different Fully Convolution Networks (FCN) and FCN-based methods are used. However, there are also numerous new studies with a strong focus on deep networks that are individually adapted to the task of segmenting PV systems. Zhu et al. [14] introduce a method to enhance PV segmentation from satellite imagery with a detail-oriented deep learning network using a Deeplabv3+ based Network. The network combines a Split-Attention Network and a Dual-Attention module with Atrous Spatial Pyramid Pooling (ASSP). Furthermore, the network has been proved to have favorable generalization capacity. Data from Google Earth Satellite with a resolution at 0.15 m in the area of Northern Baden-Württemberg in Germany was used to train the network. The authors also developed a new feature called the Constraint Refinement Module. The module enables the refinement tasks of localizing the PV panel region and thereby increasing the regulation of the PV panel shape [15]. Wang et al. [16] de-

veloped a size-aware deep-learning-based network for segmenting small-scale rooftop solar PV systems from high-resolution images. The size-aware network has performed well when it comes to the transfer of the application of the network to different datasets of similar pixel resolution.

While these studies achieve high accuracy, there is a notable caveat. The datasets used within this studies vary from 0.1 m to 30 m image resolution. Looking at the accuracy of the studies' result metrics, it can be observed that the detection accuracy is very high regardless of the image resolution chosen. However, this is only the case as long as the image data of the training match the data of the validation as well as in the following real-life application. If trained networks are applied to new image data with a different image resolution, for example, this can lead to very poor accuracy without further training [17]. This leads to the models only be restrictively used in practice, which is a considerable disadvantage. In addition, care should be taken when combining image data from different sources such as Microsoft Bing Maps and Google Earth Engine (GEE). This is particularly true when using very-high-resolution (VHR) aerial and satellite imagery. This is due to the fact that the maximum resolution of images varies from region to region due to vastly uneven global availability [18]. Furthermore, distortions become apparent in certain areas such as the USA, Europe, and India with significant interference at political borders [19]. Therefore, in order to address the different requirements for such a PV system segmentation model, some studies have already partially investigated the suitability of combining different resolution datasets for training different methods. For example, Jiang et al. [17] investigated the suitability of "cross application at different resolutions" where models were trained with data of a specific native resolution and a subsequent test application was applied to image data of a different resolution. The results show very poor transferability. Conversely, the transfer of models from one to another native resolution can be done with very good results if networks are fine-tuned with the target resolution. However, this means that the accurate application to new image data basically requires fine-tuning of existing models to new image data. Su et al. [20] propose a

cross-scale acquisition method that combines satellite data from the GaoFen constellation at 2 m resolution, Sentinel-2 at 10 m resolution, and Google satellite imagery aggregates at approximately 2 m and 10 m resolution. The results indicate that models can benefit from training with image data of different resolutions. Wang et al. [21] were recently able to show by enriching another external training data that enrichment with additional samples can lead to an increase in performance. Guo et al. [22] developed a deep learning model called GenPV. GenPV goes a step further by making use of unbalanced datasets using three-band red, green, blue (RGB) Google Earth tiles with a spatial resolution of 0.15, 0.3, and 0.6 m, respectively. The authors succeed in showing that a network trained with a single resolution can be transferred to another resolution with a high degree of performance [22].

Since, to the best of our knowledge, there is no research conducted that presents a model of PV segmentation that uses a variety of different image data with different resolutions and sensors for training at the same time, the effect of diverse image data on performance will be examined in more detail in this paper. Contrary to existing studies, this paper will not present an individualized deep-learning-based design. The focus is on successively enriching and combining a model with a large number of different existing training data by using the already extensively applied DeepLabV3 ResNet101 network. For this purpose, both native and aggregated images of the visible spectrum and resolutions in the range of 0.1 m, 0.2 m, 0.3 m, 0.8 m, 1.6 m, and 3.2 m as well as different sensor types of drone, aerial, and satellite images are used in this work. For comparison, the enriched networks are presented alongside models that were only trained on individual resolutions. Thus, this paper aims to describe the development of an image segmentation model capable of predicting PV systems from a variety of different images at different resolutions with a high level of performance.

### 3.3 Materials

In the domain of RS, a large number of different datasets have been published in recent years that can be used to train ML methods for detection of PV systems. These datasets can basically be distinguished into two different formats: vector and raster as training data in combination with aerial or satellite imagery. For example, Kruitwagen et al. [9] provide a dataset of large PV systems distributed across the globe. Dunnett et al. [23] even extend such a dataset with additional location data for wind generation facilities. However, image data are not included in these datasets. Conversely, the raster datasets are PV systems location data in the form of ready-made masks that are available along with image data collections. Therefore, the raster datasets can be used directly for training ML-based methods. All the datasets listed here have been manually inspected and tagged by humans. The first dataset mentioned here (Table ??, dataset DOP) consists of aerial photographs from the Federal Agency for Cartography and Geodesy (BKG) in Germany covering the state of North Rhine-Westphalia with a native resolution of 0.1 m [10]. Other datasets used consist of image data located within China that were obtained from unmanned aerial vehicles (UAVs) in Hai'an county (Table 3.1, dataset PV01), area imagery from the Provincial Geomatics Center (PGC) of Jiangsu (Table 3.1, dataset PV03), and satellite imagery from Gaofen-2 as well as Beijing-2 satellite imagery (Table 3.1, dataset PV08) at resolutions of 0.1 m, 0.3 m, and 0.8 m, respectively, with masks of existing PV systems [17]. The most recent published dataset [24] consist of combinations of imagery data. The combine dataset contains overflight imagery in France from the National Institute of Geographic and Forest Information (IGN) Geoservices portal with a native resolution of 0.2 m (Table 3.1, dataset IGN) as well as images gathered using GEE with 0.1 m (Table 3.1, dataset GEE) native resolution around different places in Germany. A special feature of the latter dataset is that it contains locations in the form of masks of existing PV systems as well as other information such as roof slope and orientation. An overview of the essential data such as native resolutions, number of images, and sensors used can be found in Table 3.1.

**Table 3.1:** Summary of the information on the various datasets.

| Dataset | Sensor | Category     | Res.  | Size      | with PV | Located       | Publisher |
|---------|--------|--------------|-------|-----------|---------|---------------|-----------|
| DOP     | Areal  | Roof, Ground | 0.1 m | 320×320   | 2117    | Germany       | [10]      |
| PV01    | UAV    | Roof         | 0.1 m | 256×256   | 645     | Hai'an/China  | [17]      |
| GEE     | agg.   | Roof         | 0.1 m | 400×400   | 13303   | France/Europe | [24]      |
| IGN     | Areal  | Roof         | 0.2 m | 400×400   | 7686    | France        | [24]      |
| PV03    | Areal  | Roof, Ground | 0.3 m | 1024×1024 | 2308    | Jiangsu China | [17]      |
| PV08    | Sat    | Roof, Ground | 0.8 m | 1024×1024 | 763     | China         | [17]      |

## 3.4 Methods

The following section first describes the image data preprocessing. Subsequently, the model used, the training with different image datasets, the hyperparameter tuning, and finally, the steps of the validation are described.

### 3.4.1 Data Preprocessing

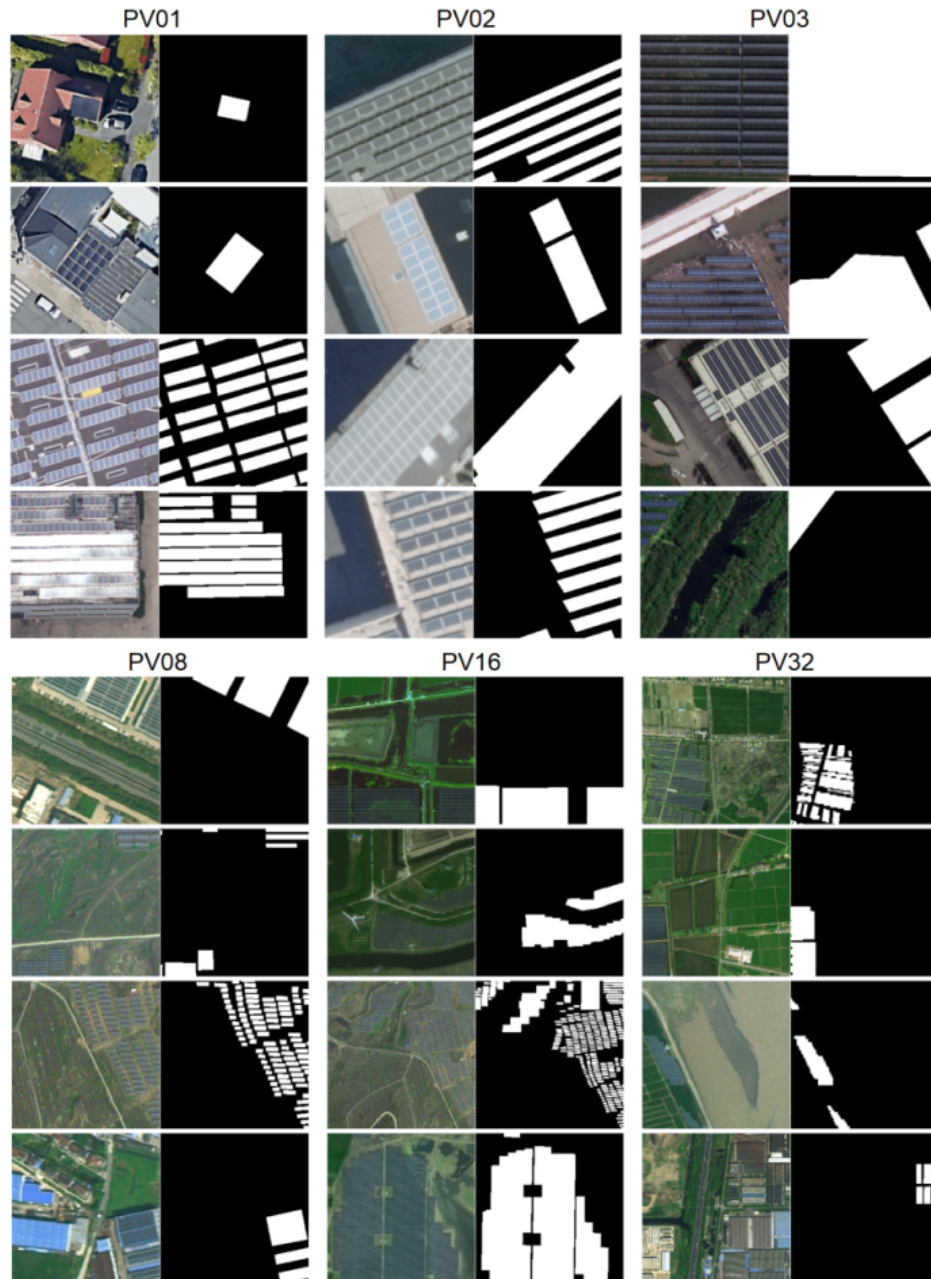
Preliminarily, six different resolution levels of 0.1 m (PV10), 0.2 m (PV20), 0.3 m (PV30), 0.8 m (PV80), 1.60 m (PV160), and 3.20 m (PV320) were defined. These resolution levels are chosen for two reasons: the limitation of the highest resolution of 0.1 m per pixel of the available training data and on the common zoom level (which ranges from 16 to 20) of the processing of image data from Bing Maps or GEE of approx. 0.1 m to 3.2 m. With the goal of dividing the available image data into the six image resolution classes, several preparatory steps were taken to prepare the image datasets. For the image data of all image resolutions, a consistent folder structure for images and masks was created that allows a static subdivision of the images for training, validation, and test. All images and masks were converted to a consistent image format. Furthermore, all masks were converted to binary format. For the dataset PV08, the images and masks (size of  $1024 \times 1024$  pixels) were adjusted into 16 frames ( $256 \times 256$ ). For the aggregated stages of 1.6 m resolution, the underlying images ( $1024 \times 1024$ ) were divided into four frames ( $512 \times 512$ ) and then, formatted to  $256 \times 256$  images using the nearest neighbor up-sampling method. Images with

3.2 m resolution were generated by directly up-sampling the PV08 images of size  $1024 \times 1024$  to  $256 \times 256$  with the nearest neighbor method. Then, the prepared image data and their corresponding masks were divided into the Training, Validation, and Test static folders by random selection. Special consideration was given to the division of the resolution levels between 0.8 and 3.2 m because the image contents show the same scenes. To avoid overfitting, care was taken to ensure that the same image content was stored in the Training, Validation, and Test folders. To keep the amount of data manageable, the amount of data for the comparisons of single-resolution and multi-resolution as well as hyperparameter tuning was limited to 1000 images per resolution level in training and 100 images per resolution level for both validation and testing phases. Thus, approximately 83.3%, 8.3%, and 8.3% of the image data is available for training, validation, and testing, respectively. The ratio of counted pixels between existing solar PV systems and the background class varies greatly within the datasets. The ratios of each dataset are shown in Table 3.2. Since several datasets with the same resolution were available, the images were combined and used for training. The only exception is the resolution level of 3.2 m with only 563 images in total in training because the whole dataset is limited to 763 images.

**Table 3.2:** Ratio of pixels containing solar PV systems per dataset.

| Dataset | Solar PV Pixels (%) |
|---------|---------------------|
| PV01    | 10.9                |
| PV02    | 19.2                |
| PV03    | 57.9                |
| PV08    | 36.7                |
| PV16    | 22.8                |
| PV32    | 12.8                |

Examples of the final image data are shown in Figure 3.1. The final model was then trained with all available images with the exception of images for validation and testing.



**Figure 3.1:** Illustration of the different training images, with each of their corresponding resolution level labeled at the top. The images with PV systems are shown on the left and the masks for training are shown on the right, with the PV systems highlighted in white.

### 3.4.2 Model

In choosing the model for the study, use was made of the existing DeepLabv3 network [18]. An essential key feature of the architecture is ASPP. ASPP is a DeepLab semantic segmentation module for resampling a given feature layer at multiple rates before convolution. The initial image is augmented with multiple filters of complementary effective fields of views. This allows objects as well as useful image context to be captured at multiple scales. A mapping strategy is implemented by using multiple parallel Atrous convolutional layers at different sampling rates instead of the resampling of features. Moreover, a spatial pyramid pooling (SPP) scheme is added. By using an SPP, parallel dilated convolutions with different rates can be applied to the input feature map. The feature map and dilated convolutions are then merged. Since objects of the same class can be represented at different sizes in the images used, ASPP can accommodate the different object sizes [26]. Finally, DeepLabv3 improves on previous versions of the model to deal with object segmentation problems at different scales. Modules have been developed to cascade or perform parallel Atrous convolutions to capture the context of multiple scales. For this purpose, several Atrous rates were introduced. Additionally, global context encoding has been added to the ASPP module, which further increases performance [18]. Combining DeepLabv3 with the deep residual neural network ResNet101 [19] as a backbone has repeatedly proven to be very effective for segmenting different image scenes [28, 29, 30]. Python 3.10, Torch 1.14.0, and Torchvision version 0.15.0 were all used for development of the model. Pre-trained weights from COCO “DeepLabV3 ResNet101 Weights COCO WITH VOC LABELS V1” were used for all training. Here, all models are evaluated with an independent test dataset that is included in the Pascal VOC dataset. The Pascal VOC dataset contains 20 categories. These weights were selected based on the performance indicated by the pixel accuracy of 92.4% as well as a mean Intersection over Union (IoU) of 67.4% on Torchvision v0.15 [31]. The segmentation model can be implemented by using a DeepLabV3 segmentation head that was set by integrating the ASPP module with an input channel of 2048 and different dilation rates (12, 24,

and 36) to capture multi-scale context information. The number of model parameters is approx. 61 million and the Giga Floating Point Operations Per Second (GFLOPS) are 258.74.

### 3.4.3 Model Training

#### Hyperparameter Tuning

During hyperparameter tuning, a full grid search was applied in which all possible parameter combinations were tested. Table 3.3 shows the search space metrics of the grid search where four loss functions, four optimizers, three learning rates, nine batch sizes, and four stride values are defined. These parameters are chosen due to the fact that the selected combined model, DeepLabV3 ResNet101, uses pre-trained weights. Therefore, hyperparameters were selected such that the model is as fine-tuned as possible to perform image segmentation. Direct parameter changes were avoided to preserve the integrity of the pre-trained weights, which ensures that the advantage of what using said weights brings is maintained. Examples of direct parameters omitted from tuning are number of layers, pooling type, and activation function types. The goal is to tune the listed parameters available with pre-trained weights such that evaluation metrics are maximized.

**Table 3.3:** Hyperparameter search space metrics.

| Hyperparameter | Search Range                            |
|----------------|---|
| Loss function  | Huber loss, BCE loss, MSE loss, CE loss |
| Optimizer      | Adagrad, Adam, RMSprop, SGD             |
| Learning rate  | 0.001, 0.0001, 0.00001                  |
| Batch size     | 2, 4, 8, 12, 16, 24, 32, 48, 64         |
| Stride         | 2, 3, 4, 5                              |

The parameters' loss function, optimizer, and learning rate were tested simultaneously. Conversely, batch size and stride parameters were tested independently and in isolation from other parameters. Additionally, an early stopping mechanism was utilized to prevent the model from overfitting to the training data. The mechanism involves evaluating the performance of

model every 10 epochs using the IoU metric. The best IoU value would be checked every 10 epochs and if the IoU value for that epoch is greater than the current best IoU, then the IoU value for that particular epoch would then be the new benchmark value for future performance checks. However, if the opposite occurs where the current best IoU is greater than the IoU value for that epoch, then the model training is stopped. Lastly, a maximum of 100 epochs is used for training.

Four loss functions are selected for tuning. The Huber loss function, invented by Huber in 1964, provides robustness as it has duality in its behavior in how to treat outlier and inlier errors [32]. According to Huber, the function will apply the squared error function (L2 loss) for error values below a particular threshold. However, if the error value is above the threshold, a scaled absolute (L1 loss) error loss function would be applied [32]. This, therefore, means that the Huber loss function behaves like a piecewise function. The threshold is called the delta variable ( $\delta$ ). The delta variable is the threshold for errors that are considered outliers. Furthermore, the Huber loss function is differentiable everywhere. When comparing different loss functions, it was concluded that dynamic loss functions (such as Huber loss) yielded better model performance during training regression CNN architecture [33]. Mathematically, the function can be described as follows:

$$Huberloss(y, f(x)) = \begin{cases} a^2 & |y - f(x)| \leq a \\ |y - f(x)| & |y - f(x)| > a \end{cases} \quad (3.1)$$

The Cross-Entropy loss (CE loss) function can be referred to as logarithmic loss. For a given set of predictions, the loss function simply averages the sum of the products of truth values ( $y$ ) and the log of the predicted values ( $p$ ) from the model.

$$CEloss(y, p) = - \sum_i y_i \log(p_i) \quad (3.2)$$

Binary Cross-Entropy (BCE) loss is a specific variation of Cross-Entropy loss tailored for binary classification tasks, where there are only two possible classes (e.g., 0 and 1). Unlike the standard CE loss function, BCE loss

calculates the error for predicting binary outcomes and is ideal for problems with binary class labels. The BCEWithLogits loss function is a combination of sigmoid and BCE loss and is numerically stable.

$$BCEWithLogitsloss(y, z) = -[y \cdot \log(\sigma(z)) + (1 - y) \cdot \log(1 - \sigma(z))] \quad (3.3)$$

Here,  $Y$  is the true label,  $z$  represents the logit or the output of your model, and  $\sigma(z)$  is the sigmoid function which converts the logits into probabilities. The BCE loss measures the difference between the predicted probabilities and the true labels. This loss function has been successfully implemented in applications such as video segmentation to minimize motion blur. Implementation involved using BCEWithLogits function in an UNet inspired encoder–decoder network [34]. Another study aimed to create a pix2pix Generative Adversarial Network (GAN) that can generate synthetic computerized tomography (sCT) images using positron emission tomography (PET) images [35]. The BCEWithLogits loss function was used for both the generator and discriminator of the GAN in combination with other loss functions. The last loss function tested is the Mean Squared Error (MSE) loss function:

$$MSEloss(y, f(x)) = \frac{1}{n} \sum_{i=1}^n (y_i - f(x_i))^2 \quad (3.4)$$

The MSE loss function is the squared difference between the ground truth ( $y$ ) and prediction value ( $f(x)$ ). Squaring the error helps the model penalize large errors. However, it should not be used with data with a sizable percentage of outliers for this reason.

The optimizers selected are the Adagrad, Adam, RMSProp, and SGD optimizers. The optimizer is responsible for backpropagation, whereby each weight is adjusted based on the error computed by the loss function. Due to backpropagation, each loss function must be differentiable as the derivative of it is found and used to adjust the weights. This means that the loss function

is the objective function as it is the function that the optimizer is attempting to minimize its output. Weights are adjusted by finding the derivative of the loss function with respect to weight and subtracting that derivative value from the original weight value. To further increase the rate of the optimization to minimize the loss function's output, the derivative value is multiplied by a factor so that the minimum is found more quickly. The multiplying factor is called the learning rate. This process continues until the minimum of the loss function is found, i.e., when the derivative of the loss function is zero and the weights can no longer be adjusted. This entire process is called Stochastic Gradient Descent (SGD). The Adagrad, Adam, and RMSProp optimizers are optimizers with adaptive learning rates. The Adaptive Gradient (Adagrad) optimizer implements an algorithm that controls proximal functions, which in turn modifies the "gradient steps of the algorithm" [36]. The study that introduced this approach concluded that adaptive optimizers' algorithms outperform non-adaptive algorithms [36]. The Adagrad optimizer is known for its adaptive learning rate approach. It is, therefore, well-suited for scenarios where the characteristics of the solar PV systems in images may vary widely. Adagrad can effectively navigate complex and varied landscapes by adjusting the learning rates for each parameter individually based on their historical gradients. This can be the case with PV segmentation. The Root Mean Square Propagation (RMSProp) algorithm was used to optimize the weights in recurrent neural network (RNN) using Long Short-term Memory (LSTM) units [37]. The model was trained to generate text sequences for handwritten text. The model was successful at both generating text as well as handwriting styles. By finding the running average of recent gradients for a particular weight, the value of that weight can be adjusted by dividing it by that average [37]. RMSProp, like Adagrad, incorporates adaptive learning rates but it further addresses the diminishing learning rate problem. In the context of PV systems segmentation, certain features may require more nuanced adjustments during optimization. RMSProp's ability could, therefore, be used to adaptively scale the learning rates for different parameters and lead to more efficient convergence. The Adaptive Moment Estimation (Adam) algorithm was introduced in 2014 by Kingma and Lei Ba [38]. It is

a gradient-based stochastic optimization algorithm. This algorithm requires fewer computational resources and less memory and is relatively straightforward to implement [38]. Using the Adam optimization takes advantage of both the Adagrad and RMSProp methods. Sparse gradients and non-stationary objectives are resolved using the intricacies deployed in both the Adagrad and RMSProp methods, respectively, [38]. Computational resources and memory efficiency are critical in the context of using large amounts of data in PV segmentation. The capability of the Adam optimizer to provide effective optimization with reduced resource requirements makes it an attractive optimizer to consider. The last optimizer to be tested is the standard SGD algorithm. This algorithm is non-adaptive, meaning that the learning rate is static. Standard SGD is a fundamental optimizer that can serve as a basis for comparison. In particular, SGD proves to be a reliable option in PV segmentation since the stability of certain features is provided for. In other words, less frequent adjustments to the learning rate are required. Due to its simplicity and stability, it serves as an important reference point for evaluating the performance of adaptive optimizers.

Three sequential learning rates are tested, each being an order of magnitude higher than its predecessor. Note that the learning rate is adjusted iteratively by multiplying factors if the optimizer is an adaptive optimizer. The magnitudes of factors depend on the selected optimizer to find the optimal adjustment value.

Batch size is the number of samples of training data inputted in the model and passes both forward and backward in the network. There is an inherent trade-off between accuracy and computational efficiency. If fast training is required, then a large batch size is preferred. However, larger batch sizes could lead to overfitting and an inaccurate model. Conversely, although a smaller batch size results in longer training times, it typically leads to higher accuracy.

Lastly, the stride of the model is tested. Stride refers to the step size over

which an image is passed over by a filter (or kernel). Higher strides lead to reduce computational requirements due to smaller feature map outputs.

### **Training with Single-Resolution Datasets**

During this task, models were trained using only single-level datasets per model. Therefore, a total of six models with single-resolution datasets of 0.1, 0.2, 0.3, 0.8, 1.60, and 3.20 m were trained. For best comparability, all parameter settings were kept constant. The parameter setting is based on the best results from the first sub-step of hyperparameter tuning. The Loss, Accuracy, Precision, Recall, F1-Score, and IoU were recorded every 10 epochs. Based on the IoU, the best model in the training process was saved and used for a subsequent independent test application. For each model, a test application was applied using all resolution levels.

### **Training with Multi-Resolution Datasets**

In the multi-resolution training, the models were not only trained using one single-resolution image dataset per model but also trained using image datasets of multiple resolutions. This was performed by extending the exposure of each model to training data step by step. Each model has therefore one more cycle of training than the previous model. During the exposure, that particular model is being trained using an image dataset that is of consecutive resolution than the previous training dataset. This process therefore yields six models in which each is trained with an additional dataset. Given that there are six datasets for each resolution, there would be six models. The training process then is as follows. The first model is trained using singularly 0.1 m resolution data. In addition to the single-resolution model for 0.1 m native resolution, the training data was first extended to include 0.2 m, then 0.3 m, 0.8 m, 1.6 m, and finally 3.2 m resolution data. Due to this experimental setup, six different models are available for subsequent comparison. The first model uses the 0.1 m dataset only, the following models then each incorporate a further dataset until all datasets are included in the last model for training. The respective models are called in the following 10, 10;

20, 10; 20; 30, ... to 10; 20; 30; 80; 160; 320.

### Validation Metrics

The quality of the models is measured by the following metrics: Accuracy, Precision, Recall, F1-Score, and IoU. Accuracy measures the overall correctness of the predictions by considering both true positive (TP) pixels representing PV systems and true negatives (TN) representing the background and comparing them to the total number of instances.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.5)$$

Precision calculates the ratio of correctly predicted PV system pixels (TP) to the total number of pixels predicted to be PV system (TP + false positives (FP)). Therefore, it provides information on how well the model correctly identifies PV system pixels and minimizes FP.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3.6)$$

Recall is a calculation of the ratio of correctly predicted PV system pixels (TP) to the total number of real PV system pixels (TP + FN). It gives an indication of how well the model captures the PV system pixels in the image, taking into account the potential false negatives (background pixels incorrectly classified as PV system pixels).

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3.7)$$

The F1-Score is a measure that combines Precision and Recall into a single metric, providing a balanced evaluation of a model's performance.

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3.8)$$

IoU measures the overlap between the predicted and ground truth regions,

which is often used in tasks like image segmentation or object detection.

$$\text{IoU} = \frac{TP}{TP + FP + FN} \quad (3.9)$$

## 3.5 Results

In the following section, the results are presented. Results are shown in the following order: the hyperparameter tuning, training with single-resolution datasets, and training with multi-resolution datasets.

### 3.5.1 Hyperparameter Tuning

The hyperparameter tuning process is divided into three steps. Firstly, the combination of optimizer, loss function, and learning rate is evaluated. Subsequently, the batch size is examined. Finally, the stride parameter is considered. The results of the combination of optimizer, loss, and learning rate is summarized in Table 3.4.

**Table 3.4:** Brief overview of the best hyperparameters tuning results.

| Loss  | Optimizer | Learning R. | Acc. (%) | Prec. (%) | Recall (%) | F1-Score (%) | IoU (%) |
|-------|-----------|-------------|----------|-----------|------------|--------------|---------|
| Huber | RMSprop   | 0.001       | 33.01    | 30.32     | 99.94      | 46.53        | 30.32   |
| BCE   | RMSprop   | 0.001       | 97.52    | 96.53     | 94.89      | 95.71        | 91.77   |
| BCE   | Adam      | 0.0001      | 97.68    | 95.90     | 96.15      | 96.02        | 92.35   |
| CE    | Adagrad   | 0.0001      | 77.76    | 56.74     | 99.94      | 72.38        | 56.72   |

The best values within each metric are highlighted in green. The full overview of hyperparameter tuning is given in Appendix Table 3.10.

The combination of Huber loss, RMSprop optimizer, and a learning rate of 0.001 as well as CE loss, Adagrad optimizer, and a learning rate of 0.0001 both achieved the best Recall of 99.94%. When using BCE loss with RMSprop optimizer and a learning rate of 0.001, the model achieved the best Precision with a score of 96.53%. The parameter setting with BCE loss as the loss function, Adam as the optimizer, and a learning rate of 0.0001 performs best on three of the five metrics, achieving an accuracy, F1-Score, and IoU of 97.68%, 96.02%, and 92.35% respectively. Full details of the results are

shown in Table 3.10 in the Appendix. The results of testing different batch sizes are summarized in Table 3.5.

**Table 3.5:** Metrics from batch size optimizations.

| Batch Size | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | IoU (%) |
|------------|--------------|---------------|------------|--------------|---------|
| 2          | 97.51        | 96.21         | 95.22      | 95.71        | 91.78   |
| 4          | 97.75        | 95.61         | 96.71      | 96.16        | 92.60   |
| 8          | 97.76        | 95.97         | 96.37      | 96.17        | 92.62   |
| 12         | 97.70        | 95.84         | 96.28      | 96.06        | 92.42   |
| 16         | 97.58        | 95.30         | 96.44      | 95.87        | 92.07   |
| 24         | 97.61        | 95.33         | 96.54      | 95.93        | 92.18   |
| 32         | 97.59        | 95.76         | 95.98      | 95.87        | 92.06   |
| 48         | 96.35        | 92.46         | 95.24      | 93.83        | 88.38   |
| 64         | 97.38        | 96.09         | 94.86      | 95.47        | 91.33   |

The best result for each metric is highlighted in green.

While batch sizes of 2 achieve the best results for precision and batch sizes of 4 the best results for recall, batch sizes of 8 are best not only for accuracy (97.76%) but also for F1-Score (96.17%) and IoU (92.62%). Finally, the different strides were tested (Table 3.6). Based on all performance metrics, a stride of 2 performs the best.

**Table 3.6:** Metrics for different stride values.

| Stride | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | IoU (%) |
|--------|--------------|---------------|------------|--------------|---------|
| 2      | 97.76        | 95.97         | 96.37      | 96.17        | 92.62   |
| 3      | 96.49        | 92.46         | 95.76      | 94.08        | 88.82   |
| 4      | 94.91        | 88.70         | 94.59      | 91.55        | 84.42   |
| 5      | 93.60        | 85.45         | 94.05      | 89.55        | 81.07   |

The best result for each metric is highlighted in green.

In summary, the best hyperparameter combination consists of BCE loss as the loss function, Adam as the optimizer, a learning rate of 0.0001, a batch size of 8, and a stride of 2, yielding optimal F1-Score and IoU outcomes. Consequently, this specific parameter configuration will be employed in all subsequent experiments. The results of the training with single image resolutions are explained in the following sub-section.

### 3.5.2 Single-Resolution Training

In the context of our analysis, we present the results of the single-resolution training utilizing the F1-Score in Table 3.7. Additionally, complementary results incorporating the IoU metric are provided in the Appendix Table 3.12. Listed against each other are the validation results with the resolutions used for training and data used for validation located on the horizontal axis and vertical axis, respectively. For an easier overview, matching resolutions from training and validation are highlighted in green. The best F1-Score of the validation is marked with \* for each validated resolution. Thus, the model with a resolution of 0.3 m achieves the best value with an F1-Score of 97.53% followed by the model with a resolution of 0.1 m (95.99%). The model with the lowest resolution (3.20 m) yields the worst performing results compared to the same resolution of training (86.59%) when validated. Overall, it can be seen from Table 3.7 that the best results are obtained in each case when validation is done using image data with a resolution that matches the resolution of each model’s respective training data.

**Table 3.7:** Single-resolution models F1-Score (measured in %) from validation.

| Resolution | 10      | 20      | 30      | 80      | 160     | 320     |
|------------|---------|---------|---------|---------|---------|---------|
| 10         | * 95.99 | 85.95   | 57.03   | 65.96   | 31.54   | 17.34   |
| 20         | 76.12   | * 95.40 | 6.45    | 78.18   | 14.12   | 2.76    |
| 30         | 68.92   | 30.14   | * 97.53 | 77.46   | 56.26   | 40.32   |
| 80         | 46.79   | 23.41   | 79.70   | * 95.53 | 85.47   | 48.82   |
| 160        | 16.81   | 8.87    | 68.64   | 91.18   | * 95.00 | 77.60   |
| 320        | 40.06   | 13.27   | 47.21   | 69.45   | 85.71   | * 86.59 |

The resolutions used in training are plotted horizontally, those used for validation vertically. The resolutions that match in both training and validation are highlighted in green. \* denotes the best F1-Score of the validation for each model.

In addition to the tabular representation of the validation, Figure 3.2 shows exemplary test applications.

### 3.5.3 Multi-Resolution Training

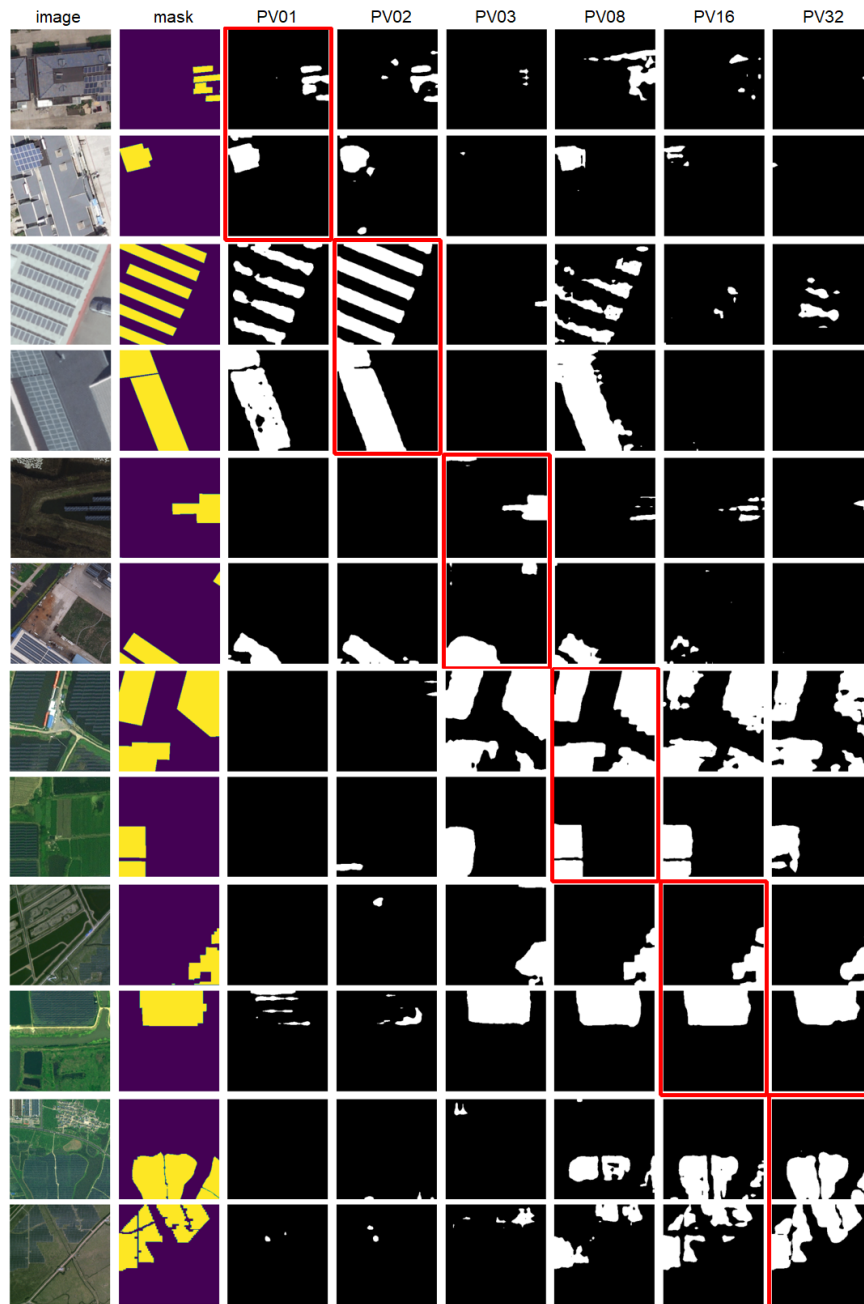
The results of the models from multi-resolution training are presented next. Firstly, the validation results using the F1-Score are summarized in Table 3.8 and the results using the IoU are summarized in Table 3.13 in the Appendix. Similar to the format of Table 3.7 in the previous section, matching resolutions from training and validation are highlighted in green for a better overview. In addition, an overall F1-Score is given, which is the validation of all datasets per model and is listed as the last row in Table 3.8 (“Overall”). This overall score can be interpreted as an average of the performances yielded from the different input image data. The tabular representation clearly illustrates that the step-wise extension of the data leads to better results, meaning that the model trained solely with 0.1 m resolution images results in the worst performance (57.45%), while the model trained with all resolutions performs best (95.27%).

**Table 3.8:** Multi-resolution models F1-Score (measured in %) from validation.

| Resolution | 10    | +20   | +30   | +80   | +160  | +320  |
|------------|-------|-------|-------|-------|-------|-------|
| 10         | 95.99 | 95.74 | 95.85 | 95.90 | 95.87 | 96.00 |
| 20         | 76.12 | 95.50 | 94.82 | 95.57 | 95.87 | 95.65 |
| 30         | 68.92 | 62.33 | 97.60 | 97.17 | 96.97 | 97.46 |
| 80         | 46.79 | 52.98 | 85.84 | 96.49 | 96.44 | 96.34 |
| 160        | 16.81 | 25.94 | 59.33 | 87.35 | 95.17 | 95.44 |
| 320        | 40.06 | 40.41 | 60.50 | 67.49 | 75.55 | 90.71 |
| Overall    | 57.45 | 62.15 | 82.32 | 89.99 | 92.64 | 95.27 |

The resolutions used in training are plotted horizontally, those used for validation vertically. The resolutions that match in both training and validation are highlighted in green.

Table 3.9 summarizes the results from both sub-sections of the specific single-resolution models and the final multi-resolution model in which the latter was trained with all training data. The multi-resolution model consistently outperforms the best single-resolution model with an average improvement of 0.93%. Only at the 0.3 m resolution level does the best single-resolution model perform better than its multi-resolution counterpart. However, it is also clear that the differences in the results are very small for almost



**Figure 3.2:** Illustration of the different predictions, where each exemplary image is labeled according to resolution levels. The images are sorted in ascending order of resolution. The horizontal label refers from left to right to the underlying image and the mask, next to the dataset used for training the respective models. The red box indicates where the resolution of the training and the test application are the same.

all models with deviations of less than 1.00%. The only exception here is the validation of the resolution of 320 with a difference of 4.12%.

**Table 3.9:** Side-by-side F1-Score from validation between single-resolution trained networks and the final multi-resolution trained network.

| Resolution | Best Single-Res Model F1-Score (%) | Multi-Res Model F1-Score (%) |
|------------|------------------------------------|------------------------------|
| 10         | 95.99                              | * 96.00                      |
| 20         | 95.40                              | * 95.65                      |
| 30         | * 97.53                            | 97.46                        |
| 80         | 95.53                              | * 96.34                      |
| 160        | 95.00                              | * 95.44                      |
| 320        | 86.59                              | * 90.71                      |
| Overall    | 94.34                              | * 95.27                      |

\* indicates the model with the best F1-Score.

With approx. 61 million model parameters, 258.7 GFLOPS, and the NVIDIA Tesla A100-SXM4 with 40 GB GPU graphics card used, the training times of the respective networks vary with the increase in datasets. For example, training the single-resolution networks with a batch size of 8 and 100 epochs takes approx. 23 min per run. In comparison, the multi-resolution training with two datasets and the same batch size and epochs took 45 min, and the multi-resolution training with all six data sets 136 min. The increase in datasets thus leads to a linear increase in run time.

Figure 3.3 shows results from a visual perspective of the applications from validation of the single-resolution trained networks as well as the final multi-resolution trained network. The input image, mask, single resolution output prediction, and multi-resolution output prediction are shown for each respective model going from left to right. From top to bottom, an example is shown for each resolution: 0.1 m, 0.2 m, 0.3 m, 0.8 m, 1.60 m, and 3.20 m.

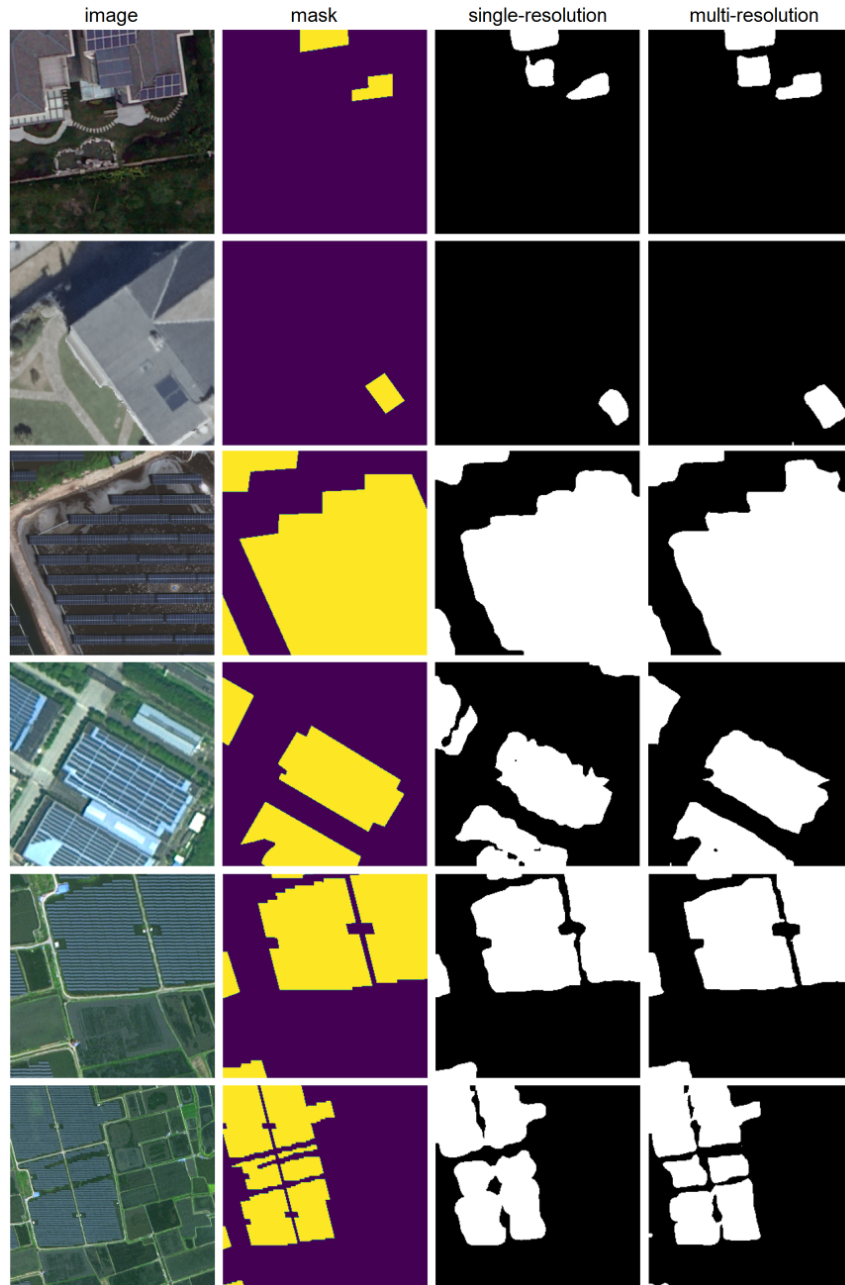
### 3.5.4 Final Model Configuration

The configurations of the final model are summarized below. BCE loss was used as the loss function, Adam as the optimizer, 0.0001 as the learning rate,

a batch size of 8, and a stride of 2. The ASSP segmentation head was set to 2048 input channels and 12, 24, and 36 dilation rates. The data were split into training, validation, and testing sets with ratios of 83.3%, 8.3%, and 8.3%, respectively. The training was set to 100 epochs. A single NVIDIA Tesla A100-SXM4 with 40 GB GPU memory and 512 GB CPU memory was used.

### 3.6 Discussion

Hyperparameter tuning results shown in Table 3.10 indicate that the best combination yielding the highest performance across most evaluation metrics was BCE loss, Adam optimizer, and 0.0001 as the loss function, optimizer, and learning rate, respectively. The BCE loss function is the best-performing loss function overall. This is based on the overall results seen in Table 3.10 and Figure 3.4. The BCE loss function tends to yield the highest performance based on the Accuracy, Precision, F1-Score, and IoU metrics. When comparing the optimizers' results, it can be seen that adaptive optimizers yield better performance than the standard SGD optimizer. The Adagrad, Adam, and RMSProp optimizers all perform relatively well under most parameter values. However, by focusing on the results from the BCE loss and varying the optimizers, the Adam optimizer outperforms the other optimizers two out of three times according to the Recall, F1-Score, and IoU metrics when all other parameters are kept constant. There are some trends that can be observed when comparing changes to the learning rate. For example, the combination of Huber loss function and Adagrad optimizer yields decreasing performance across all evaluation metrics as learning rate decreases. This statement is also true for the following loss function–optimizer combinations: MSE loss–SGD and MSE loss–Adagrad. However, the same is not true for the combinations of the Huber loss–Adam and Huber loss–RMSprop, in which performance is at its highest when a learning rate of 0.0001 is used. There is, thus, a disparity between the effect of learning rate and loss function–optimizer combinations on performance. This disparity may be due to the fact that optimizers and the learning rate affect the rate of SGD. In some loss



**Figure 3.3:** Comparison of the single-resolution trained networks versus the final multi-resolution trained network. From left to right, first are the images and masks, next are the predictions of the single-resolution networks suitable for each image. The predictions of the multi-resolution network are shown on the right.

function–optimizer combinations, increasing the learning rate leads to better convergence to a minimum loss, meaning SGD is functioning as intended. However, for combinations where the performance metrics vary when the learning rate changes, the optimal point is missed and, therefore, the error increases. When looking at the BCE loss, the best performing combination is the Adam optimizer with a learning rate of 0.0001. This is based on the fact that this combination yields the highest Accuracy, F1-Score, and IoU compared to all parameter combinations.

The results of trained models with single-resolution image data allow various conclusions. Firstly, it can be assumed that a model trained with the single-resolution dataset is very capable of segmenting pixels for images with the same resolution as the training set. The models often also show transfer capabilities when attempting to segment off images that have resolutions near to those of the images used to train the models. If the target resolution in the application is too far away from that of the training data, the performance of the application decreases significantly. The results measured by the F1-Score in Table 3.7 and IoU in Table 3.12 as well as the exemplary representation in Figure 3.2 lead to this conclusion. The results of training with several different resolutions also show clear patterns. If different resolutions are combined in models to train a network, the advantages of individual networks can be combined. Thus, a network trained with multiple resolutions is able to achieve very good performance for all the resolutions used. In addition, networks trained with multiple resolutions can even outperform networks trained using single-resolution data. This can be observed in both Figure 3.3 and Table 3.9. Figure 3.3 compares the results of a trained single-resolution model and a multiple-resolution model. The multi-resolution model outperforms the single-resolution model using validation data in five out of six resolutions based on F1-Score. Furthermore, when it comes to overall model performance, the multi-resolution model yields a higher F1-Score (95.27%) than the single-resolution model (94.34%). Additionally, from a visual standpoint, Table 3.9 shows the comparison between the outputs from both single and multi-resolution models. The multi-resolution model outputs clearly

match the mask as the output appears more rigid than the output from the single-resolution model.

The weakest performance is shown by the validation of the 3.2 m resolution data. This can be seen in both the single-resolution validation with an F1-Score of 86.59% (more than 8% worse than the validation of the other resolutions in Table 3.12) and in the multi-resolution validation with an F1-Score of 90.71% (more than 4% worse compared to the other values in Table 3.8). This could reflect the fact that at a resolution of 3.2 m per pixel, the size of complete solar panels is undercut, whereby individual panels are not even represented by a pixel. In this case, the network is presumably no longer able to recognize the internal structures of the solar panels as well. It should also be noted that the sample size of the dataset for 3.2 m resolution is smaller. There are only 563 images in total for training of the 3.2 m resolution model compared to the rest of the resolutions, which had 1000 images for each respective dataset for training. However, when comparing the single-resolution models with the multiple-resolution models in Table 3.9, it is also visible that accuracy increases as a result of training with different resolutions.

Based on these results, we can postulate the following: in order to obtain a model that is able to segment pixels as accurately as possible at a variety of resolutions, it is of considerable advantage to use training data at a wide variety of resolutions, as in the present work.

## 3.7 Conclusions

This paper presents a network that incorporates the DeepLabV3 ResNet101 architecture for segmenting solar PV systems at a variety of image resolutions. Trained on a wide range of different image data, our network is able to precisely detect PV systems in all tested image datasets. Additionally, the network outperforms almost all exclusively single-resolution trained networks. Through an intensive hyperparameter tuning, an ideal parameter setting is first determined. The combination of BCE loss as the loss func-

tion, Adam as the optimizer, a learning rate of 0.0001, a batch size of 8, and a stride of 2 proved to be optimal. To see how results differ between models trained at single and multiple resolutions, the present work provide a clear insight. Thus, when training using different image data, six different single-resolution-based networks are trained first. These networks perform well on their respective resolutions but perform very weakly on other resolutions. To resolve the issue of performance, a network is then trained such that it can outperform the six different single-resolution-based networks in terms of the metrics F1-Score, and IoU by subsequently training a network using all datasets. It can be shown that the use of different resolutions improves the overall performance of the models and allows for the application of the trained model to different image data. The network we have trained is the first of its kind, as it has been trained on a variety of different image resolutions and sensors. The network can also be applied to a wide variety of image data to detect and segment existing solar PV systems. The resulting network will be made freely available for further use.

In future work, we will use the trained networks for real application tests in different case study regions of the project OASES—“Development and Demonstration of a Sustainable Open Access AU-EU Ecosystem for Energy System Modelling” within LEAP-RE - Europe-Africa Research and Innovation Call on Renewable Energy.

### 3.8 Appendix

The following table summarizes the combinations of hyperparameter tuning, testing all parameter combinations of the loss functions (HuberLoss, BCEWithLogitsLoss (BCE), MSELoss, CrossEntropyLoss (CE)), the optimizers (Adagrad, Adam, RMSprop, SGD), and learning rate (0.001, 0.0001, 0.00001). The quality of the models is measured by the metrics Accuracy, Precision, Recall, F1-Score, and IoU.

**Table 3.10:** Hyperparameter tuning results.

| Loss  | Optimizer | Learning R. | Acc. (%) | Prec. (%) | Recall (%) | F1-Score (%) | IoU (%) |
|-------|-----------|-------------|----------|-----------|------------|--------------|---------|
| Huber | Adagrad   | 0.001       | 85.49    | 66.84     | 99.68      | 80.03        | 66.70   |
| Huber | Adam      | 0.001       | 78.73    | 57.83     | 99.86      | 73.25        | 57.79   |
| Huber | RMSprop   | 0.001       | 33.01    | 30.32     | 99.94      | 46.53        | 30.32   |
| Huber | SGD       | 0.001       | 63.79    | 44.60     | 99.79      | 61.64        | 44.55   |
| BCE   | Adagrad   | 0.001       | 97.27    | 94.93     | 95.75      | 95.34        | 91.09   |
| BCE   | Adam      | 0.001       | 97.59    | 95.03     | 96.79      | 95.90        | 92.12   |
| BCE   | RMSprop   | 0.001       | 97.52    | 96.53     | 94.89      | 95.71        | 91.77   |
| BCE   | SGD       | 0.001       | 93.32    | 82.52     | 97.81      | 89.52        | 81.02   |
| MSE   | Adagrad   | 0.001       | 86.73    | 68.79     | 99.78      | 81.43        | 68.68   |
| MSE   | Adam      | 0.001       | 70.44    | 49.66     | 99.90      | 66.34        | 49.64   |
| MSE   | RMSprop   | 0.001       | 33.24    | 30.39     | 99.92      | 46.60        | 30.38   |
| MSE   | SGD       | 0.001       | 70.43    | 49.65     | 99.86      | 66.33        | 49.62   |
| CE    | Adagrad   | 0.001       | 83.28    | 63.58     | 99.85      | 77.69        | 63.52   |
| CE    | Adam      | 0.001       | 70.84    | 0.00      | 0.00       | 0.00         | 0.00    |
| CE    | RMSprop   | 0.001       | 70.84    | 0.00      | 0.00       | 0.00         | 0.00    |
| CE    | SGD       | 0.001       | 81.01    | 60.59     | 99.84      | 75.41        | 60.53   |
| Huber | Adagrad   | 0.0001      | 80.79    | 60.31     | 99.83      | 75.19        | 60.24   |
| Huber | Adam      | 0.0001      | 83.91    | 64.48     | 99.75      | 78.33        | 64.38   |
| Huber | RMSprop   | 0.0001      | 84.07    | 64.70     | 99.85      | 78.52        | 64.64   |
| Huber | SGD       | 0.0001      | 51.28    | 37.34     | 98.94      | 54.22        | 37.20   |
| BCE   | Adagrad   | 0.0001      | 96.06    | 91.87     | 94.89      | 93.36        | 87.54   |
| BCE   | Adam      | 0.0001      | 97.68    | 95.90     | 96.15      | 96.02        | 92.35   |
| BCE   | RMSprop   | 0.0001      | 97.66    | 96.34     | 95.62      | 95.98        | 92.27   |
| BCE   | SGD       | 0.0001      | 83.59    | 67.50     | 84.30      | 74.97        | 59.96   |
| MSE   | Adagrad   | 0.0001      | 77.25    | 56.19     | 99.82      | 71.91        | 56.14   |
| MSE   | Adam      | 0.0001      | 88.94    | 72.56     | 99.80      | 84.03        | 72.46   |
| MSE   | RMSprop   | 0.0001      | 37.31    | 31.74     | 99.90      | 48.17        | 31.73   |
| MSE   | SGD       | 0.0001      | 53.64    | 38.55     | 99.33      | 55.55        | 38.45   |
| CE    | Adagrad   | 0.0001      | 77.76    | 56.74     | 99.94      | 72.38        | 56.72   |
| CE    | Adam      | 0.0001      | 84.54    | 65.37     | 99.90      | 79.03        | 65.33   |
| CE    | RMSprop   | 0.0001      | 87.82    | 70.62     | 99.71      | 82.68        | 70.48   |
| CE    | SGD       | 0.0001      | 84.63    | 65.52     | 99.82      | 79.12        | 65.45   |
| Huber | Adagrad   | 0.00001     | 71.36    | 50.45     | 99.55      | 66.97        | 50.34   |
| Huber | Adam      | 0.00001     | 79.81    | 59.11     | 99.83      | 74.26        | 59.05   |
| Huber | RMSprop   | 0.00001     | 89.82    | 74.19     | 99.80      | 85.11        | 74.08   |
| Huber | SGD       | 0.00001     | 42.96    | 33.34     | 95.63      | 49.44        | 32.84   |

**Table 3.11:** *Cont.*

| Loss | Optimizer | Learning R. | Acc. (%) | Prec. (%) | Recall (%) | F1-Score (%) | IoU (%) |
|------|-----------|-------------|----------|-----------|------------|--------------|---------|
| BCE  | Adagrad   | 0.00001     | 90.51    | 77.09     | 95.95      | 85.49        | 74.66   |
| BCE  | Adam      | 0.00001     | 96.96    | 94.58     | 95.02      | 94.80        | 90.11   |
| BCE  | RMSprop   | 0.00001     | 97.25    | 95.46     | 95.09      | 95.28        | 90.98   |
| BCE  | SGD       | 0.00001     | 74.73    | 66.51     | 26.85      | 38.25        | 23.65   |
| MSE  | Adagrad   | 0.00001     | 64.66    | 45.20     | 99.79      | 62.22        | 45.16   |
| MSE  | Adam      | 0.00001     | 85.13    | 66.28     | 99.73      | 79.64        | 66.16   |
| MSE  | RMSprop   | 0.00001     | 90.45    | 75.41     | 99.77      | 85.90        | 75.28   |
| MSE  | SGD       | 0.00001     | 45.61    | 34.58     | 97.07      | 51.00        | 34.23   |
| CE   | Adagrad   | 0.00001     | 79.24    | 58.47     | 99.43      | 73.64        | 58.27   |
| CE   | Adam      | 0.00001     | 77.67    | 56.63     | 99.92      | 72.29        | 56.61   |
| CE   | RMSprop   | 0.00001     | 73.20    | 52.11     | 99.86      | 68.49        | 52.08   |
| CE   | SGD       | 0.00001     | 80.54    | 60.00     | 99.80      | 74.94        | 59.93   |

The best values within each metric are highlighted in green.

**Table 3.12:** Single-resolution models IoU performance (measured in %) from validation.

| Resolution | 10      | 20      | 30      | 80      | 160     | 320     |
|------------|---------|---------|---------|---------|---------|---------|
| 10         | * 92.28 | 75.37   | 39.89   | 49.21   | 18.72   | 9.49    |
| 20         | 61.45   | * 91.21 | 3.33    | 64.18   | 7.6     | 1.40    |
| 30         | 52.58   | 17.74   | * 95.18 | 63.22   | 39.14   | 25.25   |
| 80         | 30.54   | 13.26   | 66.25   | * 91.44 | 74.63   | 32.30   |
| 160        | 9.18    | 4.64    | 52.26   | 83.78   | * 90.48 | 63.39   |
| 320        | 25.04   | 7.10    | 30.90   | 53.20   | 74.99   | * 76.35 |

The resolutions used in training are plotted horizontally, those used for validation vertically. The resolutions that match in both training and validation are highlighted in green.

\* denotes the best IoU of the validation for each model.

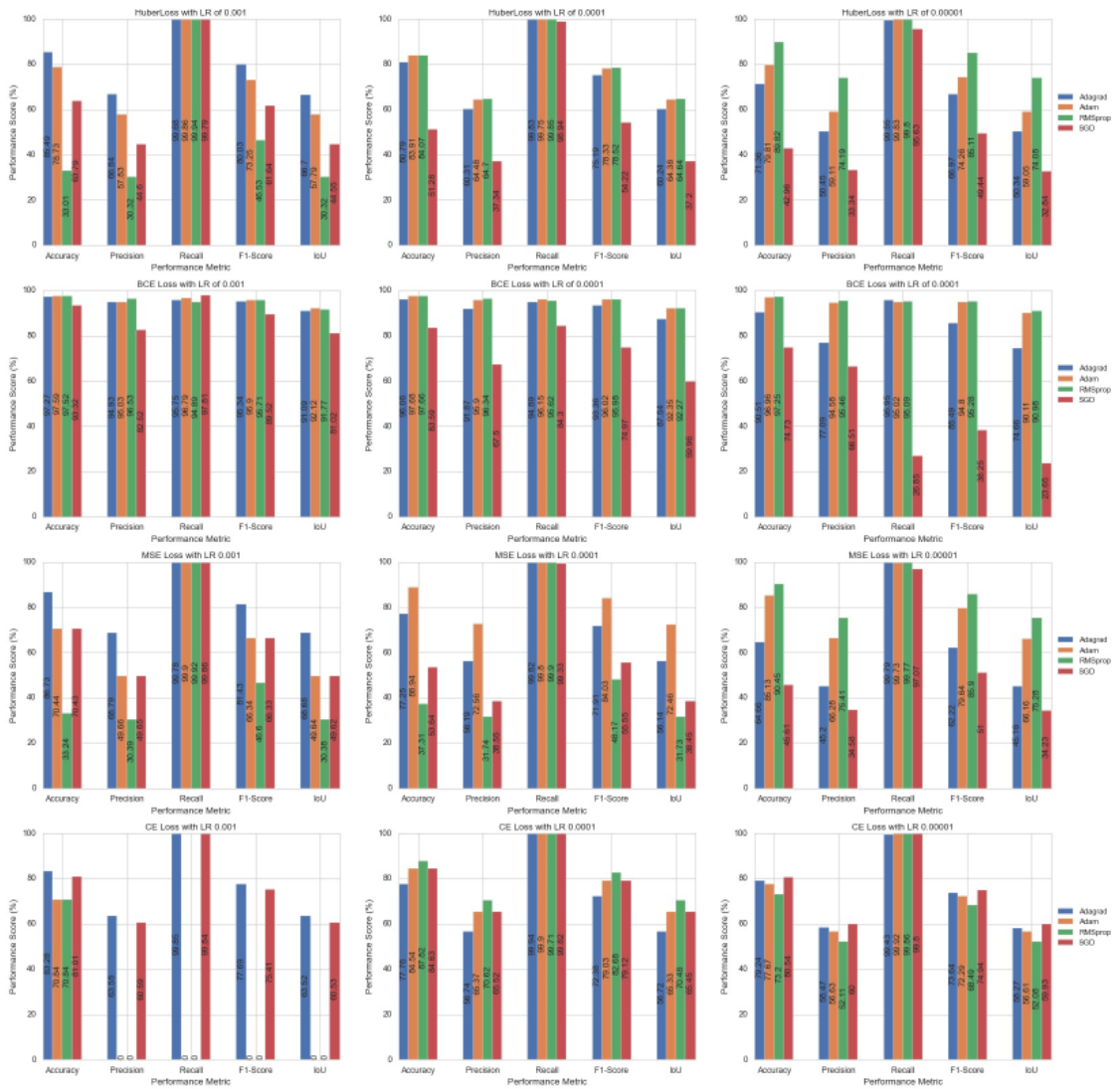


Figure 3.4: Illustration of the different hyperparameter tuning results

**Table 3.13:** Multi-resolution models IoU performance (measured in %) from validation.

| Resolution | 10    | +20   | +30   | +80   | +160  | +320  |
|------------|-------|-------|-------|-------|-------|-------|
| 10         | 92.28 | 91.82 | 92.04 | 92.12 | 92.08 | 92.31 |
| 20         | 61.45 | 91.38 | 90.15 | 91.51 | 92.07 | 91.66 |
| 30         | 52.58 | 45.27 | 95.32 | 94.49 | 94.12 | 95.05 |
| 80         | 30.54 | 36.04 | 75.2  | 93.21 | 93.13 | 92.94 |
| 160        | 9.18  | 14.91 | 42.17 | 77.55 | 90.78 | 91.27 |
| 320        | 25.04 | 25.32 | 43.37 | 50.93 | 60.71 | 82.99 |
| Overall    | 45.18 | 50.79 | 73.04 | 83.30 | 87.15 | 91.04 |

The resolutions used in training are plotted horizontally, those used for validation vertically. The resolutions that match in both training and validation are highlighted in green.

## Bibliography

- [1] Ibrahim Dincer. Renewable energy and sustainable development: a crucial review. *Renewable and sustainable energy reviews*, 4(2):157–175, 2000.
- [2] Markus Schlott, Bruno Schyska, Dinh Thanh Viet, Vo Van Phuong, Duong Minh Quan, Ma Phuoc Khanh, Fabian Hofmann, Lueder von Bremen, Detlev Heinemann, and Alexander Kies. Pypsa-vn: An open model of the vietnamese electricity system. In *2020 5th International Conference on Green Technology and Sustainable Development (GTSD)*, pages 253–258. IEEE, 2020.
- [3] Nelli Putkonen, Tomi Lindroos, Eimantas Neniškis, Diāna Žalostība, Egidijus Norvaiša, Arvydas Galinis, Jana Teremranova, and Juha Kiviluoma. Modeling the baltic countries’ green transition and desynchronization from the russian electricity grid. *International Journal of Sustainable Energy Planning and Management*, 34:45–62, 2022.
- [4] Maximilian Parzen, Hazem Abdel-Khalek, Ekaterina Fedotova, Matin Mahmood, Martha Maria Frysztacki, Johannes Hampp, Lukas Franken, Leon Schumm, Fabian Neumann, Davide Poli, Aristides Kiprakis, and Davide Fioriti. Pypsa-earth. a new global open energy system optimization model demonstrated in africa. *Applied Energy*, 341(4):121096, 2023.
- [5] Jorge Felipe Gaviria, Gabriel Narváez, Camilo Guillen, Luis Felipe Giraldo, and Michael Bressan. Machine learning in photovoltaic systems: A review. *Renewable Energy*, 2022.
- [6] Jordan M. Malof, Leslie M. Collins, and Kyle Bradbury. A deep convolutional neural network, with pre-training, for solar photovoltaic array detection in aerial imagery. In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 874–877, 2017.

- 
- [7] Jiafan Yu, Zhecheng Wang, Arun Majumdar, and Ram Rajagopal. Deep-solar: A machine learning framework to efficiently construct a solar deployment database in the united states. *Joule*, 2(12):2605–2617, 2018.
- [8] Marcus Vinícius Coelho Vieira da Costa, Osmar Luiz Ferreira de Carvalho, Alex Gois Orlandi, Issao Hirata, Anesmar Olineo de Albuquerque, Felipe Vilarinho e Silva, Renato Fontes Guimarães, Roberto Arnaldo Trancoso Gomes, and Osmar Abílio de Carvalho Júnior. Remote sensing for monitoring photovoltaic solar plants in brazil using deep semantic segmentation. *Energies*, 14(10):2960, 2021.
- [9] Lucas Kruitwagen, KT Story, J Friedrich, L Byers, S Skillman, and Cameron Hepburn. A global inventory of photovoltaic solar energy generating units. *Nature*, 598(7882):604–610, 2021.
- [10] Kevin Mayer, Zhecheng Wang, Marie-Louise Arlt, Dirk Neumann, and Ram Rajagopal. Deepsolar for germany: A deep learning framework for pv system mapping from aerial imagery. In *2020 International Conference on Smart Energy Systems and Technologies (SEST)*, pages 1–6. IEEE, 2020.
- [11] Maximilian Kleebauer, Daniel Horst, and Christoph Reudenbach. Semi-automatic generation of training samples for detecting renewable energy plants in high-resolution aerial images. *Remote Sensing*, 13(23):4793, 2021.
- [12] Simiao Ren, Jordan Malof, Rob Fetter, Robert Beach, Jay Rineer, and Kyle Bradbury. Utilizing geospatial data for assessing energy security: Mapping small solar home systems using unmanned aerial vehicles and deep learning. *ISPRS International Journal of Geo-Information*, 11(4):222, 2022.
- [13] Hongzhi Mao, Xie Chen, Yongqiang Luo, Jie Deng, Zhiyong Tian, Jinghua Yu, Yimin Xiao, and Jianhua Fan. Advances and prospects on estimating solar photovoltaic installation capacity and potential based

- on satellite and aerial images. *Renewable and Sustainable Energy Reviews*, 179(7830):113276, 2023.
- [14] Rui Zhu, Dongxue Guo, Man Sing Wong, Zhen Qian, Min Chen, Bisheng Yang, Biyu Chen, Haoran Zhang, Linlin You, Joon Heo, et al. Deep solar pv refiner: A detail-oriented deep learning network for refined segmentation of photovoltaic areas from satellite imagery. *International Journal of Applied Earth Observation and Geoinformation*, 116:103134, 2023.
- [15] Hongjun Tan, Zhiling Guo, Haoran Zhang, Qi Chen, Zhenjia Lin, Yuntian Chen, and Jinyue Yan. Enhancing pv panel segmentation in remote sensing images with constraint refinement modules. *Applied Energy*, 350:121757, 2023.
- [16] Jianxun Wang, Xin Chen, Weiyue Shi, Weicheng Jiang, Xiaopu Zhang, Li Hua, Junyi Liu, and Haigang Sui. Rooftop pv segmenter: A size-aware network for segmenting rooftop photovoltaic systems from high-resolution imagery. *Remote Sensing*, 15(21), 2023.
- [17] Hou Jiang, Ling Yao, Ning Lu, Jun Qin, Tang Liu, Yujun Liu, and Chenghu Zhou. Multi-resolution dataset for photovoltaic panel segmentation from satellite and aerial imagery. *Earth System Science Data*, 13(11):5389–5401, 2021.
- [18] About Bing and Microsoft News Data Suppliers, <https://bingexplore.azurewebsites.net/bing-data-suppliers/en/>, Accessed: 2023-08-17.
- [19] Myroslava Lesiv, Linda See, Juan Carlos Laso Bayas, Tobias Sturn, Dmitry Schepaschenko, Mathias Karner, Inian Moorthy, Ian McCallum, and Steffen Fritz. Characterizing the spatial and temporal availability of very high resolution satellite imagery in google earth and microsoft bing maps as a source of reference data. *Land*, 7(4):118, 2018.
- [20] Buyu Su, Xiaoping Du, Haowei Mu, Chen Xu, Xuecao Li, Fang Chen, and Xiaonan Luo. Fepvnet: A network with adaptive strategies for cross-

- scale mapping of photovoltaic panels from multi-source images. *Remote Sensing*, 15(9):2469, 2023.
- [21] Yinda Wang, Danlu Cai, Luanjie Chen, Lina Yang, Xingtong Ge, and Ling Peng. A downscaling methodology for extracting photovoltaic plants with remote sensing data: From feature optimized random forest to improved hrnet. *Remote Sensing*, 15(20), 2023.
- [22] Zhiling Guo, Zhan Zhuang, Hongjun Tan, Zhengguang Liu, Peiran Li, Zhengyuan Lin, Wen-Long Shang, Haoran Zhang, and Jinyue Yan. Accurate and generalizable photovoltaic panel segmentation using deep learning for imbalanced datasets. *Renewable Energy*, 219:119471, 2023.
- [23] Sebastian Dunnett, Alessandro Sorichetta, Gail Taylor, and Felix Eigenbrod. Harmonised global datasets of wind and solar farm locations and power. *Scientific Data*, 7(1):1–12, 2020.
- [24] Gabriel Kasmi, Yves-Marie Saint-Drenan, David Trebosc, Raphaël Jolivet, Jonathan Leloux, Babacar Sarr, and Laurent Dubus. A crowd-sourced dataset of aerial images with annotated solar photovoltaic arrays and installation metadata. *Scientific data*, 10(1):59, 2023.
- [25] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*, 2017.
- [26] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4):834–848, 2017.
- [27] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

- [28] Yaya Heryadi, Edy Irwansyah, Eka Miranda, Haryono Soeparno, Kiyota Hashimoto, et al. The effect of resnet model as feature extractor network to performance of deeplabv3 model for semantic satellite image segmentation. In *2020 IEEE Asia-Pacific Conference on Geoscience, Electronics and Remote Sensing Technology (AGERS)*, pages 74–77. IEEE, 2020.
- [29] Yuyuan Liu, Yu Tian, Yuanhong Chen, Fengbei Liu, Vasileios Belagiannis, and Gustavo Carneiro. Perturbed and strict mean teachers for semi-supervised semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4258–4267, June 2022.
- [30] Jia-Ji Wang, Yu-Fei Liu, Xin Nie, and YL Mo. Deep convolutional neural networks for semantic segmentation of cracks. *Structural Control and Health Monitoring*, 29(1):e2850, 2022.
- [31] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019.
- [32] Peter J. Huber. *Robust Estimation of a Location Parameter*, pages 492–518. Springer New York, New York, NY, 1992.
- [33] Fares Bougourzi, Fadi Dornaika, Nagore Barrena, Cosimo Distante, and Abdelmalik taleb ahmed. Cnn based facial aesthetics analysis through dynamic robust losses and ensemble regression. *Applied Intelligence*, 53:1–18, 08 2022.
- [34] Zijian Kuang and Xinran Tie. Flow-based video segmentation for human head and shoulders. *CoRR*, abs/2104.09752, 2021.

- 
- [35] Kouhei Nakanishi, Seiichi Yamamoto, and Tadashi Watabe. Prediction of ct images from pet images using deep learning approach for small animal systems. In *2021 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, pages 1–3, 2021.
- [36] John Duchi, Elad Hazan, and Yoram Singer. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12(61):2121–2159, 2011.
- [37] Alex Graves. Generating sequences with recurrent neural networks. *arXiv preprint arXiv:1308.0850*, 2013.
- [38] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

## Chapter 4

# Globally Scalable, QGIS-Integrated Workflow for Solar Photovoltaic System Segmentation and Capacity Estimation: A Case Study in Algeria

## Bibliographic Information

Kleebauer, M., Hafdaoui, H., Bouchakour, S., Hückner, B., Lindenmeyer, M. (2025). Globally Scalable, QGIS-Integrated Workflow for Solar Photovoltaic System Segmentation and Capacity Estimation: A Case Study in Algeria. IGARSS 2025 - 2025 IEEE International Geoscience and Remote Sensing Symposium, Brisbane, Australia, 2025.

## Author's contribution

In the following paragraph, the individual contributions of the authors are briefly broken down with regard to the publication. The authors are abbreviated as follows: Maximilian Kleebauer (M.K.), Hichem Hafdaoui (H.H.), Salim Bouchakour (S.B.), Benedikt Hückner (B.H.), and Malte Lindenmeyer (M.L.). Concept, M.K.; methodology, M.K.; software, M.K.; validation, M.K.; formal analysis, M.K.; investigation, M.K. and H.H.; resources, M.K.; data curation, M.K.; writing—original draft, M.K.; writing—review and editing, M.K., H.H., S.B., B.H., and M.L.; visualization, M.K.; supervision, M.K.; project administration, M.K., M.L.; funding acquisition, M.K. All authors have read and agreed to the published version of the manuscript.

## Copyright Notice

*Author accepted manuscript (AAM)*. This is the author's version of the work accepted for presentation at the 2025 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2025), Brisbane, Australia, 3–8 August 2025. The Version of Record will be available on IEEE Xplore upon publication. © 2025 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

## 4.1 Abstract

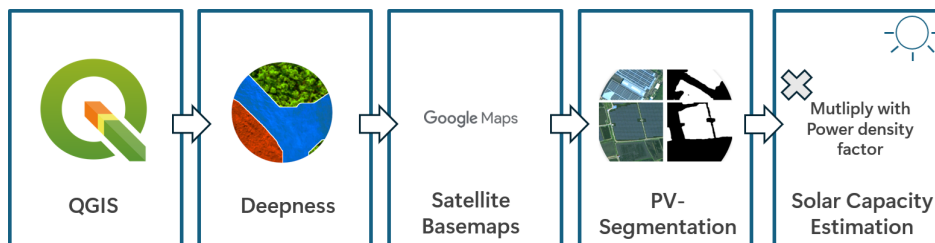
This study presents a deep learning-based approach for the segmentation of Solar Photovoltaic (PV) systems using a ResNet-based DeepLabV3 architecture. The method is seamlessly integrated into the open-source Geographic Information System QGIS environment through the user-friendly Deepness plugin and allows intuitive application on geospatial base maps anywhere in the world. A case study was conducted in the Piat region of Algeria, focusing on seven large PV systems within an autonomous power grid. After segmentation, a power density factor in megawatt-peak (MWp) was used to calculate the capacity of each identified PV system. The calculated capacities were compared with the official values of the operators, showing a mean relative error of 8.5 %. To further refine the capacity estimates, we introduce a locally adjusted factor of 73.5 MWp/km<sup>2</sup>, which reflects the expected results for the region more precisely. The approach is globally applicable, easily transferable to different geographical contexts and accessible to users without programming knowledge, offering significant potential for renewable energy analysis. The dataset is publicly available for download [1] and can be used interactively [2].

**Keywords** solar photovoltaic systems; renewable energy; image segmentation; remote sensing application

## 4.2 Introduction

The global transition towards renewable energy sources has increased the demand for efficient and scalable methods for analyzing and monitoring PV systems. Accurate segmentation of PV installations from remote sensing data is a critical step in enabling their rapid deployment and effective integration into energy systems, as highlighted in IRENA's 'Global Renewables Outlook: Energy Transformation 2050' [3]. Remote sensing and deep learning have emerged as powerful tools for object detection [4], regression [5], and segmentation [6] of PV systems from aerial and satellite imagery. De-

spite significant progress, most deep learning based models are trained on geographically limited datasets, which leads to challenges in their global applicability. In particular, regions in Northern Africa have a high potential for solar energy, but are still underrepresented in training datasets. This may affect the model’s performance. To address this challenge, we present a practical workflow for detecting and analyzing PV systems from satellite imagery. This paper presents the application for the segmentation of PV systems, combining high-resolution satellite data with deep learning modeling. Figure 4.1 shows the five key steps of the workflow. Using QGIS and the Deepness plugin, satellite basemaps such as Google Maps can be processed directly, allowing the segmentation of PV systems without the need for programming knowledge. The segmented areas are then used to estimate installed capacity by applying a solar-specific power density factor. The segmentation model used in this study was previously developed and trained on diverse datasets, as described in [7]. This workflow enables scalable and user-friendly analysis of solar installations from remote sensing data.



**Figure 4.1:** Workflow for segmentation and capacity estimation of PV systems.

This study is conducted within the framework of the Development and Demonstration of a Sustainable Open Access African Union-European Union Ecosystem for Energy System Modelling (OASES) project. The project focuses on creating an integrated open-source model chain for energy system modeling, ensuring user-friendliness and accessibility for users globally. The ecosystem enables renewable energy analyses with minimal barriers and supports transparency and adoption of open-source tools for energy planning and analysis. Key components of the OASES model chain include renewable energy system detection [7], high-resolution time series generation [8],

and energy system modeling integrated into platforms like the IRENA Flex-Tool [9, 10]. These advancements showcase the potential of open-source strategies to streamline renewable energy planning and forecasting processes, making them accessible and effective for diverse user groups.

### 4.3 Study Area

Algeria, located in North Africa, is the largest country on the continent, covering an area of approximately 2.38 million square kilometers. The country is geographically diverse, with vast desert landscapes in the south and a Mediterranean coastline in the north. Its climate varies significantly from the dry desert regions in the south to the temperate zones in the northern coastal areas, influenced by both the Saharan climate and the Mediterranean Sea. The southern desert regions, such as the Piat (Plateau of the High Atlas) region, are characterized by hot, dry and sunny conditions with high solar radiation, making it an ideal location for solar energy projects. This region, situated in central and southern Algeria, is particularly known for its low population density, vast flat terrain, and minimal vegetation. In fact, the average solar radiation in the region is between 5 to 7 kWh/m<sup>2</sup> per day, which is among the highest in the world [11, 12].

Algeria's Piat region relies heavily on gas turbine-based power generation, which poses logistical and environmental challenges. The region's enormous solar potential, estimated at over 22,000 MWp, remains largely untapped, with a capacity of just 53 MWp in 2024. In total, there are seven systems that were installed between 2015 and 2016 [13]. Transitioning to solar energy offers a sustainable solution to reduce fuel dependence, lower costs, and improve energy security. Algeria aims to generate 27 % of its electricity from renewables by 2030, with solar as a key contributor [14, 15, 16]. Integrating hybrid systems—combining gas turbines with solar and wind power—could enhance grid reliability while reducing the carbon footprint, aligning with national climate goals [17].

## 4.4 Materials and Methods

### 4.4.1 PV-Segmentation Model

The PV-Segmentation model was developed based on an approach that uses a DeepLabV3 ResNet101 architecture. The architecture leverages the DeepLabV3 network [18] with Atrous Spatial Pyramid Pooling (ASPP) to capture multi-scale context effectively. By integrating ResNet101 [19] as backbone, the model is well-suited for complex segmentation tasks. Various image datasets were used for training, including Unmanned Aerial Vehicles, aerial, and satellite imagery, at six different resolutions ranging from 0.1 m to 3.2 m. The training data sets were collected from different regions around the world, including France, Germany, and China, to ensure a broad representation of PV system characteristics. However, no training data from African countries was included, which could limit the model's performance in Algeria.

To achieve high segmentation accuracy and generalizability across different image sources, an extensive hyperparameter search was conducted. Details of this tuning process are described in [7], which identified the following final parameter combination as most effective. The model was trained for 100 epochs using binary cross-entropy (BCE) as the loss function, with the Adam optimizer, a learning rate of 0.0001, a batch size of 8, and a stride of 2. The ASPP segmentation head was configured with 2048 input channels and dilation rates of 12, 24, and 36. The data were split into training, validation, and testing sets with ratios of 83.3 %, 8.3 %, and 8.3 %, respectively. Model training was executed on a single NVIDIA Tesla A100-SXM4 with 40 GB memory and 512 GB CPU memory.

### 4.4.2 Deepness: Deep Neural Remote Sensing Plugin

The Deepness plugin is an open-source extension for QGIS that allows users to apply machine learning models directly to raster layers such as geospatial data or other matrix-based data [20]. It supports tasks such as object detection, regression, and segmentation, seamlessly integrating computer vision

techniques into GIS workflows. With a model zoo that provides ready-to-use neural network models—including the freely available one used in this study—Deepness makes deep learning accessible to users without prior machine learning experience, adding modern deep learning based capabilities to traditional geospatial analyses. Deepness supports the use of remote sensing-based deep learning methods in the format open neural network exchange (ONNX). These models can be applied directly through the tool’s graphical user interface to perform tasks such as detection or segmentation on basemaps like Google Satellite.

The following parameter settings were used for the segmentation task: *Input Channels Mapping* - Advanced settings with Band 1 assigned to red, Band 2 to green, and Band 3 to blue; *Resolution* - 80 cm/px; *Tile Size* - 1000 pixels; *Batch Size* - 1; *Tile Overlap* - 10 %; *Class Probability Threshold* - Applied based on self-validation with values between 0.3 and 0.5; *Removal of Small Segments* - Areas smaller than 20 pixels were filtered out; These settings and post-processing steps ensured accurate segmentation results while maintaining computational efficiency and minimizing noise in the output.

### 4.4.3 Google Satellite Basemap

Google Satellite provides high-resolution satellite imagery that can be used for geospatial analysis in a variety of disciplines. This dataset is accessible via Google Maps Satellite [21], this dataset offers detailed and realistic visual representations of the Earth’s surface, including features such as buildings, roads, and natural landscapes. The imagery includes true-color (red-green-blue) channels, which facilitate the accurate identification and distinction of surface features. In the context of renewable energy analysis, Google Satellite imagery is instrumental for validating and correcting geo-datasets, particularly for identifying and verifying PV systems. The high resolution and frequent updates of this dataset provide comprehensive coverage and support the integration of machine learning techniques for scalable and reliable geospatial data analysis. Furthermore, Google Satellite imagery can be integrated as a basemap in QGIS through the XYZ Tiles service. This

allows users to overlay and analyze high-resolution imagery within the QGIS environment.

#### 4.4.4 Validation and Solar Capacity Estimation

To determine the most suitable resolution, the segmentation model was initially tested on one PV system using the six trained input resolutions (ranging from 10, 20, 30, 80, 160 to 320 cm). The calculation of installed solar capacity was based on an estimated power density factor of approx. 80 MWp per kilometer<sup>2</sup> for fixed-tilt PV systems in 2016. This estimate reflects land requirements for utility-scale PV systems, as outlined in the study "Land Requirements for Utility-Scale PV: An Empirical Update on Power and Energy Density" [22]. In this study, a detailed analysis of power and energy densities in large-scale PV systems is carried out, taking into account differences in efficiency, design and space requirements. The derived power density factor allows for a straightforward conversion of segmented PV areas into their estimated capacity, offering a practical approach for renewable energy analysis. By applying this factor to the results, the installed capacity of PV systems in the Piat region is estimated.

## 4.5 Results

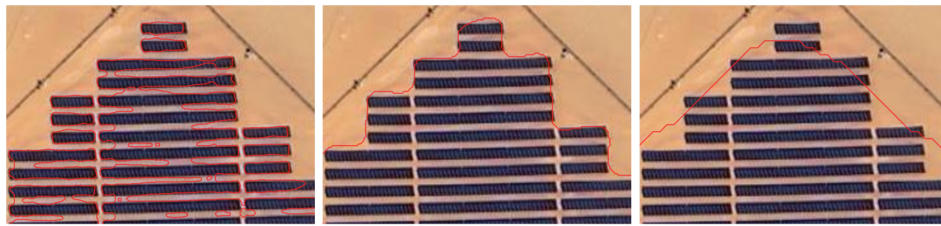
This section presents the findings of the segmentation and capacity estimations. To determine the most suitable resolution, the model was initially tested on the PV system in Reggane using all available input resolutions, as shown in Figure 4.2.

A visual inspection of the segmentation results revealed that large-scale PV system boundaries are clearly identified even at coarser resolutions. With regard to the level of detail, however, significant differences can be observed in Figure 4.3: While the high-resolution inputs of 10 cm, 20 cm and 30 cm led to detection at panel level, the results at 80, 160 and 320 cm predominantly segmented the large-scale arrangement of PV systems.

After examining various input resolutions, we chose 80 cm/px for all re-



**Figure 4.2:** Comparison of segmentation results at different resolutions for the Reggane PV system. From left to right: 20, 80, and 320 cm. The red overlay shows the results from segmentation.



**Figure 4.3:** Detailed comparison of segmentation results for the Reggane PV system at three selected input resolutions (20, 80, and 320 cm). The red overlays illustrate how segmentation performance varies between fine-grained panel-level detection and large-scale structure identification.

maining analyses. This setting balances segmentation quality with computational efficiency, accurately capturing system boundaries while limiting runtime. Since processing time increases quadratically with resolution, 80 cm/px is especially suitable for large, ground-mounted PV parks, where it preserves key structural details without becoming prohibitively slow. Figure 4.4 illustrates the segmentation results for the six remaining PV systems in the Piat region, overlaid on a high-resolution satellite basemap.

The largest system, Adrar, covers 28.41 hectares, while the smallest system, Kabertene, is 3.89 hectares in size. The total area is 75.86 hectares. The distribution of system sizes indicates a dominance of medium-to-large installations in the Piat region. Table 4.1 summarizes the segmented areas of each PV system, the corresponding capacities estimated using a static power density of 80 MWp/km<sup>2</sup>, and the official capacity values reported by operators. The table also includes the relative error, illustrating how closely the estimated capacities align with official data. Adrar emerges as the largest site, covering 28.41 ha and yielding an estimated capacity of 22.81 MWp,



**Figure 4.4:** Satellite imagery with segmentation results for the PV systems in the Piat region. Red outlines indicate the predicted PV areas. From left to right and top to bottom: Timimoune, Adrar, Z Kounta, Kabertene, Aoulef, and In Salah.

**Table 4.1:** Comparison of segmented areas (in hectares), calculated capacities using a factor of 80 MWp/km<sup>2</sup>, official capacities, and relative errors in percent for the PV systems.

| PV System | Segm. Area (ha) | Capa. Calc. (MWp) | Capa. Off. (MWp) | Rel. Error (%) |
|-----------|-----------------|-------------------|------------------|----------------|
| Z Kounta  | 8.13            | 6.53              | 6                | 8.81           |
| Reggane   | 6.35            | 5.10              | 5                | 2.02           |
| In Salah  | 6.49            | 5.21              | 5                | 4.19           |
| Adrar     | 28.41           | 22.81             | 20               | 14.05          |
| Timimoune | 12.06           | 9.69              | 9                | 7.66           |
| Kabertene | 3.89            | 3.12              | 3                | 4.00           |
| Aoulef    | 6.53            | 5.24              | 5                | 4.83           |

whereas Kabertene is the smallest at 3.89 ha (3.12 MWp). Across all seven sites, the relative error ranges from 2.02 % (Reggane) up to 14.05 % (Adrar), with most errors below 10 %. When aggregated, the method estimates a total of 57.51 MWp for the entire Piat region—approximately 8.5 % higher than the official combined capacity of 53 MWp. This discrepancy underscores the need for locally refined power densities to achieve more accurate estimates.

## 4.6 Discussion

The results confirm that the proposed method is capable of accurately estimating the capacities of PV systems. However, several observations show significant differences.

The resolution analysis revealed that input resolutions of 10 to and 30 cm primarily captured small-scale features at the panel level, whereas resolutions of 80 cm and above focused on the large-scale structural layout of the PV systems. This correlation is certainly highly dependent on the training sample used. The estimated total installed capacity of 57.51 MWp exceeds the official total capacity of 53 MWp by 8.5 %. This overestimation likely reflects the generic nature of the static power density factor of 80 MWp/km<sup>2</sup>, derived from an average of installations in the US for 2016. Some PV systems in the Piat region were installed in 2015, when average power densities were lower due to less efficient technologies, highlighting the importance of applying year-specific power density factors.

Regional variations in factors such as land coverage, module efficiency, and site-specific design are also not fully taken into account. While the factor provides a practical baseline, local adjustments are essential to improve the accuracy of capacity estimates in regions with unique environmental conditions, such as the Piat region. The relative error varied by system size, with smaller systems like Kabertene showing lower deviations (4.00 %) and larger systems like Adrar showing higher deviations (14.05 %). This suggests that inaccuracies in segmentation and simplifications in capacity conversion have a proportionally greater impact on larger systems. Incorporating system-specific parameters such as the ground coverage ratio could further refine capacity estimates.

The lack of training data from African regions, including Algeria, may have limited the model's ability to adapt to local conditions. Increasing the geographic diversity of the training data could significantly improve the model's performance in underrepresented regions. To address the overestimation, we propose a localized correction factor of 73.5 MWp/km<sup>2</sup>, which was derived by incorporating the mean relative error of 8.5 %. This adjusted fac-

tor better reflects the specific conditions of the Piat region and could reduce the discrepancies between calculated and official capacities. The application of such regional adjustments could improve the applicability of the model in other areas with similar characteristics.

In summary, the results demonstrate the robustness of the method and its potential for analyzing renewable energy at the global level. However, addressing the identified limitations is crucial to improving its accuracy and applicability in different regional contexts.

## 4.7 Conclusions

This study demonstrates a scalable and user-friendly approach for the segmentation and capacity estimation of PV systems using deep learning integrated into the QGIS environment. The proposed methodology achieves high accuracy with deviations around 8.5 %, despite the absence of local training data. Integrating the Deepness plugin ensures accessibility for non-technical users, enabling global applications. A local correction reduces the generic 80 MWp/km<sup>2</sup> power density to 73.5 MWp/km<sup>2</sup>, based on the mean relative error in the Piat region. Note that “power density” here denotes installed capacity per unit area (MWp/km<sup>2</sup>), not to be confused with “capacity factor,” the ratio of actual energy output to its theoretical maximum. Future work will expand regional datasets and refine area-to-capacity conversions. Overall, the approach holds strong potential for advancing renewable energy analysis and planning worldwide.

## Bibliography

- [1] Maximilian Kleebauer. Piat Solar Photovoltaic Systems, Algeria. Zenodo, <https://doi.org/10.5281/zenodo.15294535>, 2025. Data set.
- [2] Maximilian Kleebauer. Piat Solar Photovoltaic Systems, Algeria (Map). Google My Maps, [https://www.google.com/maps/d/edit?mid=1mJtANHKUGB81otvNTuTCEah\\_QbvybtU](https://www.google.com/maps/d/edit?mid=1mJtANHKUGB81otvNTuTCEah_QbvybtU), 2025. Accessed: 2025-04-28.
- [3] International Renewable Energy Agency (IRENA). Global renewables outlook: Energy transformation 2050, April 2020. Accessed: 2024-12-09.
- [4] Jordan M Malof, Kyle Bradbury, Leslie M Collins, and Richard G Newell. Automatic detection of solar photovoltaic arrays in high resolution aerial imagery. *Applied energy*, 183:229–240, 2016.
- [5] Maximilian Kleebauer, Daniel Horst, and Christoph Reudenbach. Semi-automatic generation of training samples for detecting renewable energy plants in high-resolution aerial images. *Remote Sensing*, 13(23):4793, 2021.
- [6] Jiafan Yu, Zhecheng Wang, Arun Majumdar, and Ram Rajagopal. Deep-solar: A machine learning framework to efficiently construct a solar deployment database in the united states. *Joule*, 2(12):2605–2617, 2018.
- [7] Maximilian Kleebauer, Christopher Marz, Christoph Reudenbach, and Martin Braun. Multi-resolution segmentation of solar photovoltaic systems using deep learning. *Remote Sensing*, 15:5687, 2023.
- [8] Nicolene Botha, Toshka Coleman, Gert Wessels, Maximilian Kleebauer, and Stefan Karamanski. Power generation time series for solar energy generation: Modelling with atlite in south africa. In *Solar*, volume 5, page 8. MDPI, 2025.
- [9] Juha Kiviluoma, Arttu Tupala, Antti Soinen, and International Renewable Energy Agency (IRENA). Irena flextool. *Computer software*.

Available online: <https://github.com/irena-flextool/flextool>,  
Accessed: 2024-12-02.

- [10] Anni Niemi, Salim Bouchakour, Bendaas Ismail, Kada Bouchouicha, Abdelhak Razagui, Nelli Putkonen, and Juha Kiviluoma. The curious case of wind power in the desert. *IET Conference Proceedings*, 2024:536–541, 02 2025.
- [11] F. Sahnoune, M. Belhamel, M. Zelmat, and R. Kerbachi. Climate change in algeria: Vulnerability and strategy of mitigation and adaptation. *Energy Procedia*, 36:1286–1294, 2013. TerraGreen 13 International Conference 2013 - Advancements in Renewable Energy and Clean Environment.
- [12] Mohammed Guezgouz, Jakub Jurasz, Mohamed Chouai, Hannah Bloomfield, and Benaissa Bekkouche. Assessment of solar and wind energy complementarity in algeria. *Energy Conversion and Management*, 238:114170, 2021.
- [13] Ahmed Bouraiou et al. Status of renewable energy potential and utilization in algeria. *Journal of Cleaner Production*, 246:119011, 2020.
- [14] Mohamed Teggat, Ali Elbar, Abdelghani Laouer, Aissa Atia, Ameer Mechraoui, Saad Mekhilef, Kamal A. R. Ismail, El Hacene Mezaache, Mohamed Souici, and Fatima A. M. Lino. Challenges and prospects of concentrated solar power deployment in algeria. *European Journal of Sustainable Development Research*, 8(4), 2024.
- [15] A. Boudghene Stambouli, Z. Khiat, S. Flazi, and Y. Kitamura. A review on the renewable energy development in algeria: Current perspective, energy scenario and sustainability issues. *Renewable and Sustainable Energy Reviews*, 16(7):4445–4460, 2012.
- [16] Ali Abderrazak Tadjedine et al. Vre integrating in piat grid with afr using pss, mppt, and pso-based techniques: A case study kabertene. *EAI Endorsed Transactions on Energy Web*, 10, 2023.

- 
- [17] Mohamed Ahmedbelbachir. Renewable energies, transition and prospects: The case of algeria. *The Eurasia Proceedings of Educational and Social Sciences*, 32:62–70, Nov. 2023.
- [18] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*, 2017.
- [19] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [20] Przemysław Aszkowski, Bartosz Ptak, Marek Kraft, Dominik Pieczyński, and Paweł Drapikowski. Deepness: Deep neural remote sensing plugin for qgis. *SoftwareX*, 23:101495, 2023.
- [21] Google. Google satellite imagery. <https://mt1.google.com/vt/lyrs=s&x={x}&y={y}&z={z}>, 2024. Accessed: 2024-11-25.
- [22] Mark Bolinger and Greta Bolinger. Land requirements for utility-scale pv: An empirical update on power and energy density. *IEEE Journal of Photovoltaics*, 12(2):589–594, 2022.

## Chapter 5

# Enhancing Wind Turbine Location Accuracy: A Deep Learning-Based Object Regression Approach for Validating Wind Turbine Geo-Coordinates

## Bibliographic Information

Kleebauer, M., Braun, A., Horst, D., Pape, C. (2024). Enhancing Wind Turbine Location Accuracy: A Deep Learning-Based Object Regression Approach for Validating Wind Turbine Geo-Coordinates. IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium, Athens, Greece, 2024, pp. 7863-7867.

## Author's contribution

In the following paragraph, the individual contributions of the authors are briefly broken down with regard to the publication. The authors are abbreviated as follows: Maximilian Kleebauer (M.K.), Axel Braun (A.B.), Daniel Horst (D.H.), and Carsten Pape (C.P.). Conceptualization, M.K.; methodology, M.K.; software, M.K.; validation, M.K., C.P.; formal analysis, M.K.; investigation, M.K.; resources, D.H., C.P.; curation, M.K.; writing—original draft preparation, M.K., A.B.; writing—review and editing, M.K.; visualization, M.K.; supervision, M.K., C.P.; project administration, C.P.; funding acquisition, M.K. All authors have read and agreed to the published version of the manuscript.

## Copyright Notice

*Author accepted manuscript (AAM)*. This is the author's version of the work accepted for presentation/publication at the 2024 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2024). © 2024 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The Version of Record is available at IEEE Xplore:  
<https://doi.org/10.1109/IGARSS53475.2024.10641018>.

## 5.1 Abstract

Remote sensing and deep learning-based methods can be combined to obtain location information automatically on a large scale. This paper introduces an approach for enhancing the geo-coordinate accuracy of existing wind turbines. By employing a RetinaNet-based method for regressive object localization, turbines can be precisely located in images in addition to being identified. Utilizing semi-automatically processed and manually filtered high-resolution image data, a model is trained with an average precision of 96 %. Subsequently, the model is applied to Germany's MaStR wind turbine dataset. The application illustrates the advantageous implementation of the method and emphasizes its considerable potential for improving the accuracy of geo-coordinates. While 73.72 % of existing coordinates can be confirmed as correct with a deviation of less than 10 meter, for more than 15 % of the turbine locations coordinates between 10 and 100 meters can be corrected, and for 5.6 % locations a deviation of more than 100 meter can be determined. This showcases the real-world application of the proposed methodology and underscores its significant potential for enhancing the quality of geo-coordinates.

**Keywords** wind turbines; renewable energy systems; object regression; geo-coordinate validation

## 5.2 Introduction

The future of energy supply faces a pivotal challenge, necessitating a substantial expansion of renewable energy sources. This shift towards renewable energy often occurs at the local level, characterized by a strong decentralization trend. This development is evident in the spatial fragmentation of renewable energy expansion. Accurate representation of existing producers and consumers is crucial, particularly in grid operations but also for forecasting wind potentials. Effective planning, considering priority areas and land availability for further wind turbine installations, becomes feasible through

this detailed depiction. Ensuring the reliability of wind turbine facilities requires precise predictions.

Looking at datasets of existing renewable energy systems on a global scale, different approaches to providing such datasets can be found. For example, datasets derived from aggregated and digitally processed information on the locations and performance of wind and solar farms have already been published by Dunnet et al. [1]. In contrast, Zhang et al. [2] created a global dataset of offshore wind turbines using Sentinel-1 Synthetic Aperture Radar time series images. Hoerer et al. [3] also used Sentinel-1 Synthetic Aperture Radar data for the automatic derivation of offshore wind turbine locations. Other published studies perform segmentation of onshore wind turbines in high-resolution aerial imagery [4, 5] and wind turbine detection using Sentinel-2 RGB imagery [6]. An improvement in detection accuracy was achieved through multiple acquisition times with Sentinel-2, as shown in a multimodal approach [7]. According to our research, it has not yet been achieved to derive a precise data set of existing energy facilities such as wind turbines, especially onshore, on a large scale based solely on remote sensing data and machine learning methods.

In Germany, the Core Energy Market Data Register provides an overview of existing energy facilities. Location information, performance values or specific plant characteristics are provided for wind turbines, solar photovoltaic systems, and other plant types [8]. However, certain coordinates given here are very inaccurate. Studies have consequently made several revisions to improve data quality [9, 10].

Our approach aims to enhance the accuracy of existing wind turbine location data through a deep learning-based object detection approach using aerial images. There are a variety of DL techniques and methods for object detection, starting with early approaches of the Single Shot MultiBox Detector [11], a variety of real-time object detection variants based on YOLO [12], or RetinaNet [13], which combines the ResNet [14] with an FPN [15]. After

studies have already shown high accuracy [16, 17], RetinaNet will be trained with semi-automatically processed and manually filtered high-resolution imagery and then applied to all wind turbines in Germany. This case study is designed to demonstrate the potential for improving the location accuracy of wind turbines.

### 5.3 Materials

The Core Energy Market Data Register (German: Marktstammdatenregister, MaStR), administered by the Federal Network Agency for the German electricity and gas market, serves as a comprehensive database for energy market. Commencing operations in 2019, the register undergoes daily updates and encompasses detailed information on entities and facilities within the grid-bound energy supply market. Alongside registered electricity generation units, the register also includes extensive listings of large-scale consumers. On the producer side, the MaStR provides location information, performance values, and specific plant characteristics for various energy sources, such as wind turbines, solar photovoltaic systems, biomass plants, hydro power plants, and conventional plants. The register offers a wealth of information specifically tailored to wind turbines like registration date, commissioning date, rated power, remote controllability, current operating status, manufacturer, type designation, hub height, and rotor diameter. Publicly accessible address data is available at the zip code level, while detailed coordinates are provided for the majority of the turbines. As of November 21, 2023, out of the 32,788 listed turbines in operation, 31,892 include coordinate information [8].

The Digital Orthophotos (DOP) of Germany, as documented by the Federal Agency for Cartography and Geodesy [18], constitute georeferenced and differentially rectified aerial imagery, sourced from the surveying administrations of Germany's federal states. These images faithfully represent the Earth's surface, within the confines of the Federal Republic of Germany, employing a ground resolution of 0.2 meter for the purposes of this investigation.

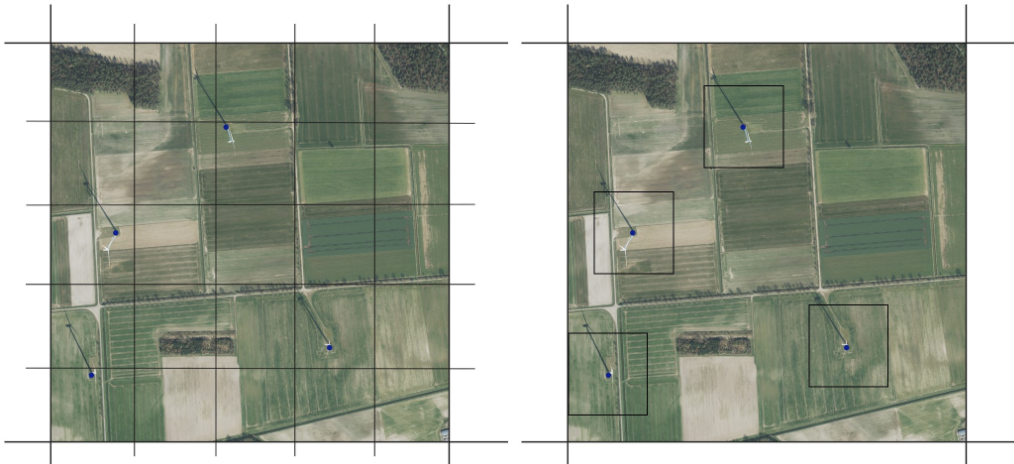
The dataset encompasses both color images in the RGB spectrum. The images maintain a positional accuracy of  $\pm 0.4$  meter standard deviation. The entire dataset covers Germany and is presented in tiles measuring  $1,000 \times 1,000$  meters, equivalent to  $5,000 \times 5,000$  pixels each. Each tile is accompanied by a file containing meta-information, notably the timestamp indicating when the respective image was captured. The image data is updated cyclically, usually available every 3 years in site-specific overflight intervals.

## 5.4 Methods

### 5.4.1 Data preprocessing

The preparation of training data comprises multiple steps. Initially, all wind turbines registered in the MaStR are loaded. A pre-filtering process is then applied, focusing exclusively on wind turbines with operational status labeled as "in operation". Additionally, the turbines must be categorized as "onshore", and only systems with available geo-coordinates are considered. Finally, turbines situated outside the German federal border are excluded under the assumption that their location data is inaccurate. Existing coordinates of the turbines are provided with a static buffer of 30 meter radius in order to obtain an area-like imprint of the point coordinates. These are required for the subsequent regressive localization method. To generate training and application image data, the wind turbine location data is integrated with DOPs. To align with the requirements of RetinaNet, DOP tiles, each measuring  $5,000 \times 5,000$  pixels, are further divided into  $1,000 \times 1,000$  pixel tiles. A methodical differentiation is applied to training and application images. For training, the cut edges are statically selected to generate 15 tiles from each original tile, as illustrated on the left side in Figure 5.1.

This approach ensures that wind turbines are not consistently positioned at the center of the image sections. Conversely, for application images, the wind turbine location is designated as the centroid of the image, as depicted on the right side in Figure 5.1. Approximately 12,000 images are produced, each containing at least one wind turbine.



**Figure 5.1:** The two methods of processing the DOPs are shown in comparison. The left-hand side shows an example of static cutting of the training images, while the right-hand side shows cutting using the coordinates of the wind turbines as a centroid. The black lines represent the cutting edges, the blue dots the coordinates of the wind turbines.

This dataset is utilized for the initial training. To ensure the utilization of highly suitable image data in the subsequent, second training, samples generated automatically undergo manual inspection. This process identifies and removes instances with incorrect coordinates stored in MaStR, imprecise coordinates, and image scenes with insufficient resolution. The re-selection leads to the reduction of a further 5,000 unsuitable images, yielding a dataset of 7,000 images for the second training. The primary emphasis is on the precise localization of wind turbines, ensuring that the center of the regression boxes accurately represents the tower's exact ground location. As illustrated in Figure 5.2, several samples are depicted to exemplify their suitability. The training is divided into two parts. First, all 12,000 samples automatically derived from the data preprocessing are used, whereas in the second training, the number of samples is reduced to 7,000 highly suitable samples by manual filtering. All other parameters remained the same for both the first and second training: 100 epochs, 100 steps, 80 % training and 10 % independent validation and test data set each. Further parameter configurations are the default settings defined in the package [19].



**Figure 5.2:** Samples based on their suitability for training. The images marked in red are unsuitable due to incorrect position or poor image resolution, the images marked in yellow contain wind turbines that are clearly visible but were rejected for fine-tuning due to their inaccurate position. The images marked in green contain turbines whose tower base is located directly in the center of the respective boxes.

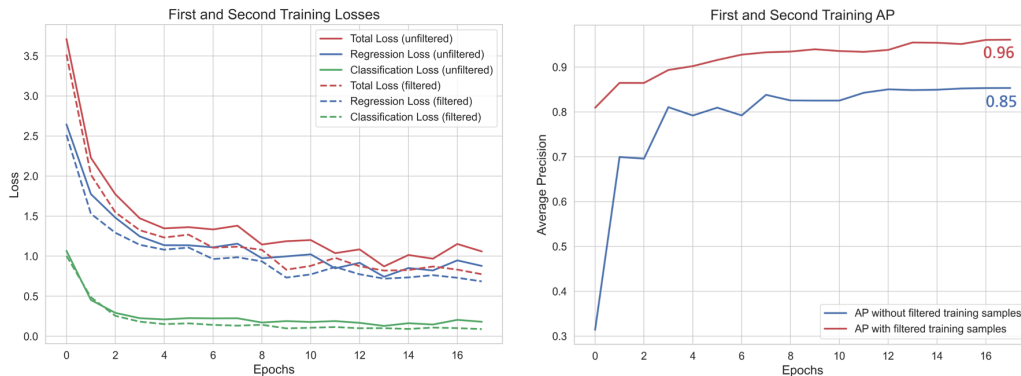
### 5.4.2 Deep learning approach

The machine learning method used in the work is called RetinaNet. The network combines several common variants of deep learning into a classification and regression network. One key feature is the ResNet architecture, including Convolutional Neural Network (CNN) [13]. Furthermore, adding skip connections between the layers enables residual learning as a widely used basis for deep learning [14]. A Feature Pyramid Net is also used. It consists of a top-down architecture with horizontal connections to create high-level semantic feature maps at all scales [15]. The classification task is performed at the output of the backbone network using Focal Loss, which was developed to train extremely unevenly distributed foreground and background classes [13]. The regression task, on the other hand, is implemented for the regressive delineation of the objects. The regression targets are output as rectangles that are located entirely within the images shown and have predefined aspect ratios of 1:2, 1:1, 2:1 [13]. The regression loss uses the Smooth L1 Loss, which was originally developed as part of the Fast R-CNN [20]. Average Precision (AP) is used to measure the performance of the model. The sum of correctly detected objects in relation to all detected objects is considered. RetinaNet uses the COCO detection evaluation metric under the AP,

whereby the overlap of the areas from the regression and the ground truths must be greater than 50 %, in order to be considered a correctly detected object. The keras-retinanet [19] package is used, which was specially developed for the use of RetinaNet. To quickly achieve good generalization of the network, pre-trained weights on 500 classes from the Open Image Dataset where used [21].

## 5.5 Results

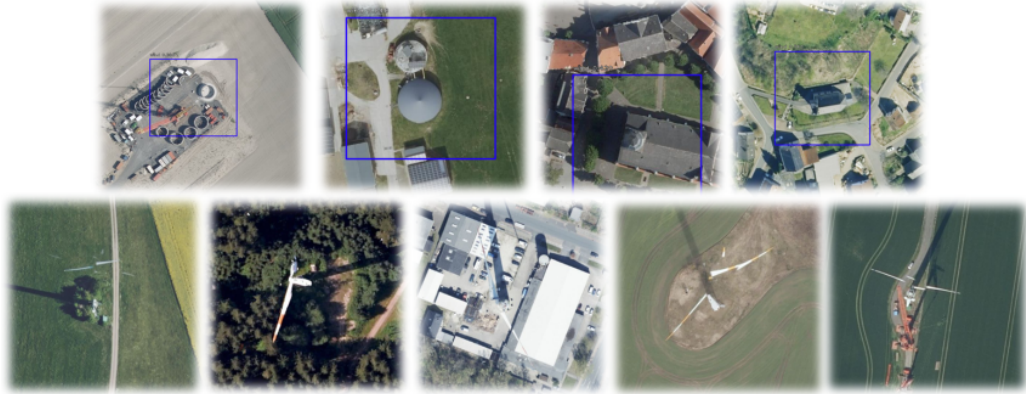
As Training Progress Summary, the progression of the two losses from the classification and regression networks, as well as the AP, were validated to determine the networks' performance. Figure 5.3 shows the losses on the left-hand and the AP on the right-hand side.



**Figure 5.3:** The illustrations depict the losses and the AP.

The training is terminated by early stopping after 17 epochs in each case, indicating no further progress in training. The trend in losses exhibits a nearly constant decrease for both the initial and subsequent training. This is observed for both the focal loss in classification and the Smooth L1 loss in regression. The total loss represents the cumulative sum of the individual losses. A consistent upward trend can be observed in the AP. Finally, the AP is 85 % for the first training and 96 % for the second training with post-filtered samples. Overall, the loss and AP's curves clearly show the strong generalization of the network based on the training examples. Incorporating

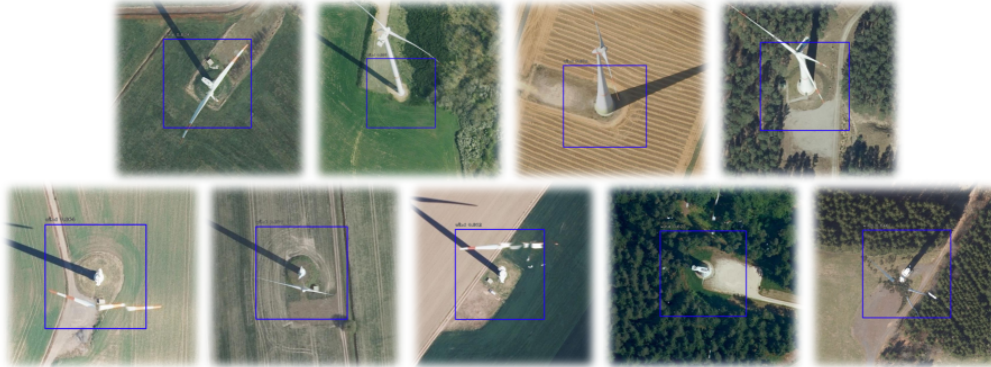
rect recognition are shown in Figure 5.4.



**Figure 5.4:** False positive and false negative examples from the application are summarized in the following. The top line represents incorrectly identified wind turbine, false positives. The bottom line shows turbines that have not been detected.

This includes a construction site, a biogas plant and two churches. Secondly, some of the poorly represented turbines are not recognized by the network. This applies to different backgrounds, so that turbines in open fields, in the forest and also in the settlement are not recognized. However, they are also hard to identify during a visual inspection. Examples of correctly recognized wind turbines, conversely, are shown in Figure 5.5. In addition to turbines with good resolution, poorly resolved turbines can also be identified in the images. All images show that the regression locates the towers of the turbines exactly in the centers of the bounding boxes. In other words, the centers of the regression boxes can be interpreted as exact geo-coordinates of the wind turbines.

The results of the geo-location site correction for onshore wind turbines are described below. The dataset comprised a total of 30,326 operational onshore wind turbines. Through geographical selection, it was revealed that 55 wind turbine (0.18 %) coordinates were situated outside the country's borders. Additionally, 1,397 turbines (4.61 %) were flagged for having outdated aerial images captured before their commissioning. The validation methodology was applied to identify 1,699 instances (5.60 %) of incorrect lo-



**Figure 5.5:** True positive examples from the application are presented as follows. The upper row displays instances featuring clearly visible and accurately identified wind turbines, where the centroid of the regression boxes serves as the base of the tower. The bottom row, shows correctly detected wind turbines, with less clear representation in the images.

cation data, where a deviation of at least 100 meter from the coordinates in the MaStR was observed. Consequently, 27,130 wind turbine location data points (89.46 %) were identified within 100 meter accuracy and are treated with sensitivity in subsequent analyses due to their deviation characteristics. Table 5.1 visually depicts the dispersion of wind turbines across various distance ranges, offering crucial insights into the efficacy of the applied coordinate corrections. The categorization of turbines based on deviation ranges in meter provides a detailed view. Notably, a substantial percentage (73.72 %) of turbines exhibits a deviation of less than 10 meter, underscoring the high precision of their coordinates. In contrast, 7.66 % of turbines deviate between 10 and 20 meter, while more than 8 % show deviations exceeding 20 meter.

## 5.6 Discussion

Comparing the results of the first and second training reveals a significant increase in recognition accuracy due to the manual post-filtering of the training samples. The high accuracy of the second model suggests a strong generalization of the network based on the validation data. This can also be illustrated by Figure 5.5. It should be noted, however, that the data used in the val-

**Table 5.1:** Distances between MaStR and validated coordinates.

| Range in meter | wind turbines | wind turbines in % |
|----------------|---------------|--------------------|
| < 10           | 22,357        | 73.72              |
| 10 - 20        | 2,323         | 7.66               |
| 20 - 30        | 872           | 2.88               |
| 30 - 40        | 514           | 1.70               |
| 40 - 50        | 319           | 1.05               |
| > 50           | 745           | 2.46               |

idation is very homogeneous due to the experimental setup. A large-scale application would result in a reduction with regard to the AP. Nevertheless, the misidentified objects observed in the test application, as shown in Figure 5.4, provide insights into two important considerations. Several objects are incorrectly recognized as wind turbines. Facing this issue, an improvement in recognition accuracy could be achieved by including misidentified objects as part of the negative class during training. However, the unrecognized wind turbines also illustrate the limitations of the application based on DOPs. Moreover, Figure 5.4 demonstrates that turbines that lack consistent representation in the images may go unnoticed by the network during the application. Supplementing the image data with additional acquisition times could address this issue.

The application shows a strong variance in the accuracy of the geo-coordinates. The locations of almost three quarters of the wind turbines can be confirmed. However, nearly 8 % of the locations deviate between 10 and 20 meter, and more than 8 % deviate by more than 20 meter. With 5.6 % of the completely incorrectly stored coordinate locations, a considerable proportion of the wind turbines are incorrectly located by more than 100 meter. Due to outdated aerial images, the 4.61 % of turbines with outdated aerial images could not be validated. Expanding the methodology to regularly updated image datasets, such as the utilization of Sentinel-2 imagery, is conceivable. Nevertheless, the adoption of Sentinel-2 imagery, characterized by considerably lower resolution, may result in a substantial decline in accuracy during

facility identification.

## 5.7 Conclusions

This paper introduces an approach for enhancing the geo-coordinate accuracy of existing wind turbines. Employing a RetinaNet-based approach for regressive object localization, wind turbines can be precisely located in images in addition to being identified. The validation results demonstrate the network's high generalization capability. The exemplified application, focusing on enhancing wind turbine location accuracy through MaStR data, illustrates the practical implementation of the method and emphasizes its considerable potential for improving the accuracy of geo-coordinates.

## Bibliography

- [1] Sebastian Dunnett, Alessandro Sorichetta, Gail Taylor, and Felix Eigenbrod. Harmonised global datasets of wind and solar farm locations and power. *Scientific Data*, 7(1):1–12, 2020.
- [2] Ting Zhang, Bo Tian, Dhritiraj Sengupta, Lei Zhang, and Yali Si. Global offshore wind turbine dataset. *Scientific Data*, 8(1):191, 2021.
- [3] Thorsten Hoerer, Stefanie Feuerstein, and Claudia Kuenzer. Deepowt: A global offshore wind turbine data set derived with deep learning from sentinel-1 data. *Earth System Science Data*, 14(9):4251–4270, 2022.
- [4] Min Han, Huabin Wang, Guanghui Wang, and Yu Liu. Targets mask u-net for wind turbines detection in remote sensing images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42:475–480, 2018.
- [5] Narayana Darapaneni, A Jagannathan, Vigneshwaran Natarajan, Guruprasadh Vadakkupattu Swaminathan, S Subramanian, and Anwesh Reddy Paduri. Semantic segmentation of solar pv panels and wind turbines in satellite images using u-net. In *2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS)*, pages 7–12. IEEE, 2020.
- [6] Michael Mommert, Linus Scheibenreif, Joëlle Hanna, and Damian Borth. Power plant classification from remote imaging with deep learning. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pages 6391–6394. IEEE, 2021.
- [7] N Mandroux, S Drouyer, and R Grompone von Gioi. Multi-date wind turbine detection on optical satellite images. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2:383–390, 2022.

- 
- [8] Federal Network Agency (BNetzA). Core energy market data register (mastr). <https://www.marktstammdatenregister.de/MaStR/>. Accessed on 02 November 2023.
- [9] David Manske, Lukas Grosch, Julius Schmiedt, Nora Mittelstädt, and Daniela Thrän. Geo-locations and system data of renewable energy installations in germany. *Data*, 7(9):128, 2022.
- [10] Deniz Tepe, Florian Kotthoff, Christoph Muschner, Esther Vogt, and Ludwig Hülk. Improving data reliability in germany’s energy system: A validation of unit locations of the marktstammdatenregister. *arXiv preprint arXiv:2304.10581*, 2023.
- [11] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. Ssd: Single shot multibox detector. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*, pages 21–37. Springer, 2016.
- [12] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.
- [13] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollar. Focal loss for dense object detection. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2999–3007, 22.10.2017 - 29.10.2017.
- [14] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [15] Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 936–944. IEEE, 21.07.2017 - 26.07.2017.

- 
- [16] Maximilian Kleebauer, Daniel Horst, and Christoph Reudenbach. Semi-automatic generation of training samples for detecting renewable energy plants in high-resolution aerial images. *Remote Sensing*, 13(23):4793, 2021.
- [17] Wenwen Qi. Object detection in high resolution optical image based on deep learning technique. *Natural Hazards Research*, 2(4):384–392, 2022.
- [18] Bundesamt für Kartographie und Geodäsie. Dokumentation digitale orthophotos. 2023.
- [19] Hans Gaiser, Maarten de Vries, Valeriu Lacatusu, vcarpani, Ashley Williamson, Enrico Liscio, Andrés, Yann Henon, jjiun, Cristian Gratie, Mihai Morariu, Charles Ye, Martin Zlocha, Ben Weinstein, Rodrigo Meira de Andrade, Pedro Conceição, Alexander Pacha, hannedevartsen, Daniyal Shahrokhian, Wudi Fang, Mike Clark, meagerYak, Iver Jordal, Max Van Sande, Jin, Etienne-Meunier, Andrew Grigorev, Guillaume Erhard, Eduardo Ramos, and Denis Dowling. *fizyr/keras-retinanet 0.5.1*. 2019.
- [20] Ross Girshick. Fast r-cnn. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 1440–1448. IEEE, 07.12.2015 - 13.12.2015.
- [21] GitHub. Github - zfturbo/keras-retinanet-for-open-images-challenge-2018: Code for 15th place in kaggle google ai open images - object detection track, 2021.

## Chapter 6

# A Wind Turbines Dataset for South Africa: OpenStreetMap Data, Deep Learning Based Geo-Coordinate Correction and Capacity Analysis

## Bibliographic Information

Kleebauer, M., Karamanski, S., Callies, D., Braun, M. (2025). A Wind Turbines Dataset for South Africa: OpenStreetMap Data, Deep Learning Based Geo-Coordinate Correction and Capacity Analysis. *ISPRS International Journal of Geo-Information*, 14(6), 232.

## Author's contribution

In the following paragraph, the individual contributions of the authors are briefly broken down with regard to the publication. The authors are abbreviated as follows: Maximilian Kleebauer (M.K.), Stefan Karamanski (S.K.), Doron Callies (D.C.), and Martin Braun (M.B.). Conceptualization, M.K., S.K.; methodology, M.K.; software, M.K.; validation, M.K., S.K.; formal analysis, M.K.; investigation, M.K.; resources, M.K.; curation, M.K., S.K.; writing—original draft preparation, M.K.; writing—review and editing, M.K., M.B., D.C., S.K.; visualization, M.K.; supervision, M.K.; project administration, M.K.; funding acquisition, M.K., D.C. All authors have read and agreed to the published version of the manuscript.

## Copyright Notice

© 2025 by the authors. Published by MDPI on behalf of the International Society for Photogrammetry and Remote Sensing. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>). This is the accepted manuscript version of the article published (post-peer-review) in *ISPRS International Journal of Geo-Information*, MDPI. For the final published version, please see:

<https://doi.org/10.3390/ijgi14060232>.

## 6.1 Abstract

Accurate and detailed spatial data on wind energy infrastructure is essential for renewable energy planning, grid integration, and system analysis. However, publicly available datasets often suffer from limited spatial accuracy, missing attributes, and inconsistent metadata. To address these challenges, this study presents a harmonized and spatially refined dataset of wind turbines in South Africa, combining OpenStreetMap (OSM) data with high-resolution satellite imagery, deep learning-based coordinate correction, and manual curation. The dataset includes 1,487 turbines across 42 wind farms, representing over 3.9 GW of installed capacity as of 2025. Of this, more than 3.6 GW is currently operational. The Geo-Coordinates were validated and corrected using a RetinaNet-based object detection model applied to both Google and Bing satellite imagery. Instead of relying solely on spatial precision, the curation process emphasized attribute completeness and consistency. Through systematic verification and cross-referencing with multiple public sources, the final dataset achieves a high level of attribute completeness and internal consistency across all turbines, including turbine type, rated capacity, and commissioning year. The resulting dataset is the most accurate and comprehensive publicly available dataset on wind turbines in South Africa to date. It provides a robust foundation for spatial analysis, energy modeling, and policy assessment related to wind energy development. The dataset is publicly available.

**Keywords** Wind turbine location; Renewable energy; Deep learning; Geo-coordinate correction; OpenStreetMap

## 6.2 Introduction

Wind energy is one of the fastest-growing renewable energy sources worldwide. In 2023, wind energy recorded its highest ever growth: in a single year, more than 100 GW of new onshore capacity and over 11 GW of offshore wind capacity were added globally. Total installed capacity worldwide

exceeded the symbolic milestone of 1 TW for the first time and is expected to reach 2 TW before the end of this decade if current growth trends continue [1]. In addition, the International Energy Agency (IEA) forecasts scenarios in which wind energy could meet more than 20 % of global electricity demand by 2030, provided that ambitious climate protection measures are implemented [2]. The transition to renewable energy sources presents major challenges. Accurate mapping and monitoring of wind turbine locations and meta-information on the turbine characteristics (e. g. turbine types, nominal power, hub height, or rotor diameter) are critical for effective integration into electricity grids and sustainable infrastructure planning.

Despite its growing global importance, detailed and spatially accurate datasets of wind turbine infrastructure remain scarce in many regions of the world. Existing global datasets often focus on aggregated capacities or rough location data, lacking precision for localized planning and operational decision-making. Recent research efforts address these limitations through advanced remote sensing and machine learning approaches. For instance, global offshore wind turbine locations were mapped using Sentinel-1 radar images [3, 4], while segmentation methods utilizing high-resolution aerial images [5, 6] and Sentinel-2 RGB imagery [7, 8] improved the detection accuracy of onshore wind turbines. Moreover, the integration of multimodal data sources [9, 10], further enhances detection accuracy and completeness. Even approaches to enable global detection are being researched [11, 12].

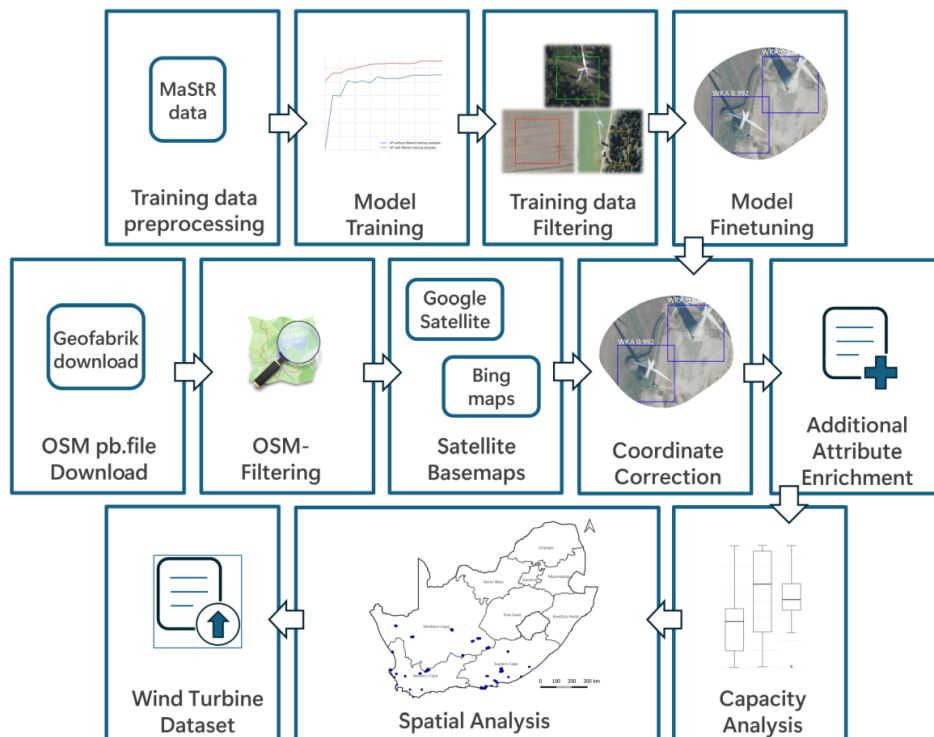
In the specific context of South Africa, the national Renewable Energy Independent Power Producer Procurement Programme (REIPPPP) plays a central role in realizing the country’s long-term energy infrastructure goals. Launched in 2011 by the Department of Mineral Resources and Energy in cooperation with National Treasury and the Development Bank of Southern Africa, the REIPPPP was designed to facilitate private sector investment into grid-connected renewable energy generation through competitive bidding. The programme has since led to the procurement of more than 6.3 GW of renewable capacity, including wind, solar photovoltaic (PV), and other sources [13]. Recent regulatory changes have further expanded the landscape of wind energy development in South Africa. In particular, the

lifting of the 100 MW licensing cap for private generation in January 2023 has enabled the construction of wind farms outside the REIPPPP framework [14]. As part of this programme, the Independent Power Producers (IPP) Projects Database is maintained by the IPP Office and provides a structured overview of utility-scale renewable energy projects, including wind farms. The database includes information such as project names, capacities, and commissioning dates. However, it does not contain detailed geospatial information on individual turbines and typically excludes smaller or non-utility-scale developments. At the same time, it does not provide any technical information such as turbine types, hub height, or rotor diameter [15]. This study aims to fill this data gap and provide a spatially refined and attribute-based dataset that captures the full extent of wind turbine infrastructure in the country. This includes both large utility-scale farms and smaller, decentralized installations, enabling more comprehensive and accurate energy system analyses.

To overcome these data limitations, this article builds upon the methodologies initially presented in the conference paper entitled *Enhancing Wind Turbine Location Accuracy: A Deep Learning-Based Object Regression Approach for Validating Wind Turbine Geo-Coordinates* [16]. Here, the original methods are further developed, combining OSM data, DL-based object detection with RetinaNet, high-resolution satellite imagery from Google and Bing, and manual attribute enrichment, to produce a comprehensive, spatially precise dataset of wind turbines in South Africa. This multi-step pipeline ensures robust validation and enrichment, significantly enhancing data quality and applicability for detailed infrastructure planning and energy modelling. Structured as following, this study introduces a multi-step data processing pipeline that combines open data sources, deep learning-based geo-coordinate correction, and manual validation. For better readability, the term “coordinate” will be used synonymously with “geo-coordinate” in the following.

As illustrated in Figure 6.1, the construction of the dataset follows a multi-stage workflow. First, training data is prepared using the German Core Energy Market Data Register (MaStR) and high resolution aerial imagery. A RetinaNet-based deep learning model is trained and fine-tuned to detect

turbines based on this reference data. Preparing the South African wind turbine dataset starts with downloading, extracting and filtering the raw wind turbine data from OSM. High-resolution satellite imagery from both Bing Maps and Google Satellite is then integrated to provide visual context for turbine locations. The model is then applied to correct the spatial positions of turbines, improving the coordination accuracy. Subsequently, a manual attribute enrichment step ensures the inclusion of key turbine information such as name, turbine type, turbine capacity and total wind farm capacities. A capacity analysis and a spatial analysis are then carried out for further description and evaluation. This leads to the final high-quality, geo-referenced dataset of wind turbines in South Africa.



**Figure 6.1:** Workflow for creating the wind turbine dataset, including model training, coordinate correction, and attribute enrichment.

In the larger project context, a comprehensive open-source strategy was developed to ensure barrier-free access to tools and data for energy system modeling. This ecosystem promotes transparency and supports the wider

use of open-source solutions for renewable energy planning and analysis. The methodological chain includes renewable energy system detection [17], high-resolution time series generation [18], and energy system modelling with integration into IRENA FlexTool [19].

## 6.3 Materials

### 6.3.1 OpenStreetMap

OSM is a collaborative, open-access mapping platform that provides geospatial data contributed and maintained by a global community of volunteers [20]. Established in 2004, OSM has become a leading source of freely available geographic information, covering diverse features such as roads, buildings, land use, and energy infrastructure. Its community-driven model ensures that the data is continuously updated and enriched, offering a valuable resource for research, spatial planning and analysis. For renewable energy studies, OSM often serves as an initial source for identifying the location and distribution of wind farms and other energy infrastructure. While OSM data provides extensive coverage, it frequently lacks consistent accuracy in spatial precision and technical detail. This is particularly evident for features such as wind turbine coordinates and capacity specifications, which may be missing or imprecise [21, 22]. To address these limitations, this study validates and refines the locational data using complementary methodologies, such as high-resolution satellite imagery analysis and DL-based object detection approaches. To enable a comprehensive and regionally consistent extraction of the energy-related infrastructure, the complete OSM file for South Africa was downloaded from the Geofabrik [23]. Geofabrik offers daily updated and freely accessible OSM extracts for all regions worldwide. The file for South Africa, was used as the base dataset to ensure that no relevant entries were omitted due to incomplete or outdated online queries.

### 6.3.2 Google Satellite Data and Bing Image Data

Google Satellite Data and Bing Image Data provide high-resolution satellite imagery that is widely utilized for spatial analysis across various disciplines. Google Satellite imagery is accessible through platforms such as Google Earth Engine [24] or Google Maps [25], while Bing imagery is accessed through the Bing Maps API [26]. Both offer detailed views of the Earth’s surface, capturing features such as buildings, roads, and natural landscapes. For accurate validation or correction of existing geo-datasets of renewable energy sources, Google Satellite and Bing Image Data can be used as important resources for identifying and validating infrastructure locations, including wind farms and individual turbines. These datasets provide high-resolution imagery with true-color (red-green-blue) channels that provide realistic visual representations of surface features. This color channel information facilitates the identification and distinction of objects and their properties. In addition, the two satellite image datasets complement each other by providing additional perspectives and different dates of acquisition, which is useful for validating results and improving coverage in areas where one source may have better resolution or more recent data. By integrating Google and Bing satellite data with deep learning and remote sensing techniques, the accuracy, scalability, and reliability of spatial data analysis are significantly improved, providing a comprehensive and multiply-verified dataset for renewable energy infrastructure analysis. They were used for the application and coordinate correction of wind turbine locations in South Africa. In contrast, the training of the DL model was based on turbine coordinates from the German Core Energy Market Data Register (German: Marktstammdatenregister, MaStR) and high-resolution Digital Orthophotos (DOP).

### 6.3.3 Core Energy Market Data Register

The MaStR, administered by the Federal Network Agency for the German electricity and gas market, serves as a comprehensive database for energy market. Commencing operations in 2019, the register undergoes daily updates and encompasses detailed information on entities and facilities within

the grid-bound energy supply market [27]. Alongside registered electricity generation units, the register also includes extensive listings of large-scale consumers. On the producer side, the MaStR provides location information, performance values, and specific plant characteristics for various energy sources, such as wind turbines, PV systems, biomass plants, hydro power plants, and conventional plants. The register offers a wealth of information specifically tailored to wind turbines like registration date, commissioning date, rated power, remote controllability, current operating status, manufacturer, type designation, hub height, and rotor diameter. While freely accessible address data is generally available for all facilities at zip code level, for most wind turbines there are even detailed coordinates for determining the location of the individual turbines. As of November 21, 2023, out of the 32,788 listed turbines in operation, 31,892 include coordinate information. Although MaStR provides very comprehensive technical information, it is limited to turbines in Germany. Therefore, a model can only learn visual and structural features that are specific to the German wind energy industry, such as turbine design, landscape features, or image capture conditions. The application of the model in South Africa may be affected by differences in environmental conditions or local visual characteristics, which could impair recognition performance.

### 6.3.4 Digital Orthophotos

The DOP of Germany, as documented by the Federal Agency for Cartography and Geodesy, constitute georeferenced and differentially rectified aerial imagery, sourced from the surveying administrations of Germany's federal states [28]. These images faithfully represent the Earth's surface, within the confines of the Federal Republic of Germany, employing a ground resolution of 0.2 m for the purposes of this investigation. The dataset encompasses both color images in the RGB spectrum. The images maintain a positional accuracy of  $\pm 0.4$  m standard deviation. The entire dataset covers Germany and is presented in tiles measuring  $1,000 \times 1,000$  m, equivalent to  $5,000 \times 5,000$  pixels each. Each tile is accompanied by a file containing meta-

information, notably the timestamp indicating when the respective image was captured. The image data is updated cyclically, usually available every 3 years in site-specific overflight intervals.

## 6.4 Methods

### 6.4.1 Training data preprocessing

The preparation of training data comprises multiple steps. Initially, all wind turbines registered in Germanys MaStR are loaded. A pre-filtering process is then applied, focusing exclusively on wind turbines with operational status labeled as "in operation". Additionally, the turbines must be categorized as "onshore", and only systems with available coordinates are considered. Finally, turbines situated outside the German federal border are excluded under the assumption that their location data is inaccurate. Existing coordinates of the turbines are provided with a static buffer of 30 m radius in order to obtain an area-like imprint of the point coordinates. These are required for the subsequent regressive localization method.

To generate training image data, the wind turbine location data is combined with DOPs. To align with the requirements of RetinaNet, DOP tiles, each measuring  $5,000 \times 5,000$  pixels, are further divided into  $1,000 \times 1,000$  pixel tiles. For training, the cut edges are statically selected to generate 15 tiles from each original tile, as illustrated on the left side in Figure 6.2. This approach ensures that wind turbines are not consistently positioned at the center of the image sections. Conversely, for application images, the wind turbine location is designated as the centroid of the image. Approximately 12,000 images are produced, each containing at least one wind turbine.

This dataset is utilized for the initial training. To ensure that highly suitable image data is used in the subsequent second training session, automatically generated examples are subjected to manual review. This process identifies and removes instances with incorrect coordinates stored in MaStR, imprecise coordinates, and image scenes with insufficient resolution. The re-selection leads to the reduction of a further 5,000 unsuitable images, yielding



**Figure 6.2:** The method of the static cutting of the training images is shown. The black lines represent the cutting edges, the red dots the coordinates of the wind turbines.

a dataset of 7,000 images for the second training. The primary emphasis is on the precise localization of wind turbines, ensuring that the center of the regression boxes accurately represents the tower's exact ground location. As illustrated in Figure 6.3, several samples are depicted to exemplify their suitability. The training is divided into two parts. First, all 12,000 samples automatically derived from the data preprocessing are used, whereas in the second training, the number of samples is reduced to 7,000 highly suitable samples by manual filtering. All other parameters remained the same for both the first and second training: 100 epochs, 100 steps, 80 % training and 10 % independent validation, and 10 % test data set.

### 6.4.2 Deep Learning Approach

Several object detection frameworks were considered for the coordinate correction task, including Faster R-CNN, YOLOv3, and RetinaNet. RetinaNet was selected due to its balance between high detection accuracy and computational efficiency, particularly in scenarios with class imbalance, such as the



**Figure 6.3:** Samples based on their suitability for training. The images marked in red are unsuitable due to incorrect position or poor image resolution, the images marked in yellow contain wind turbines that are clearly visible but were rejected for fine-tuning due to their inaccurate position. The images marked in green contain turbines whose tower base is located directly in the center of the respective boxes.

detection of sparsely distributed wind turbines in large aerial images. The Focal Loss mechanism employed by RetinaNet has proven to significantly improve detection performance for rare objects compared to conventional cross-entropy loss in other architectures [29]. In addition, RetinaNet achieves competitive results in common object detection benchmarks (e.g., COCO dataset), while offering simpler training requirements compared to two-stage detectors like Faster R-CNN. RetinaNet is an object detection model that combines classification and regression within a unified architecture. It integrates several well-established DL techniques to enable high-precision object localization and classification. A key component is the Residual Network (ResNet) architecture, a variant of Convolutional Neural Networks (CNNs), which utilizes skip connections between layers to facilitate residual learning and improve gradient flow in deep networks [29, 30]. To handle multi-scale object detection, a Feature Pyramid Network (FPN) is employed on top of the backbone. The FPN uses a top-down architecture with lateral connections to generate semantically rich feature maps at multiple scales [31]. This allows the network to detect objects of varying sizes effectively. The classification subnetwork is trained using the Focal Loss, which was specifically developed to address the problem of class imbalance between foreground and background objects in dense detection tasks [29]. Unlike the standard cross-

entropy loss, Focal Loss introduces a modulation factor to down-weight easy examples and focus training on hard negatives. The equation 6.1 for Focal Loss is:

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t) \quad (6.1)$$

where  $p_t$  is the model's estimated probability for the true class,  $\alpha_t$  is the weighting factor for class imbalance ( $\alpha = 0.25$ ), and  $\gamma$  is the focusing parameter ( $\gamma = 2.0$ ). This formulation ensures that well-classified examples receive less weight, allowing the model to focus on misclassified or more difficult samples. The regression subnetwork is responsible for predicting bounding boxes around detected objects. It uses the Smooth L1 Loss function, which combines the benefits of L1 and L2 losses and is less sensitive to outliers. This loss was originally introduced in the Fast Region-based Convolutional Network Network (Fast R-CNN) architecture [32]. The regression loss  $L_{loc}$ , as introduced in equation 6.2, is computed for the predicted bounding box tuple  $t^u = (t_x^u, t_y^u, t_w^u, t_h^u)$  and the ground truth box  $v = (v_x, v_y, v_w, v_h)$  as:

$$L_{loc}(t^u, v) = \sum_{i \in \{x, y, w, h\}} smooth_{L_1}(t_i^u - v_i) \quad (6.2)$$

The Smooth L1 function itself is defined by the following equation 6.3:

$$smooth_{L_1}(x) = \begin{cases} 0.5(\sigma x)^2 & \text{if } |x| < \frac{1}{\sigma^2} \\ |x| - \frac{0.5}{\sigma^2} & \text{otherwise} \end{cases} \quad (6.3)$$

The smoothing parameter  $\sigma$  was set to its commonly used default value of 3.0. This loss formulation enables stable training and effective bounding box regression even in the presence of noisy labels. RetinaNet outputs bounding boxes with predefined aspect ratios of 1:2, 1:1, and 2:1 [29].

Model performance is evaluated using the Average Precision (AP) metric, as defined in equation 6.4. Following the Common Objects in Context (COCO) detection benchmark, a prediction is considered correct if its Intersection over Union (IoU) with the ground truth exceeds 50 %. To calculate AP, two basic metrics are first needed, Precision and Recall. Precision mea-

sures the proportion of correctly identified objects (true positives) among all identified objects (true positives and false positives):

$$\text{Precision} = \frac{TP}{TP + FP} \quad (6.4)$$

Recall quantifies the proportion of correctly identified objects among all actual objects (true positives and false negatives), as defined in equation 6.5:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (6.5)$$

As shown in equation 6.6, the AP is then obtained by calculating the area under the precision-recall curve:

$$\text{AP} = \int_0^1 \text{Precision}(\text{Recall}) d\text{Recall} \quad (6.6)$$

The implementation is based on the open-source `keras-retinanet` package [33], which was developed specifically for RetinaNet applications.

### 6.4.3 South Africa Wind Turbine Pre-Dataset

Initially, OSM was used to derive the dataset of all wind turbines in South Africa, based on downloading a complete set of OSM data for South Africa via the Geofabrik download service [23], followed by filtering using the “`esy-osmfilter`” tool [34]. The filtering is done to extract relevant energy infrastructure, with a focus on renewable energy facilities such as wind turbines and PV systems. The filtering process involves using predefined filter criteria to ensure that only the required energy infrastructure elements are selected:

- **Prefilter:** The ‘prefilter’ is used to identify nodes, ways, and relations tagged with attributes like “`power`”: [“`generator`”, “`plant`”, “`solar`”, “`photovoltaic`”]‘ to capture all relevant renewable energy installations.
- **Blackfilter:** A ‘blackfilter’ is applied to exclude certain types of infrastructure that are not of interest, such as those associated with fossil fuels or hydro-based generation. Examples include (“`generator:source`”,

"gas")', ('generator:method', "combustion")', and ('generator:source', "coal")'.

- **Whitefilter:** A 'whitefilter' is also used to ensure that elements explicitly tagged with ('power', "generator")' are retained in the dataset.

This process provides a refined dataset that filters out non-relevant elements and focuses on renewable energy facilities, improving the quality and relevance of the geospatial analysis.

#### 6.4.4 South Africa's wind turbines coordinate correction

The model is applied to high-resolution satellite images provided by Google Satellite as well as Bing image. The initial coordinates for the images are taken from the OSM dataset, with each wind turbine's OSM coordinate serving as the centroid for image extraction. The prepared tiles for analysis are each  $640 \times 640$  pixels with a resolution of  $20 \times 20$  cm. This allows for a focused examination of each turbine location and subsequent adjustment based on the model's predictions. The restriction on the size of the images depends on the permitted options of the services. To generate the coordinates of the final dataset, the resulting coordinates with highest confidence score are adopted. If the confidence score falls below the threshold of 0.2, an additional manual verification is applied.

#### 6.4.5 Additional attribute Enrichment

The additional step of attribute enrichment was to manually add detailed information about the wind farms, including wind farm names, turbine capacities, total farm capacities, turbine types and commissioning years. These attributes were collected through manual enrichment, which was crucial to ensure the accuracy and completeness of the dataset. For this purpose, the operators' websites or publications about the construction of the farms were searched for where possible and used our information enrichment. In addition

to the technical attributes, we assigned spatial information by intersecting each wind farm site with administrative boundaries using the Global Administrative Areas (GADM) dataset [35]. For each site, the corresponding country and first and second level administrative units were identified and added to the dataset. For this purpose, the coordinates of each wind turbine are spatially allocated to corresponding administrative polygons. If a wind farm location was outside a defined polygon due to geometric inaccuracies, it was mapped to the nearest administrative unit to ensure completeness. This referencing simplifies regional analysis and the combination with other datasets. While the enrichment presented here is performed manually, future implementations could benefit from automated techniques such as Natural Language Processing (NLP) or web scraping to extract structured attribute information from online sources and thereby reduce the manual workload.

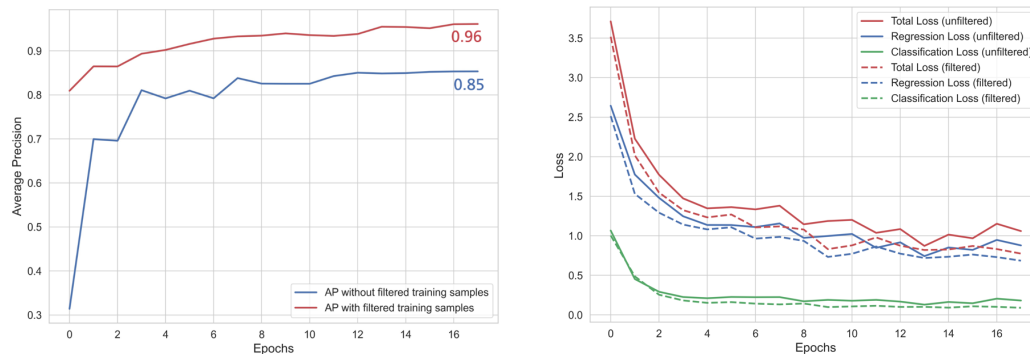
## 6.5 Results

Initially, we briefly present the results from model training, the data extracted and processed from OSM, followed by the results of the location correction. Finally, we present the results of the additional attribute enrichment.

### 6.5.1 Performance and Results of Deep Learning Training

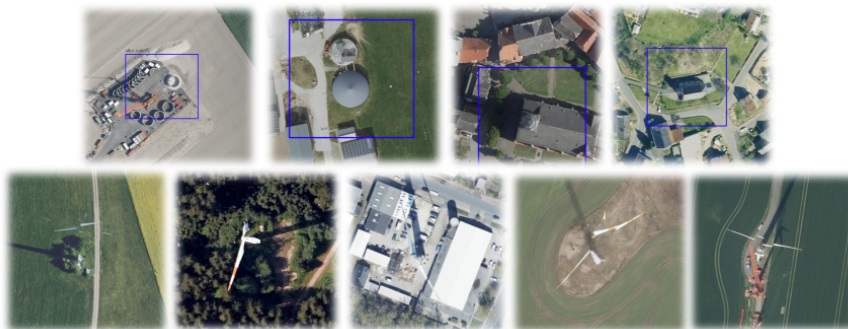
This section presents the results of the DL training, including the loss functions and the accuracy achieved. These results provide insight into the robustness and performance of the applied RetinaNet approach. As Training Progress Summary, the progression of the two losses from the classification and regression networks, as well as the AP, were validated to determine the networks' performance, as displayed in Figure 6.4.

Shown in blue are the results of the first training session, in which all training data was used, and in red the second training session, in which the training data was used after filtering. A consistent upward trend can be



**Figure 6.4:** Training statistics, on the left side the AP during training, on the right side the changing loss.

observed in the AP. Finally, the AP is 85 % for the first training and 96 % for the second training with manually post-filtered samples. In addition, the following Figure 6.4 shows the losses during training phase. Both the regression loss  $L_{loc}$  used to localize the objects and the Smooth L1 loss used for classification decrease significantly and almost evenly in both training runs. The total loss represents the cumulative sum of the individual losses. The training is terminated by early stopping after 17 epochs in each case, indicating no further progress in training. Overall, the loss and AP's curves clearly show the strong generalization of the network based on the training examples. Incorrect recognition are shown in Figure 6.5.



**Figure 6.5:** False positive and false negative examples from the application are summarized in the following. The top line represents incorrectly identified wind turbine, false positives. The bottom line shows turbines that have not been detected.

This includes a construction site, a biogas plant and two churches. Sec-

only, some of the poorly represented turbines are not recognized by the network. This applies to different backgrounds, so that turbines in open fields, in the forest and also in the settlement are not recognized. However, they are also hard to identify during a visual inspection.



**Figure 6.6:** True positive examples from the application using the DOPs images are presented as follows. The upper row displays instances featuring clearly visible and accurately identified wind turbines, where the centroid of the regression boxes serves as the base of the tower. The bottom row, shows correctly detected wind turbines, with less accurate regressive identification on the images.

Examples of correctly recognized wind turbines, conversely, are shown in Figure 6.6. In addition to turbines with good resolution, poorly resolved turbines can also be identified in the images. All images show that the regression locates the towers of the turbines exactly in the centers of the bounding boxes. In other words, the centers of the regression boxes can be interpreted as exact coordinates of the wind turbines.

### 6.5.2 OSM Data Extraction

The initial dataset for South Africa's wind turbines, extracted from OSM, contained a total of 1,546 point features. After a manual review and refinement process, this number was reduced to 1,487 verified wind turbines. Point features with the tags generator and diesel as well as solar were excluded and deleted. However, 55 turbines in the OSM data are not assigned to any wind farm. These are added manually. Among the wind farms, Longyuan Mulilo de Aar 2 North has the highest number of turbines with 96 individual units,

while the smallest wind farm, Buffeljags Abalone Farm, consists of only two turbines. For all turbines without an associated wind farm, a manual assignment to the respective farms was carried out to ensure the completeness of the data. A capacity is given for 351 of the 1,487 turbines, while no capacity data is available for 1,144 turbines. This ensures that all wind turbines are assigned to a wind farm and capacity information if possible.

### 6.5.3 Coordinate Correction

The accuracy of the neural network’s predictions heavily depends on the domain-specific characteristics of the training and application datasets. To analyze this effect, we compare the confidence scores of the predictions for onshore wind turbines in South Africa. Table 6.1 presents the results of the coordinate correction process using both Bing and Google satellite imagery.

**Table 6.1:** Comparison of onshore wind turbine data distributions across South Africa using Bing images and Google images, including count and percentage of coordinates within different Confidence Scores (Confidence S.).

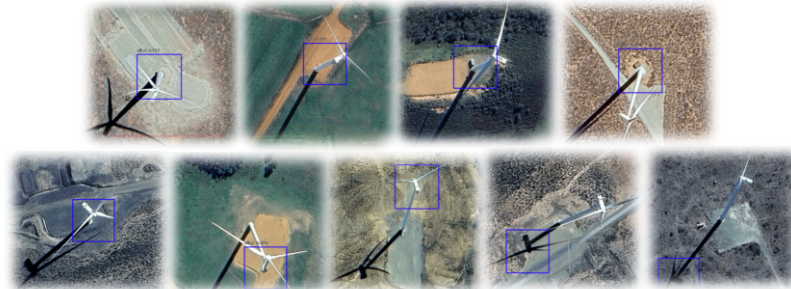
| Confidence S. | Bing Count | Bing (%) | Google Count | Google (%) |
|---------------|------------|----------|--------------|------------|
| < 0.1         | 207        | 13.86    | 116          | 7.77       |
| 0.1 – 0.2     | 362        | 24.25    | 289          | 19.36      |
| 0.2 – 0.3     | 245        | 16.41    | 224          | 15.00      |
| 0.3 – 0.4     | 183        | 12.26    | 223          | 14.94      |
| 0.4 – 0.5     | 142        | 9.51     | 157          | 10.52      |
| 0.5 – 0.6     | 125        | 8.37     | 145          | 9.71       |
| 0.6 – 0.7     | 85         | 5.69     | 129          | 8.64       |
| 0.7 – 0.8     | 51         | 3.42     | 122          | 8.17       |
| > 0.8         | 3          | 0.20     | 45           | 3.01       |
| NULL          | 90         | 6.03     | 43           | 2.88       |
| <b>Total</b>  | 1493       | 100.00   | 1493         | 100.00     |

The Table summarizes results for 1,487 wind turbines, showing that the overall distribution of confidence scores differs considerably between Bing and Google imagery. While only a small fraction of detections reaches confidence scores above 0.8 (0.2 % for Bing and 3.0 % for Google), the majority falls below 0.5 , indicating potential challenges in image consistency or domain

transfer. Despite this, visual inspection confirms the accurate detection of turbines in both datasets, as illustrated in Figure 6.7 and Figure 6.8.



**Figure 6.7:** True positive examples from the application using the Bing images are presented as follows. The upper row displays instances featuring clearly visible and accurately identified wind turbines, where the centroid of the regression boxes serves as the base of the tower. The bottom row, shows correctly identified wind turbines, with less accurate regressive delineation in the images.



**Figure 6.8:** True positive examples from the application using the Google images are presented as follows. The upper row displays instances featuring clearly visible and accurately identified wind turbines, where the centroid of the regression boxes serves as the base of the tower. The bottom row, shows correctly identified wind turbines, with less accurate regressive delineation in the images.

A total of 90 turbines (6.05 %) on the Bing images and 43 turbines (2.89 %) on the Google images are not detected and thus fall into the null category. The analysis shows that 36 of the non-detected South African wind turbines are matched by Bing and Google. All these overlaps are exclusively located within four specific farms: San Kraal Wind Farm, Phezukomoya, Cookhouse Wind Farm, and Wolf Wind Farm. The visual inspection of the zero category shows that there are often construction sites for wind turbines

at the locations, which means that some of the images are not up-to-date enough to show the existing wind turbine. In addition to the accuracy of the detection, the accuracy of the regression is examined in the following. Table 6.2 summarizes the distances between pre-dataset coordinates and regression analysis.

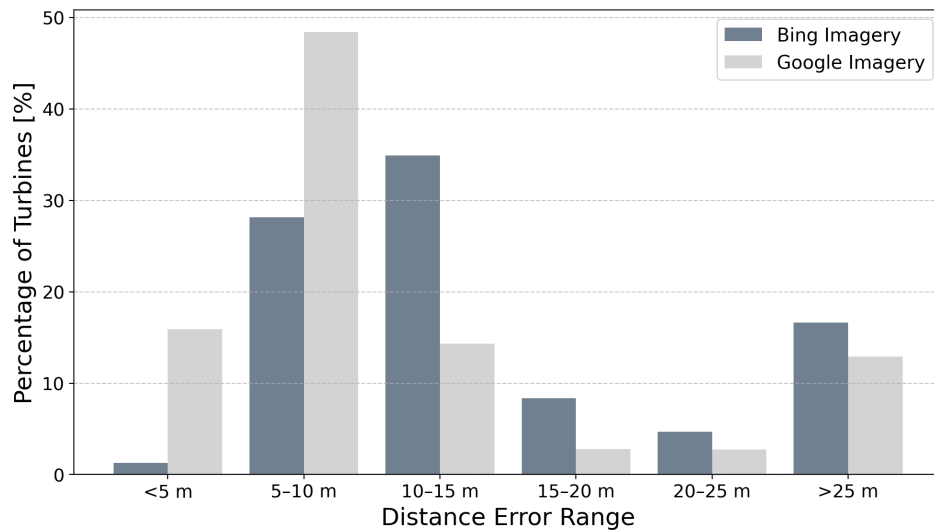
**Table 6.2:** Comparison of wind turbine coordinate deviations in South Africa based on image sources. The percentage of wind turbines that lie within certain distance ranges between the original OSM coordinates and the coordinates corrected using Bing and Google satellite images are shown.

| Distance Range [m]  | Bing (%) | Google (%) |
|---------------------|----------|------------|
| < 5                 | 1.27     | 15.87      |
| 5–10                | 28.13    | 48.43      |
| 10–15               | 34.90    | 14.33      |
| 15–20               | 8.37     | 2.81       |
| 20–25               | 4.69     | 2.75       |
| >25                 | 16.61    | 12.93      |
| Not Detected (NULL) | 6.03     | 2.88       |

The Table 6.2 presents the distribution of coordinate deviations for wind turbines in South Africa, comparing results derived from Bing and Google Maps. The deviations are categorized into six distance intervals: <5 m, 5–10 m, 10–15 m, 15–20 m, 20–25 m, and >25 m. A significant portion (64.3 %) of the Google-based coordinates fall within 10 m of the reference, whereas only 29.4 % of the Bing-based coordinates achieve this accuracy. The largest deviations (>25 m) occur in 16.6 % of Bing and 12.9 % of Google. To provide a visual summary of the distribution of location errors, a histogram of the distance deviations was created (Figure 6.9). It shows the proportion of turbines falling within specific distance ranges for both Bing and Google images.

#### 6.5.4 Wind turbine dataset

An overview of the existing wind farms in South Africa is provided below. The summarizing Table 6.3 combines spatial information with key technical attributes for each wind turbine. It includes both operational and under-



**Figure 6.9:** Histogram of wind turbine location errors based on Bing and Google imagery. It shows the percentage of turbines whose corrected coordinates fall within different distance ranges compared to their original OSM positions.

construction sites and was cross-checked and harmonized based on multiple publicly available sources. Listed are commissioning years, the number of turbines, the total installed capacity in MW, the rated capacity per turbine in MW and the type of turbine installed in each wind farm.

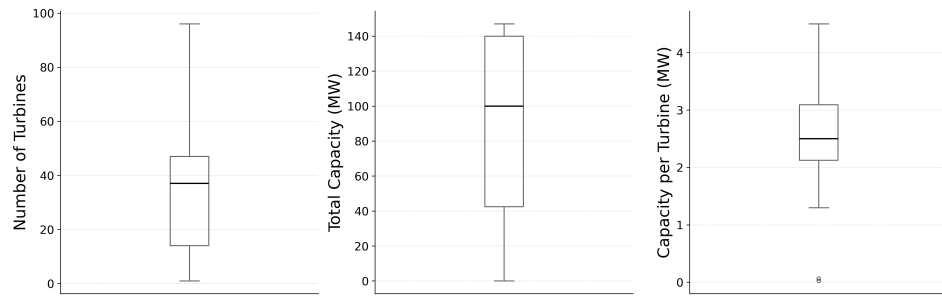
Two wind farms, Phezukomoya and San Kraal, are still under construction. In these cases, not all turbines have yet been built or identified, which explains deviations from the detailed point-based turbine dataset. A more detailed graphical evaluation is summarized in Figure 6.10. Boxplots illustrate three key parameters from left to right: the number of turbines per wind farm, the total installed capacity, and the specific capacity per turbine.

The number of turbines varies significantly, ranging from small farms with only 2 to 4 turbines to large-scale farms hosting up to 96 turbines. However, the majority of wind farms contain between around 15 and under 50 turbines. On average, there are 37 turbines within a farm. The total installed capacity per wind farm ranges from as little as 0.1 MW to 147 MW. The majority of projects lie within the interquartile range of 35 to 140 MW, the median is 100 MW. The nominal capacity per turbine spans a wide range, from small-scale units with 25 kW to modern high-capacity turbines rated at 4.5 MW.

**Table 6.3:** Summary of Wind Turbines in South Africa, including the commissioning year, number of turbines, total capacity, capacity per turbine, and turbine type for each wind farm.

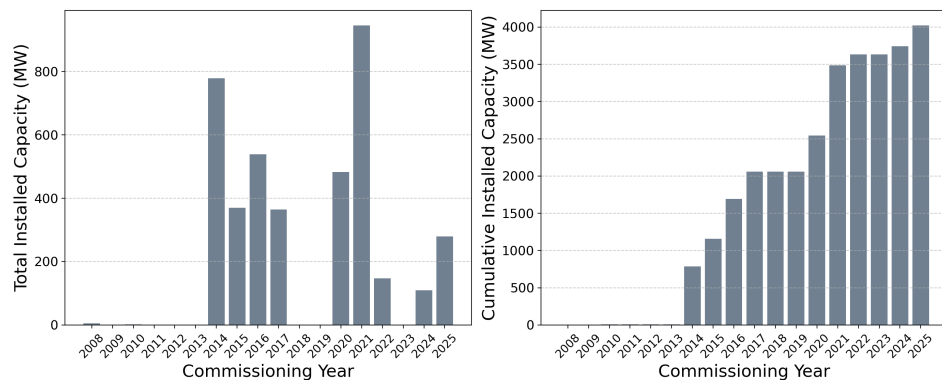
| Name of Farm                | Comm. Year | Turbines | Tot. Cap. (MW) | Cap./Turbine (MW) | Turbine Type             |
|-----------------------------|------------|----------|----------------|-------------------|--------------------------|
| Amakhala Emoyeni            | 2016       | 56       | 134.4          | 2.4               | Nordex N117/2400         |
| Buffeljags Abalone          | 2012       | 2        | 0.13           | 0.065             | Horizontal Axis Turbine  |
| Caledon - Dassieklip        | 2015       | 9        | 27             | 3                 | Sinovel SL 3000/90       |
| Chaba Wind Farm             | 2015       | 7        | 21             | 3                 | Vestas V112-3.0          |
| Coega Wind Farm             | 2010       | 2        | 3.6            | 1.8               | General Electric GE2.5XL |
| Cookhouse Wind Farm         | 2014       | 66       | 138.6          | 2.1               | Suzlon S88               |
| Copperton Wind Farm         | 2021       | 34       | 102            | 3                 | Nordex N117/2400         |
| Darling Wind Farm           | 2008       | 4        | 5.2            | 1.3               | Fuhrländer FL1250        |
| Dorper Wind Farm            | 2014       | 40       | 100            | 2.5               | Nordex N100/2500         |
| Excelsior Energy Facility   | 2020       | 13       | 32.5           | 2.5               | Vestas V100-2.5          |
| Garob Wind Farm             | 2021       | 46       | 145            | 3.15              | Nordex AW125/3150        |
| Golden Valley Wind          | 2020       | 48       | 120            | 2.5               | Vestas V120-2.5          |
| Gouda Wind Facility         | 2015       | 46       | 138            | 3                 | Acciona AW3000/116       |
| Grassridge Wind Farm        | 2016       | 20       | 61.5           | 3.075             | Vestas V112-3.0          |
| Jeffreys Bay Wind Farm      | 2014       | 60       | 138            | 2.3               | Siemens SWT-2.3-108      |
| Kangnas Wind Farm           | 2020       | 61       | 140            | 2.3               | Siemens SWT-2.3-108      |
| Karusa Wind Farm            | 2021       | 35       | 147            | 4.2               | Vestas V136-4.2          |
| Khobab Wind Farm            | 2017       | 61       | 140            | 2.3               | Siemens SWT-2.3-108      |
| Loeriesfontein 2            | 2017       | 61       | 140            | 2.3               | Siemens SWT-2.3-108      |
| Longyuan M. De Aar 2 North  | 2017       | 96       | 144            | 1.5               | Guodian UP86             |
| Longyuan M. De Aar Maanh.   | 2016       | 67       | 100            | 1.5               | Guodian UP86/1500        |
| Noblesfontein Wind Farm     | 2014       | 41       | 73.8           | 1.8               | Vestas V100-1.8          |
| Nojoli Wind Farm            | 2016       | 44       | 88             | 2                 | Vestas V100-2.0          |
| Noupoort Mainstream         | 2016       | 35       | 80             | 2.3               | Siemens SWT-2.3-108      |
| Nxuba Wind Farm             | 2020       | 47       | 147            | 3.6               | Nordex AW 125/3150       |
| Oyster Bay Wind Farm        | 2021       | 41       | 140            | 3.4               | Vestas V117-3.45         |
| Perdekraal East Wind Farm   | 2020       | 48       | 110            | 2.3               | Siemens SWT-2.3-108      |
| Phezukomoya                 | 2025*      | 35**     | 140            | 4                 | Vestas V136-4.0          |
| Red Cap - Gibson Bay        | 2017       | 37       | 111            | 3                 | Nordex N117/3000         |
| Red Cap Kouga - Oyster Bay  | 2015       | 32       | 80             | 2.5               | Nordex N90/2500          |
| Roggeveld Wind Farm         | 2022       | 47       | 147            | 3.15              | Nordex AW125/3150        |
| San Kraal Wind Farm         | 2025*      | 35**     | 140            | 4                 | Vestas V136-4.0          |
| Sere Wind Farm              | 2015       | 46       | 100            | 2.2               | Siemens SWT-2.3-108      |
| Silo District's Sustainable | 2024       | 4        | 0.1            | 0.025             | Vertical Axis Turbine    |
| Soetwater Wind Farm         | 2021       | 35       | 147            | 4.2               | Vestas V136-4.2          |
| Tsitsikamma Community       | 2016       | 31       | 95             | 3.07              | Vestas V112-3.0          |
| Umoya Hopefield Farm        | 2014       | 37       | 66.6           | 1.8               | Vestas V100-1.8          |
| Van Stadens Wind Farm       | 2014       | 9        | 27             | 3                 | Sinovel SL 3000/113      |
| Waainek Wind Farm           | 2016       | 8        | 24.6           | 3.075             | Vestas V112-3.075        |
| Wesley-Ciskei Farm          | 2021       | 10       | 34.5           | 3.45              | Vestas V126-3.45         |
| West Coast One Farm         | 2015       | 47       | 94             | 2                 | Vestas V90-2.0           |
| Wolf Wind Farm              | 2024       | 17       | 85             | 5                 | Vestas V162/V163         |

\* under construction. \*\* not all wind turbines have been built yet, thus do not match the detailed dataset.



**Figure 6.10:** Summary statistics of key parameters of South African wind farms. Number of turbines per wind farm (left side), total installed capacity (MW) per wind farm (in the middle), and capacity per turbine (MW) (on the left). The boxplots contains the median, interquartile range, and outliers in the dataset

Most turbines, however, fall within the interquartile range of 2.3 to 3.1 MW, with mean capacity of a turbine is 2.5 MW, typical for recent onshore turbine installations.

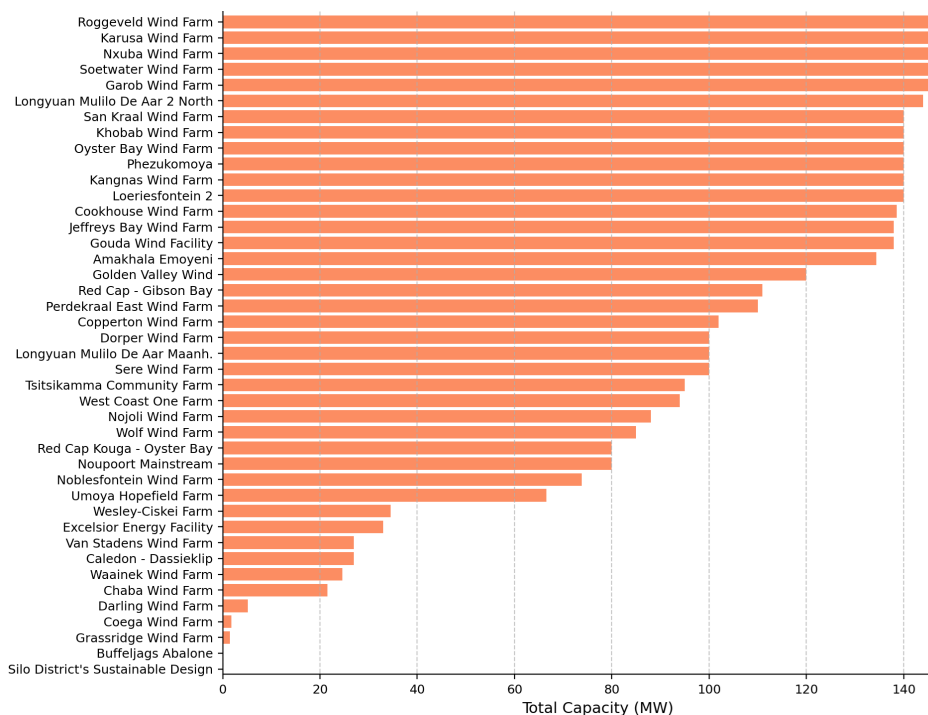


**Figure 6.11:** Development of wind power capacity in South Africa by year. The annual installed wind power capacity from 2008 to 2025 is shown on the left-hand side, and the cumulative installed capacity on the right-hand side.

The number of turbines varies significantly, ranging from small farms with only 2 to 4 turbines to large-scale farms hosting up to 96 turbines. However, the majority of wind farms contain between around 15 and under 50 turbines. On average, there are 37 turbines within a farm. The total installed capacity per wind farm ranges from as little as 0.1 MW to 147 MW. The majority of projects lie within the interquartile range of 35 to 140 MW, the median is 100 MW. The nominal capacity per turbine spans a wide range, from small-scale units with 25 kW to modern high-capacity turbines rated at 4.5 MW.

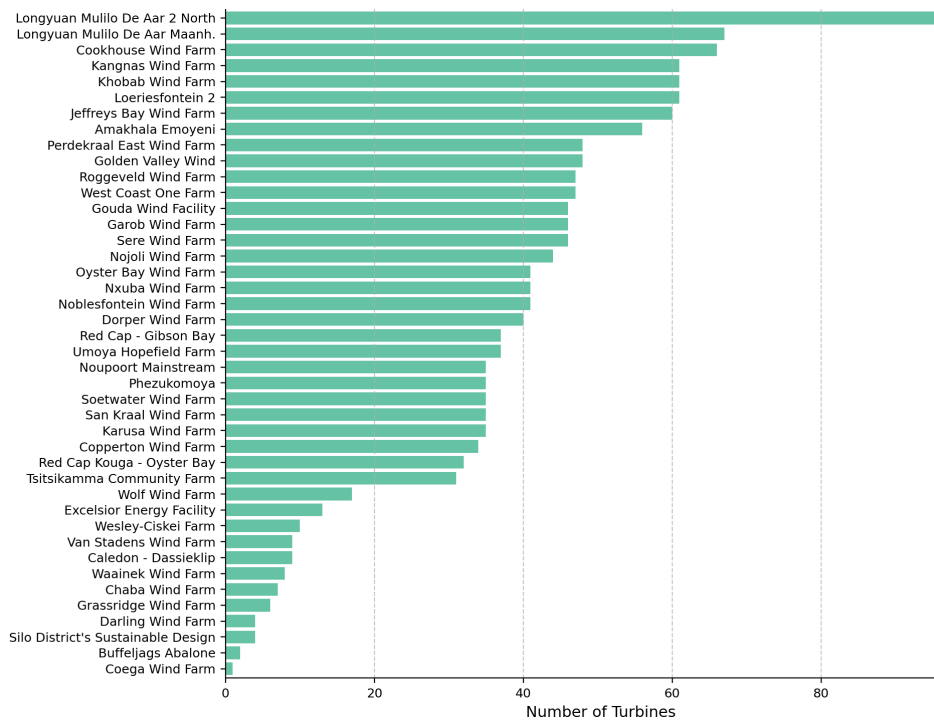
Most turbines, however, fall within the interquartile range of 2.3 to 3.1 MW, with mean capacity of a turbine is 2.5 MW, typical for recent onshore turbine installations.

Figure 6.11 shows the development of wind power capacity in South Africa over time, starting with the first installations in 2008 through to 2025. To illustrate the growth trend in recent years, the left panel shows the annual installed capacity between 2008 and 2025 based on the commissioning years of the individual wind farms. At least three different phases of capacity growth can be observed: an initial phase with isolated installations between 2008 and 2012, a first strong expansion phase from 2014 to 2021 with significant annual growth and a second expansion phase since 2022. The largest annual increases were in 2016 with around 580 MW and in 2021 with almost 570 MW of newly installed capacity. The right panel shows the cumulative installed capacity over the same period. By 2025, the total installed capacity will reach over 3.9 GW.



**Figure 6.12:** The total installed capacity (MW) of the individual wind farms is shown. The wind farms are listed by size, starting with the largest.

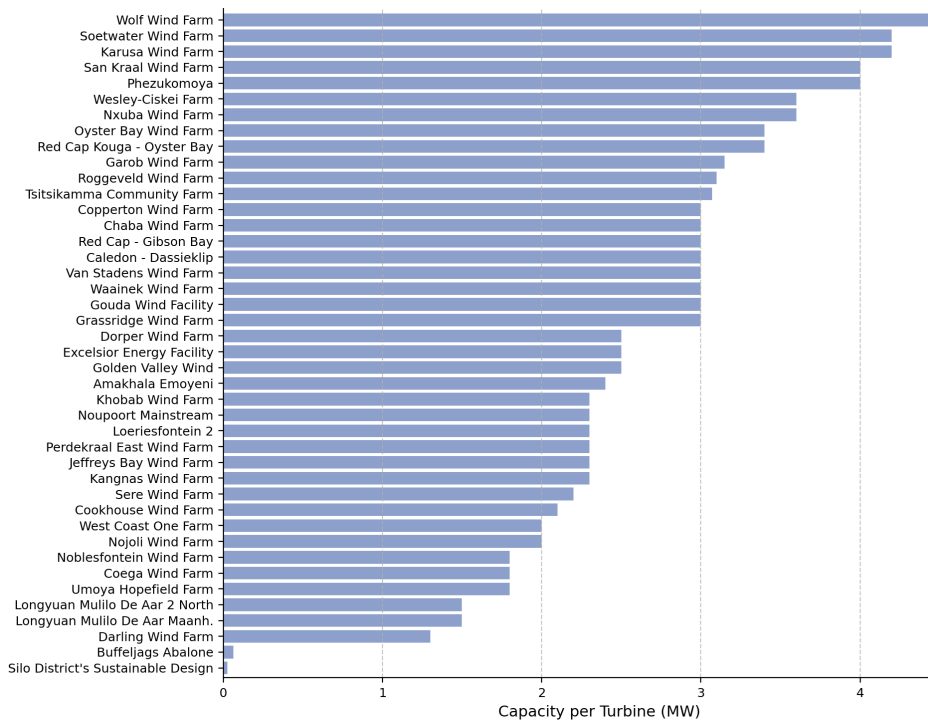
Figure 6.12 shows the total installed capacity per wind farm in descending order, distributed across 42 different wind farms with capacities ranging from 147 MW to 0.1 MW. The bar lengths provide a quick indication of the relative capacity of the individual wind farms. This ranking makes it easier to identify the wind farms in South Africa with the highest rated capacity. The largest farms - such as Roggeveld, Karusa, Nxuba or Soetwater - reach around 140-150 MW. The smallest wind farms such as Coega, Buffeljags Abalone Farm and Silo District's Sustainable Design have significantly lower total capacities of less than 2 MW.



**Figure 6.13:** This figure shows the number of wind turbines per wind farm. Wind farms with more turbines are shown at the top, while smaller farms with fewer turbines are listed further down.

Alongside the total installed capacity, the Figure 6.13 shows the number of wind turbines installed in the individual wind farms in descending order. The order provides a quick overview of the locations with a particularly high amount of turbines. Longyuan Mulilo De Aar 2 North stands out with 96 turbines, while Longyuan Mulilo De Aar Maanhaarberg with 67 turbines and

Cookhouse Wind Farm with 66 turbines are the next largest farms. Coega Wind Farm has only two turbines. In combination with the capacity data, this also gives an indication of the average turbine size in each wind farm.



**Figure 6.14:** The capacity per wind turbine in megawatts (MW) for the wind farms in South Africa is shown. The values are sorted in ascending order, so that farms with lower capacity per turbine are shown at the bottom and more powerful farms at the top.

The Figure 6.14 shows the nominal capacity per wind turbine at each wind farm. This overview can be used to determine which sites mainly use smaller turbines and which rely on turbines with a higher rated capacity. The frequent use of turbines with a capacity of 2.3 MW (here with Siemens SWT-2.3 turbines) in the Jeffreys Bay Wind Farm, Kangnas Wind Farm, Khobab Wind Farm, Loeriesfontein 2, Noupoort Mainstream, and Perdekraal East Wind Farm is particularly evident. However, turbines with a capacity of 3 MW are also widely used in Dassieklip, Chaba Wind Farm, Copperton Wind Farm, Gouda Wind Facility, Red Cap - Gibson Bay, and Van Stadens Wind Farm. The lower end of the scale includes turbines with relatively

small capacities, such as those at Buffeljags Abalone Farm or the vertical axis turbines in the Silo District. Higher bars correspond to larger capacity turbines, such as the Vestas V136 and V162 models with capacities with up to 5 MW.

The following section of the results focuses on the spatial distribution of wind turbines in South Africa. The installed wind power capacity is concentrated in just three of the country's nine provinces, Northern Cape, Eastern Cape, and Western Cape. Table 6.4 provides a summary of wind energy infrastructure at the provincial level.

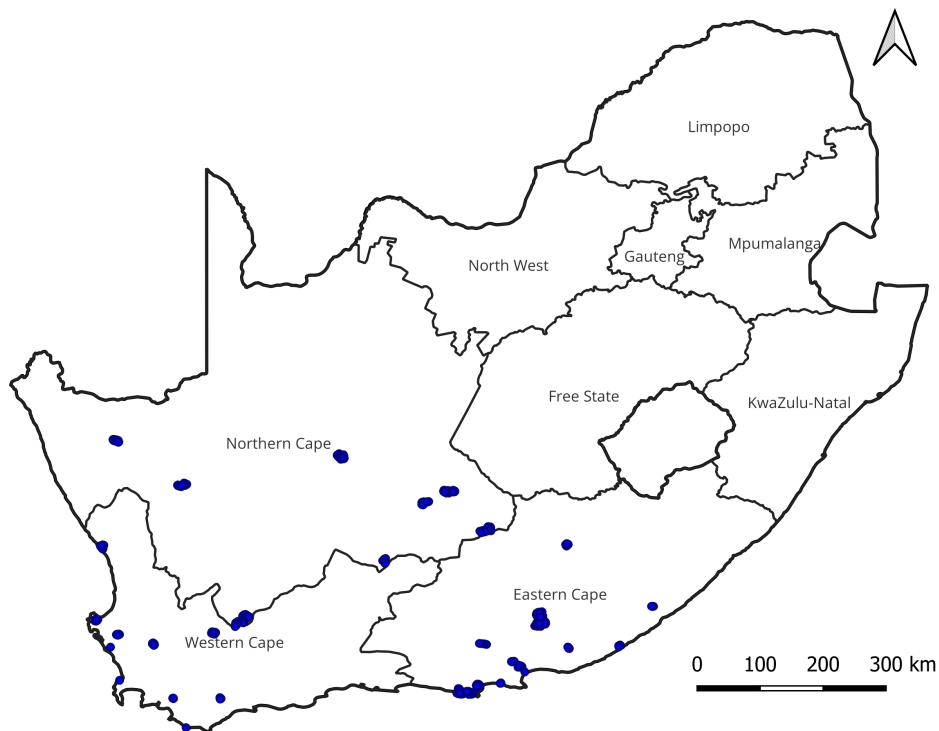
**Table 6.4:** Overview of wind farms in South Africa by province. The number of different wind farms, the total number of turbines and the aggregated installed capacity are shown.

| Province      | Wind Farms | Turbines | Total Capacity (MW) |
|---------------|------------|----------|---------------------|
| Eastern Cape  | 18         | 575      | 1,571               |
| Northern Cape | 14         | 656      | 1,670               |
| Western Cape  | 10         | 256      | 575                 |

The majority of capacity is located in the Northern Cape and Eastern Cape, which together host 32 wind farms and 1,231 turbines. The Western Cape follows with 10 wind farms. Together, the Eastern Cape and the Northern Cape account for 1,571 MW and 1,670 MW of installed capacity, respectively. The Western Cape contributes 575 MW, bringing the total installed capacity in these three provinces to more than 3,800 MW. The Roggeveld Wind Farm represents a special case, as it spans across two provinces. Since the majority of its 42 turbines are located in the Northern Cape and only five fall within the Western Cape, the entire wind farm is attributed to the Northern Cape for consistency in the provincial analysis.

Figure 6.15 illustrates the spatial distribution of all 42 existing wind farms in South Africa. It clearly shows that the facilities are exclusively located in the southwestern provinces, particularly in the Northern Cape, Eastern Cape, and Western Cape.

To supplement the analysis at provincial level, a more detailed spatial aggregation was carried out at district municipality level. This approach enables a finer resolution of the spatial distribution and highlights the dif-

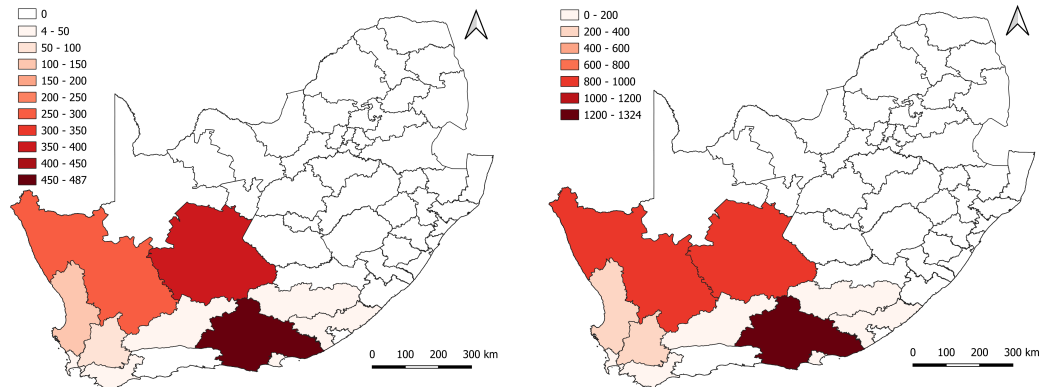


**Figure 6.15:** Spatial distribution of all existing wind turbines in South Africa, marked in blue, highlighting their locations across the country.

ferences within the provinces in the expansion of wind energy. Figure 6.16 shows the total installed capacity on the one hand and the number of wind turbines per municipality on the other. The results show a very uneven distribution, with a limited number of municipalities hosting the majority of turbines and installed capacity. In contrast, many regions are still completely undeveloped, indicating a significant spatial concentration of wind energy infrastructure.

### 6.5.5 Validation against Official Capacity Figures

In order to assess the accuracy of the data compiled in the publication with regard to installed capacity, the total installed capacity of wind farms in operation was compared with the official IPP project database [15]. According to our data, a total of 3,627 MW is currently in operation. The IPP database lists an installed capacity of 3,428 MW (Wolf Wind Farm is considered to



**Figure 6.16:** Spatial distribution of wind energy infrastructure by municipality. Map (on the left side) displays the number of wind turbines, while map (on the right side) shows the total installed wind capacity (MW). The patterns reveal significant regional clustering, with a small number of municipalities concentrating the majority of infrastructure.

be already in operation). The slight deviation of less than 200 MW can be explained by the inclusion of additional wind farms in our dataset that are not part of the projects supported by the REIPPPP, such as small or privately financed farms. According to the official database, three additional wind farms, each with a capacity of 140 MW, are currently in the planning phase but have not yet been commissioned and are therefore not included in our dataset. This comparison confirms both the consistency of our data with national figures and the added value of including additional data sources.

## 6.6 Discussion

This study presents a comprehensive and spatially validated dataset of wind power infrastructure in South Africa. With 1,487 turbines across 42 wind farms and a total installed capacity exceeding 3.9 GW, the dataset offers both spatial and technical detail, with a total of 3,6 GW currently in operation. Most turbines are concentrated in the Northern Cape, Eastern Cape, and Western Cape provinces, reflecting the regional clustering of wind development in the country. In addition to the spatial information, the dataset includes harmonized metadata such as commissioning year, turbine type,

wind farm capacity, and per-turbine capacity. These attributes were manually collected and cross-checked from various sources.

Although labor-intensive, this enrichment process significantly increases the usability and reliability of the dataset—enabling advanced applications in energy system modelling, infrastructure planning, and policy design. However, manually collecting turbine-specific information also revealed common challenges regarding the availability and quality of public data. The information on operators’ websites was often unstructured, inconsistently formatted, or partially incomplete. In several cases, additional sources such as press releases, freely accessible news articles, and energy-related databases were consulted. While these secondary sources were useful for cross-checking, they sometimes contained unverifiable or contradictory data, highlighting the limitations of public reporting on renewable energy infrastructure. These challenges underline the crucial role of manual processing within the overall pipeline, which, despite advances in automation, remains indispensable for ensuring technical completeness and high data quality.

While most of the data processing, including the localization of the turbines for coordinate correction using DL methods, was automated, manual steps were essential to ensure the technical completeness and reliability of the dataset. In particular, turbine attributes such as turbine type, capacity and year of commissioning were manually enriched by comparing several publicly available sources (e.g. operator website, project reports, press releases). This manual effort was necessary because the detailed technical metadata in open datasets such as OSM or national databases is almost completely missing, incomplete or inconsistent. If the pipeline were transferred to other countries or regions, a similar manual enrichment step would probably be required due to the heterogeneous availability of data and the different reporting standards worldwide.

Automated extraction of attributes from semi-structured text sources (e.g. using NLP methods) could be investigated as a future extension to partially automate this step. However, full automation is currently only possible to a limited extent due to the lack of standardized and structured publication of turbine metadata. Furthermore, regular updates of the dataset (e.g. every

1-2 years) would require re-verification of new wind farm projects and updating of technical attributes, meaning that some level of manual verification and enrichment will still be essential to maintain data quality. Nevertheless, further improvements, such as the integration of automated web scraping techniques combined with manual quality checks, could significantly reduce the manual workload while ensuring high standards of data accuracy.

The dataset was systematically checked against several external sources to ensure its completeness. A comparison with the official South African IPP database [15] confirms that all 34 large wind farms currently in operation are included in this dataset. In addition, two projects under construction and several smaller wind farms not listed in the official database have been included. The dataset thus shows that it not only covers large infrastructures but also takes into account smaller and emerging projects. It is noteworthy that the aggregate installed capacity of the wind farms currently included in our dataset is largely consistent with the total capacity reported in official IPP sources, further supporting the validity and representativeness of the dataset.

The coordinate correction process based on RetinaNet was trained on German aerial imagery and applied to South African wind turbine locations using both Bing and Google satellite data. The application resulted in a notable drop in confidence scores, which can be attributed to the domain shift between training and application imagery—a typical challenge in DL when transferring models across data sources. Despite this, the visual and statistical evaluation confirms a high localization accuracy. More than 60 % of Google-based predictions and 29 % of Bing-based predictions fall within a 10 m range from the reference coordinates.

The model’s ability to correctly identify turbine locations across different landscapes and image types confirms its practical value as a scalable validation tool. Due to the lack of official, publicly available data on wind turbines in South Africa, the spatial validation of the turbine coordinates was carried out by visual comparison with high-resolution satellite images from Google and Bing. Although this method does not replace GPS-based ground validation, it improves accuracy compared to the raw OSM data. Furthermore, the

high degree of agreement between the visually validated and corrected coordinates suggests that the original OSM point data already provides relatively high positional accuracy in many cases.

However, some aspects of the detection and correction process could be improved in future applications. First, the exclusive use of a RetinaNet architecture could limit performance in more complex or visually diverse environments. Although RetinaNet has demonstrated high accuracy in correcting wind turbine coordinates, its performance is sensitive to variations in image quality and background complexity. This may reduce its generalizability when applied to unknown regions or alternative satellite image sources. These limitations become more apparent in large-scale applications where wind turbines need to be detected across large areas without predefined coordinate references. In such contexts, it can be difficult for the model to distinguish wind turbines from visually similar structures such as high-voltage pylons, cranes, or communication towers, especially in complex environments.

Alternative approaches—such as modern transformer-based models—could offer greater robustness and accuracy, particularly under conditions of visual ambiguity or clutter. Second, the image data itself could be further diversified. The current approach is limited to single time frames from Bing and Google images, which may not capture seasonal variations or recent changes in infrastructure. The use of time series imagery or higher-resolution commercial datasets could improve model generalization and enable the detection of newer or smaller installations.

From a methodological perspective, the study highlights the importance of combining open spatial data, deep learning, and manual curation to overcome the usual limitations of public datasets. OSM offers broad coverage but lacks standardization and, in some cases, location accuracy. The integration of DL fills this gap by refining the location data, while manual enrichment ensures the completeness and technical detail required for meaningful application. Together, these components form a transferable and reproducible workflow for the creation of high-quality renewable energy datasets in data-poor regions.

## 6.7 Conclusions

This study presents the most accurate, comprehensive, and up-to-date dataset on wind turbines and wind farms currently available for South Africa. By integrating publicly available OSM data, high-resolution satellite imagery, and advanced DL-based coordinate correction using RetinaNet, the spatial accuracy of turbine locations has been significantly improved. The dataset has been further enhanced through manual enrichment with important technical and temporal attributes such as wind farm names, turbine types, capacities, and commissioning years—information that is often missing or inconsistent in existing sources. Spatial metadata has been mapped to administrative boundaries from the GADM database, enabling regional analysis and integration with other relevant datasets.

This dataset thus provides accurate turbine coordinates, technical specifications, and harmonized metadata. It includes not only all large wind farms currently listed in the South African IPP project database, but also smaller and emerging wind farms that are not covered by official sources. The result is a high-quality, freely accessible dataset that provides a solid foundation for research, energy system modelling, infrastructure planning, and policy evaluation. It makes an important contribution to the open energy data landscape and provides a transferable methodology for creating similarly detailed datasets in other countries and for other renewable energy technologies.

Keeping the data up to date is particularly important given the rapid expansion of wind energy infrastructure and evolving project developments. In order to continue to provide valuable support to this ongoing development in South Africa’s dynamic wind energy sector, we are currently in discussions with national stakeholders to facilitate regular updates to the datasets. The aim is to establish a process that ensures updates every 1-2 years, including the review of new wind farms and the enrichment of technical attributes.

The dataset is freely available for download [36]. We strongly encourage its reuse and further development by the broader research and planning community.

## Bibliography

- [1] Global Wind Energy Council. Global wind report 2024, 2024. Accessed on April 22, 2025.
- [2] International Energy Agency (IEA). World energy outlook 2022. <https://www.iea.org/reports/world-energy-outlook-2022>, 2022. Accessed on January 17, 2025.
- [3] Ting Zhang, Bo Tian, Dhritiraj Sengupta, Lei Zhang, and Yali Si. Global offshore wind turbine dataset. *Scientific Data*, 8(1):191, 2021.
- [4] Thorsten Hoerer, Stefanie Feuerstein, and Claudia Kuenzer. Deepowl: A global offshore wind turbine data set derived with deep learning from sentinel-1 data. *Earth System Science Data*, 14(9):4251–4270, 2022.
- [5] Min Han, Huabin Wang, Guanghui Wang, and Yu Liu. Targets mask u-net for wind turbines detection in remote sensing images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42:475–480, 2018.
- [6] Narayana Darapaneni, A Jagannathan, Vigneshwaran Natarajan, Guruprasadh Vadakkupattu Swaminathan, S Subramanian, and Anwesh Reddy Paduri. Semantic segmentation of solar pv panels and wind turbines in satellite images using u-net. In *2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS)*, pages 7–12. IEEE, 2020.
- [7] Michael Mommert, Linus Scheibenreif, Joëlle Hanna, and Damian Borth. Power plant classification from remote imaging with deep learning. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pages 6391–6394. IEEE, 2021.
- [8] Tingting He, Yihua Hu, Fashuai Li, Yuwei Chen, Maoxin Zhang, Qiming Zheng, Yukan Jin, and He Ren. Mapping land-and offshore-based wind turbines in china in 2023 with sentinel-2 satellite data. *Renewable and Sustainable Energy Reviews*, 214:115566, 2025.

- 
- [9] N Mandroux, S Drouyer, and R Grompone von Gioi. Multi-date wind turbine detection on optical satellite images. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2:383–390, 2022.
- [10] Pukaiyuan Yang, Zhigang Zou, and Wu Yang. Mapping wind turbine distribution in forest areas of china using deep learning methods. *Remote Sensing*, 17(5), 2025.
- [11] Caleb Robinson, Anthony Ortiz, Allen Kim, Rahul Dodhia, Andrew Zolli, Shivaprakash K. Nagaraju, James Oakleaf, Joe Kiesecker, and Juan M. Lavista Ferres. Global renewables watch: A temporal dataset of solar and wind energy derived from satellite imagery. Preprint, arXiv:2503.14860, 2025. Accessed on 4 June 2025.
- [12] Yasen Fei, Yongnian Gao, Hongyuan Gu, Yongqi Sun, and Yanjun Tian. Yolov5\_cdb: A global wind turbine detection framework integrating cbam and dbscan. *Remote Sensing*, 17(8):1322, 2025.
- [13] Anton Eberhard and Raine Naude. The south african renewable energy independent power producer procurement programme: A review and lessons learned. *Journal of Energy in Southern Africa*, 27(4):1–14, 2016.
- [14] Shipokosa Paulus Mashatile. Remarks by deputy president shipokosa paulus mashatile at the south africa-ireland business forum, May 2023. Accessed on 6 Mai 2025.
- [15] Department of Electricity and Energy, Republic of South Africa. Ipp projects database. <https://www.ipp-projects.co.za/ProjectDatabase>, 2025. Accessed on April 23, 2025.
- [16] Maximilian Kleebauer, Axel Braun, Daniel Horst, and Carsten Pape. Enhancing wind turbine location accuracy: A deep learning-based object regression approach for validating wind turbine geo-coordinates. In *IGARSS 2024-2024 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2024.

- 
- [17] Maximilian Kleebauer, Christopher Marz, Christoph Reudenbach, and Martin Braun. Multi-resolution segmentation of solar photovoltaic systems using deep learning. *Remote Sensing*, 15:5687, 2023.
- [18] Nicolene Botha, Toshka Coleman, Gert Wessels, Maximilian Kleebauer, and Stefan Karamanski. Power generation time series for solar energy generation: Using atlite in south africa. *Solar*, 4, 2024.
- [19] Anni Niemi, Salim Bouchakour, Bendaas Ismail, Kada Bouchouicha, Abdelhak Razagui, Nelli Putkonen, and Juha Kiviluoma. The curious case of wind power in the desert. *IET Conference Proceedings*, 2024:536–541, 02 2025.
- [20] OpenStreetMap contributors. OpenStreetMap, 2024. Accessed on November 27, 2024.
- [21] Muki Haklay. How good is volunteered geographical information? a comparative study of openstreetmap and ordnance survey datasets. *Environment and Planning B: Planning and Design*, 37(4):682–703, 2010.
- [22] Christopher Barrington-Leigh and Adam Millard-Ball. The world’s user-generated road map is more than 80% complete. *PLOS ONE*, 12(8):e0180698, 2017.
- [23] Geofabrik GmbH. Geofabrik download service: South africa, 2024. Accessed on December 3, 2024.
- [24] Noel Gorelick, Matt Hancher, Mike Dixon, Simon Ilyushchenko, David Thau, and Rebecca Moore. Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*, 202:18–27, 2017.
- [25] Google. Google satellite imagery. <https://mt1.google.com/vt/lyrs=s&x={x}&y={y}&z={z}>, 2024. Accessed on November 11, 2024.
- [26] Microsoft Corporation. Bing maps api. <https://dev.virtualearth.net/REST/v1/Imagery/Map/Aerial>. Accessed on December 5, 2024.

- 
- [27] Federal Network Agency (BNetzA). Core energy market data register (mastr), 2025. Accessed on January 14, 2025.
- [28] Bundesamt für Kartographie und Geodäsie. Dokumentation digitale orthophotos. 2023.
- [29] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollar. Focal loss for dense object detection. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2999–3007, 22.10.2017 - 29.10.2017.
- [30] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [31] Tsung-Yi Lin, Piotr Dollar, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 936–944. IEEE, 21.07.2017 - 26.07.2017.
- [32] Ross Girshick. Fast r-cnn. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 1440–1448. IEEE, 07.12.2015 - 13.12.2015.
- [33] Hans Gaiser, Maarten de Vries, Valeriu Lacatusu, vcarpani, Ashley Williamson, Enrico Liscio, Andrés, Yann Henon, jjiun, Cristian Gratie, Mihai Morariu, Charles Ye, Martin Zlocha, Ben Weinstein, Rodrigo Meira de Andrade, Pedro Conceição, Alexander Pacha, hannedevartsen, Daniyal Shahrokhian, Wudi Fang, Mike Clark, meagerYak, Iver Jordal, Max Van Sande, Jin, Etienne-Meunier, Andrew Grigorev, Guillaume Erhard, Eduardo Ramos, and Denis Dowling. *fzyr/keras-retinanet 0.5.1*. 2019.
- [34] Adam Pluta and Ontje Lünsdorf. esy-osmfilter – A Python Library to Efficiently Extract OpenStreetMap Data. *Journal of Open Research Software*, 8(1):19, 2020.

- [35] Global Administrative Areas (GADM). Gadm database of global administrative areas, version 4.1, 2023.
- [36] Maximilian Kleebauer. Dataset according to "a wind turbines dataset for south africa: Open street map data, deep learning based geo coordinate correction and capacity analysis", 2025. Accessed on April 28, 2025.

# Chapter 7

## Conclusions

In this dissertation, modern deep learning methods are examined in combination with remote sensing imagery to improve the recognition, correction, and analysis of renewable energy systems. The work is guided by three central research questions, which form the core of the dissertation: (1) how training data for deep learning models can be generated efficiently and reproducibly, (2) how segmentation approaches perform across different spatial resolutions and geographic contexts, and (3) how deep learning methods can improve the positional accuracy of existing renewable energy system datasets. These questions provide the foundation for the research hypotheses defined in Chapter 1.4, which structure the synthesis of the results presented below. The following chapter thus revisits the hypotheses, summarizes the most important contributions of the dissertation, and discusses limitations, broader implications, and future perspectives.

## 7.1 Synthesis of the Central Research Hypotheses

The following sections revisit the three central research hypotheses and discuss them with the most important findings of the dissertation.

**Hypothesis 1:** *Semi-automated methods for generating training samples can provide reliable inputs for deep learning-based recognition of renewable energy systems and offer a possible way to reduce dependence on time-consuming manual labeling.*

This hypothesis is discussed in Chapter 2, which presents a semi-automatic workflow that combines registry information with building geometries to efficiently generate training examples. It demonstrates how address data can be combined with spatial data and aerial imagery to produce training samples for rooftop PV systems, which can then be used in a proof-of-concept application to detect buildings equipped with PV systems. Using this approach, 100,000 rooftop samples were automatically generated, from which a automatically filtered selection of 51,000 samples was used to train the

final RetinaNet-based network. In a proof-of-concept application, the model achieved a precision of 92.77 % and a recall of 84.47 % in an independent test area, thereby demonstrating that large-scale, reproducible training data can be created with minimal manual effort. The study thus shows that this reproducible method can substantially reduce the effort required for manual labeling while enabling scalable analysis across regions. Moreover, the final model proved capable of detecting PV systems in previously unseen imagery, thereby supporting the validity of the first hypothesis.

**Hypothesis 2:** *Deep learning-based segmentation methods can enable precise and robust identification of solar photovoltaic systems across multiple spatial resolutions and can be transferred to new geographic contexts without region-specific retraining.*

Chapter 3 presents a deep learning-based segmentation model based on the DeepLabV3 architecture, trained on a uniquely broad combination of satellite, aerial and UAV imagery with spatial resolutions ranging from 0.1, 0.2, 0.3, 0.8, 1.6, to 3.2 meters. Through extensive hyperparameter tuning, the study identified an optimal configuration (BCE as loss, Adam as optimizer, and a learning rate of 0.0001) and demonstrated that a multi-resolution training strategy consistently outperforms single-resolution networks. The resulting model achieved high accuracy across all tested datasets, enabling precise results of both rooftop and utility-scale PV systems as measured by the F1-Score (95.27 %) and IoU (91.04 %). While performance decreases at very coarse resolutions (e.g., 3.2 meters per pixel), the multi-resolution approach substantially improves robustness and generalizability compared to conventional single-source models.

Based on these results, the final model was implemented in a practical application workflow in Chapter 4. A test application focused on large-scale PV systems in the Piat region of Algeria. It is important to note that no additional training was provided and the model had not previously been exposed to image data from Algeria or any other African country. Nevertheless, the approach achieved high accuracy with mean relative error of 8.5 %, demonstrating that the methods are both transferable and precise

in different geographical contexts. The application has been tested through integration into the open-source environment QGIS and Deepness, allowing even non-experts without programming knowledge to use the workflow, which improves accessibility and supports broader applicability in research and planning contexts.

These results combined support the conclusion that deep learning-based segmentation methods trained on diverse image sources can enable precise and robust identification of PV systems across multiple spatial resolutions and can be transferred to new geographic contexts without region-specific retraining.

**Hypothesis 3:** *Object detection methods can systematically reduce positional errors in renewable energy system datasets, thereby enhancing their spatial accuracy and overall reliability for energy system research.*

This hypothesis is first examined in Chapter 5, which introduces a object detection-based approach for correcting wind turbine coordinates in Germany. By combining location data from the official data register with high-resolution aerial imagery and applying the RetinaNet architecture, the study demonstrates how positional errors in existing datasets can be systematically reduced. A nationwide application to onshore wind turbines highlights the robustness of the method. It detected 27,130 turbines within 100 m of the provided coordinates, verified that 74% of coordinates are correct within 10 m radius, and corrected more than 15% of turbine locations with deviations between 10 and 100 m. By correcting coordinates in these large administrative registers, this approach provides more reliable spatial information, which in turn strengthens the basis for further applications. Chapter 6 transfers the approach to South Africa, where more than 1,400 wind turbines were processed and validated using OpenStreetMap location data and satellite imagery. The successful application in a completely different data and regional context highlights the robustness and transferability of the method, demonstrating its potential to improve spatial datasets beyond a single country or register.

Together, these studies strongly support the conclusion that object detec-

tion methods can systematically reduce positional errors in renewable energy system datasets, thereby enhancing their spatial accuracy and reliability for further applications.

## 7.2 Contribution to the Research Field

The contributions of this dissertation can be summarized on three complementary levels: conceptual, methodological, and applied. Combined, they illustrate how the presented dissertation advances renewable energy system research through improved data quality, reproducible workflows, and openly available datasets, which are essential prerequisites for reliable capacity assessments, infrastructure planning or scenario modelling. The focus here lies on the integration of methodological advances into the empirical and applied context of energy research.

This dissertation makes several conceptual contributions by integrating remote sensing, deep learning, and energy-system research into a coherent framework. Hypothesis 1 (Chapter 2) establishes scalable training-data generation; Hypothesis 2 (Chapter 3 and 4) demonstrates segmentation robustness across spatial resolutions and geographic contexts; and Hypothesis 3 (Chapter 5 and 6) enables systematic positional correction—together yielding an end-to-end pipeline for reliable infrastructure datasets.

On the methodological level, the dissertation advances three innovations that map directly onto the central hypotheses. In line with Hypothesis 1, it establishes scalable, semi-automatic generation of training data for deep learning models (Chapter 2). Under Hypothesis 2, it develops and validates a multi-resolution segmentation strategy that is robust across spatial resolutions and geographic contexts, and implements this capability in open-source environments (Chapters 3 and 4). Consistent with Hypothesis 3, it introduces and transfers an object detection based workflow for systematic geo-coordinate correction of wind turbines, improving the spatial accuracy of administrative and community-sourced datasets (Chapters 5 and 6).

In applied terms, the dissertation provides openly available tools and datasets that strengthen the empirical basis of energy system research. In

line with Hypothesis 2, Chapters 3 and 4 release a segmentation network for photovoltaic systems and validated open datasets for large-scale PV installations in Algeria's Piat region, enabling transfer across resolutions and regions. Consistent with Hypothesis 3, Chapter 6 delivers a comprehensive wind-turbine dataset for South Africa with corrected coordinates and enriched metadata (turbine type, commissioning year, capacity), improving spatial reliability for downstream analyses.

### 7.3 Innovations in Remote Sensing and Deep Learning

The following section highlights the most significant innovations of this dissertation, focusing on methodological and technical advances in remote sensing and deep learning. Independent of their application for energy system research, these innovations expand the state of the art in automated data generation, image segmentation, and coordinate correction. Several contributions are particularly noteworthy.

First, in line with Hypothesis 1, a workflow was established for the semi-automatic generation of training samples in high-resolution aerial images (Chapter 2). The creation of large, labeled datasets has long been one of the main bottlenecks in applying deep learning to remote sensing, as manual annotation is highly time-consuming, costly, and often inconsistent. The proposed approach addresses this challenge by automatically combining address-based PV location data with georeferenced building footprints and aerial imagery. This integration enables the generation of large numbers of labeled samples without manual inspection, while subsequent post-filtering further improves data quality. In a proof-of-concept application, the workflow produced reliable training samples that supported the training of a RetinaNet-based detection model with high precision and recall. By reducing the need for exhaustive manual labeling, the approach accelerates dataset preparation, ensures reproducibility, and provides a scalable method for future remote sensing application.

Secondly, consistent with Hypothesis 2, a multi-resolution imagery strategy for heterogeneous data was introduced using aerial, satellite, and UAV images with spatial resolutions ranging from 10 to 320 cm per pixel (Chapters 3 and 4). Traditional segmentation approaches in remote sensing typically rely on single sensor types or fixed resolutions, limiting their robustness and generalizability. The proposed method systematically combines data from different sources and resolutions into a unified training strategy, thereby enhancing the adaptability and stability of segmentation models. Extensive hyperparameter tuning identified optimal configurations for loss functions, optimizers, and learning rates, while validation across multiple datasets showed that multi-resolution training consistently outperforms single-resolution models. This contribution presents a methodological improvement by demonstrating that utilizing data diversity rather than enforcing homogeneity can lead to more robust and transferable segmentation models.

Third, following Hypothesis 3, a object detection-based method for correcting positional errors in spatial infrastructure datasets was developed (Chapters 5 and 6). Public datasets such as administrative registers like the Energy Market Data Register or community-driven resources like OpenStreetMap data often contain substantial locational inaccuracies. The proposed approach combines high-resolution aerial and satellite imagery with a RetinaNet-based object detection model to systematically reduce these errors. A nationwide application in Germany demonstrated that positional deviations of wind turbine coordinates could be corrected at scale, while a transfer study in South Africa further validated the robustness and adaptability of the approach. By improving spatial accuracy in existing datasets rather than attempting full object detection from scratch, the method offers a scalable and efficient pathway to enhance spatial data quality across regions.

## 7.4 Practical Applications

The developed methods demonstrate practical value both within the context of energy system research and in the broader field of remote sensing and deep learning. On the one hand, they provide concrete improvements for empirical energy research by enhancing the accuracy and availability of infrastructure data. On the other hand, they illustrate how deep learning workflows can be integrated into user-friendly environments such as QGIS, thereby facilitating wider adoption of such methods for spatial applications.

In practice, the segmentation model and corresponding code associated with Hypothesis 2 from Chapter 3 were openly released and subsequently applied in Chapter 4, where large-scale PV systems in Algeria were segmented with high accuracy despite the absence of region-specific training data. By integrating the trained network into QGIS via the Deepness plugin, the workflow becomes accessible even for users without programming knowledge. This lowers the entry barrier for planners and researchers in the energy domain while also providing a transferable blueprint for applying segmentation-based deep learning methods in other remote sensing applications.

Object detection-based coordinate correction, as demonstrated in Chapters 5 and 6 and relating to Hypothesis 3, has proven effective in improving the quality of administrative and community-based wind turbine datasets. In Germany, the Core Energy Market Data Register dataset was improved by correcting position errors for a significant portion of the turbines, with approximately one quarter of the records showing substantial improvements. In South Africa, the method was successfully transferred to the OpenStreetMap-based wind turbine dataset, resulting in a validated dataset with clear relevance for site analysis as well as grid planning. At the same time, this case demonstrates the methodological flexibility of object detection-based correction approaches, which can be applied to other types of infrastructure datasets beyond the energy domain.

The practical usability of the dissertation's outputs is further enhanced by the open dissemination of the resulting datasets. Validated data on PV systems in Algeria (Chapter 4) and wind turbines in South Africa (Chapter 6)

were published on Zenodo in accordance with FAIR principles, ensuring their long-term reusability. Beyond their immediate value for energy system research, the datasets also represent open, annotated benchmarks for future studies in computer vision and remote sensing.

In summary, this dissertation provides three tangible outputs that directly benefit both research and practice:

1. A user-friendly PV segmentation model was implemented in QGIS via the Deepness plugin, lowering the entry barrier for non-programmers.
2. An openly available dataset of large-scale PV systems in Piat, Algeria was generated and validated, providing a valuable resource for further research.
3. A corrected and enriched dataset of wind turbines in South Africa was validated and openly published, supporting further research and applications.

## 7.5 Limitations and Challenges

Despite the methodological advances and practical applications demonstrated in this dissertation, it is important to address several limitations and challenges. These are closely tied to the three central hypotheses and concern data availability, bias and transferability, as well as the scope of applicability of the developed methods.

A first limitation arises from the data foundations of the semi-automatic workflow, which define the scope of Hypothesis 1. The approach in Chapter 2 relies on combining registry data with official building footprints, and such datasets are available in this form only in a few countries, such as Germany. While this combination enables large-scale, reproducible training-data generation for rooftop PV systems, it is not directly transferable to other renewable energy technologies like wind turbines. Furthermore, the reference data sets can vary significantly in terms of their origin and quality. While the

data register is based on mandatory self-reported information from operators, building areas are officially maintained.

A main limitation to Hypothesis 2 concerns bias and generalizability. The training data in Chapter 3 were compiled from several previous European and Asian datasets, but contain no examples from Africa, America, or Australia. While PV systems are globally standardized, contextual differences such as building footprints, roof types, and building materials can complicate segmentation. The model generalized well to large-scale systems in Algeria (Chapter 4), but its applicability in other contexts, for example to rooftop PV systems in regions with very different urban morphologies, remains unclear. Furthermore, the multi-resolution approach mitigates dataset biases, but does not remove them completely. Performance drops at the aggregated coarsest resolution of 3.2 meters, which represents the upper end of our scale range. The information loss from aggregation leads to notably lower scores than at finer resolutions.

In addition, the validation strategy in Chapter 3 relied on a distribution of PV and non-PV pixels that does not reflect real-world conditions, potentially leading to an overestimation of model performance. Further limitations arise from differences in data coverage and resolution between countries. For example, Chapter 4 relied on aggregated Google satellite imagery for Algeria, which does not provide native 0.2 meters data. Finally, although the integration into QGIS (Chapter 4) allows practical use of trained models even on standard desktop computers, the amount of data that can be processed in such environments is naturally limited, constraining scalability for large-area applications.

One major limitation of Hypothesis 3 relates to the dependence on existing datasets and image availability. The object detection-based correction methods introduced in Chapter 5 and 6 depend on the quality of existing reference datasets. In Germany, the Core Energy Market Data Register is based on operator self-reporting, while in South Africa OpenStreetMap is compiled by volunteers. Both data sources are subject to inconsistencies, incompleteness, and systematic biases. Another limitation concerns coverage and applicability. The object detection models improve positional accuracy

for turbines already present in the datasets, but they cannot identify turbines missing from the input data. Their application is further limited by the availability and quality of the underlying imagery, as the method requires that turbine structures are clearly visible. While the transfer of the method to the South African OpenStreetMap dataset shows that it can be adapted to different contexts, the completeness and accuracy of OpenStreetMap contributions vary widely across regions. Incorporating heterogeneous additional datasets may further lead to issues with inconsistent metadata and annotation standards.

## 7.6 Outlook and Future Research

Based on the results of this dissertation, several directions for future research can be identified. These directions involve methodological advances in remote sensing and deep learning, the expansion and harmonization of data sources as well as the integration of the developed methods into broader application contexts.

### 7.6.1 Methodological Perspective

One initial field for future research relates to methodological improvements in deep learning techniques. A promising approach involves the combination of object detection and segmentation methods within integrated workflows. Object detection-based models are well suited for identifying as well as regression of objects in large datasets, while subsequent segmentation enables a detailed assessment of the scope and characteristics of observed data. These approaches can be applied sequentially in cascades or potentially integrated into single multitask networks, thereby increasing both accuracy and performance.

Further progress is expected with the introduction of newer model architectures. Vision Transformer (ViT)-based and hybrid models combining convolutional and transformer components are particularly promising for achieving higher accuracy and robustness. In addition, multi-modal approaches

that integrate image data with complementary sources such as textual, contextual or spatial information open new perspectives for the joint analysis of heterogeneous datasets. However, the computational effort must always be taken into account. Finally, efficient lightweight architectures offer potential opportunities for applying deep learning methods at large scale, facing limited computational resources, thereby supporting broader applicability in the remote sensing and deep learning community.

In addition, future research could explore self-supervised learning methods, enabling the use of large amounts of unannotated remote sensing images. Compared to semi-supervised sampling strategies, these approaches offer a systematic method for learning robust feature representations without relying on costly annotated datasets, thereby addressing one of the key bottlenecks in applying deep learning to geospatial domains. Building on this, the increasing emergence of foundation models and large-scale pre-trained networks for deep learning and remote sensing, most of which are trained in a self-supervised manner, offers the opportunity to reduce the amount of task-specific training data required and improve transferability between regions and tasks. The use of such models could significantly lower dependence on region-specific datasets while improving performance in underrepresented contexts.

### **7.6.2 Data Enrichment and Practical Applications**

Future work should aim to further expand and diversify the data basis in order to reduce regional biases and strengthen the robustness of remote sensing workflows. This includes systematically extending training datasets to underrepresented world regions such as South America, Australia, or Oceania and incorporating complementary earth observation programs like Copernicus Sentinel combined with high resolution constellations like PlanetScope or UAV datasets to improve both spatial and temporal coverage. In parallel, the harmonization of reference datasets such as the German Core Energy Market Data Register, OpenStreetMap, or other national registers through standardized metadata and annotation practices will enhance comparability

and reusability across regions. Beyond these technical aspects, the integration of socio-economic and ecological datasets offers the opportunity to support more holistic approaches in energy and spatial planning.

In addition to data-related advances, further research should emphasize practical applications and integration into operational contexts. This includes the expansion of open tools such as QGIS plugins and interactive platforms to make the developed methods broadly accessible to diverse user groups. For wind energy, this also implies developing open validation tools that facilitate transparent quality checks of turbine datasets. Moreover, embedding the proposed methods into automated monitoring systems could enable continuous tracking of not only the existing amount of renewable energy systems but also their expansion as well as decommissioning, thereby enabling near real-time insights for further planning and monitoring.

### 7.6.3 Long-term Perspectives

In the long term, several overarching perspectives emerge from this dissertation. A central aspect concerns the sustainability of open science practices. Ensuring the long-term accessibility and usability of open datasets and workflows requires a consistent orientation toward FAIR and CARE principles as well as their embedding into international research infrastructures. At the same time, advances in cloud infrastructures and distributed computing open up possibilities for near real-time processing and large-scale monitoring of renewable energy systems.

Another important perspective lies in cross-domain integration. Remote sensing-based infrastructure data can be linked with socio-economic, ecological, and climate models to support systemic analyses and integrated planning. In addition, the methods and datasets developed in this work can be embedded into digital twins of energy systems and urban infrastructures, enabling interactive scenario exploration, multi-criteria assessments, and decision support.

Finally, this research contributes to supporting the global energy transition. By promoting globally standardized, open, and scalable workflows, it

provides the basis for improved comparability across regions and a more reliable empirical basis for international research in the field of energy systems.

## 7.7 Overall Conclusions

This dissertation examines how remote sensing imagery combined with deep learning methods can improve the recognition, correction, and analysis of renewable energy systems. The work is guided by three central hypotheses formulated, which are explicitly addressed and supported. So the dissertation directly answers the overarching research questions:

1. How can training data be generated efficiently and reproducibly? By combining registry data with building geometries, reproducible training datasets can be generated at scale, reducing the dependence on manual annotation.
2. How robust are segmentation approaches across different resolutions and contexts? Multi-resolution segmentation significantly improves robustness and transferability, enabling reliable segmentation of PV systems across diverse regions and scales.
3. How can the positional accuracy of existing datasets be improved? Object detection-based models can systematically correct and enrich infrastructure datasets, achieving substantial accuracy improvements for both administrative and community-based data sources.

In summary, the dissertation makes several contributions. For the remote sensing and deep learning community, it provides methodological innovations in training data generation, multi-resolution segmentation, and object detection-based coordinate correction, supported by openly released models and code. For the energy system research community, it delivers validated datasets, thereby strengthening the empirical basis for energy research. In doing so, the dissertation lays a methodological and empirical foundation for future research and applications that connect remote sensing, deep learning, and renewable energy system research.

## Bibliography

- [1] Ibrahim Dincer. Renewable energy and sustainable development: a crucial review. *Renewable and sustainable energy reviews*, 4(2):157–175, 2000.
- [2] Abdelrahman Azzuni and Christian Breyer. Definitions and dimensions of energy security: a literature review. *Wiley interdisciplinary reviews: Energy and environment*, 7(1):e268, 2018.
- [3] Patrick Devine-Wright and Adam Peacock. Putting energy infrastructure into place: A systematic review. *Renewable and Sustainable Energy Reviews*, 197:114272, 2024.
- [4] Jann Michael Weinand, Fabian Scheller, and Russell McKenna. Reviewing energy system modelling of decentralized energy autonomy. *Energy*, 203:117817, 2020.
- [5] Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, et al. The era5 global reanalysis. *Quarterly journal of the royal meteorological society*, 146(730):1999–2049, 2020.
- [6] Raja Pandi Perumal, Holger Voos, Florio Dalla Vedova, and Hubert Moser. Small satellite reliability: A decade in review. 2021.
- [7] Erik Kulu. Satellite constellations—2024 survey, trends and economic sustainability. In *Proceedings of the International Astronautical Congress, IAC, Milan, Italy*, pages 14–18, 2024.
- [8] Qiang Zhao, Le Yu, Zhenrong Du, Dailiang Peng, Pengyu Hao, Yongguang Zhang, and Peng Gong. An overview of the applications of earth observation satellite data: impacts and future trends. *Remote Sensing*, 14(8):1863, 2022.
- [9] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436–444, 2015.

- 
- [10] Xiao Xiang Zhu, Devis Tuia, Lichao Mou, Gui-Song Xia, Liangpei Zhang, Feng Xu, and Friedrich Fraundorfer. Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine*, 5(4):8–36, 2017.
- [11] Simiao Ren, Wayne Hu, Kyle Bradbury, Dylan Harrison-Atlas, Laura Malaguzzi Valeri, Brian Murray, and Jordan M Malof. Automated extraction of energy systems information from remotely sensed data: A review and analysis. *Applied Energy*, 326:119876, 2022.
- [12] James B Campbell and Randolph H Wynne. *Introduction to remote sensing*. Guilford press, 2011.
- [13] Arthur P Cracknell. *Advanced very high resolution radiometer AVHRR*. Crc Press, 1997.
- [14] Floyd M Henderson and Anthony J Lewis. *Principles and applications of imaging radar. Manual of remote sensing: Volume 2*. John Wiley and Sons, Inc., Somerset, NJ (United States), 1998.
- [15] Alexander FH Goetz, Gregg Vane, Jerry E Solomon, and Barrett N Rock. Imaging spectrometry for earth remote sensing. *science*, 228(4704):1147–1153, 1985.
- [16] John R Jensen and Kalmesh Lulla. *Introductory digital image processing: a remote sensing perspective*. Taylor & Francis, 1987.
- [17] Matthew F. McCabe, Matthew Rodell, Douglas E. Alsdorf, Diego G. Miralles, Remko Uijlenhoet, Wolfgang Wagner, Arko Lucieer, Rasmus Houborg, Niko E.C. Verhoest, Trenton E. Franz, Jiancheng Shi, Huilin Gao, and Eric F. Wood. The future of earth observation in hydrology. *Hydrology and Earth System Sciences*, 21(7):3879–3914, 2017.
- [18] Thomas Lillesand, Ralph W Kiefer, and Jonathan Chipman. *Remote sensing and image interpretation*. John Wiley & Sons, 2015.

- 
- [19] Josef Aschbacher and Maria Pilar Milagro-Pérez. The european earth monitoring (gmes) programme: Status and perspectives. *Remote Sensing of Environment*, 120:3–8, 2012.
- [20] Andrew J Tatem, Scott J Goetz, and Simon I Hay. Fifty years of earth observation satellites: Views from above have lead to countless advances on the ground in both scientific knowledge and daily life. *American scientist*, 96(5):390, 2008.
- [21] Michael A. Wulder, Joanne C. White, Samuel N. Goward, Jeffrey G. Masek, James R. Irons, Martin Herold, Warren B. Cohen, Thomas R. Loveland, and Curtis E. Woodcock. Landsat continuity: Issues and opportunities for land cover monitoring. *Remote Sensing of Environment*, 112(3):955–969, 2008.
- [22] Michael A. Wulder, David P. Roy, Volker C. Radeloff, Thomas R. Loveland, Martha C. Anderson, David M. Johnson, Sean Healey, Zhe Zhu, Theodore A. Scambos, Nima Pahlevan, Matthew Hansen, Noel Gorelick, Christopher J. Crawford, Jeffrey G. Masek, Txomin Hermosilla, Joanne C. White, Alan S. Belward, Crystal Schaaf, Curtis E. Woodcock, Justin L. Huntington, Leo Lymburner, Patrick Hostert, Feng Gao, Alexei Lyapustin, Jean-Francois Pekel, Peter Strobl, and Bruce D. Cook. Fifty years of landsat science and impacts. *Remote Sensing of Environment*, 280:113195, 2022.
- [23] Michael Berger, Jose Moreno, Johnny A Johannessen, Pieternel F Lev-elt, and Ramon F Hanssen. Esa’s sentinel missions in support of earth system science. *Remote Sensing of Environment*, 120:84–90, 2012.
- [24] Zbyněk Malenovský, Helmut Rott, Josef Cihlar, Michael E. Schaepman, Glenda García-Santos, Richard Fernandes, and Michael Berger. Sentinels for science: Potential of sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land. *Remote Sensing of Environment*, 120:91–101, 2012. The Sentinel Missions - New Opportunities for Science.

- 
- [25] Susan L. Ustin and Elizabeth McPhee Middleton. Current and near-term earth-observing environmental satellites, their missions, characteristics, instruments, and applications. *Sensors*, 24(11), 2024.
- [26] Amy E Frazier and Benjamin L Hemingway. A technical review of planet smallsat data: Practical considerations for processing and using planetscope imagery. *Remote Sensing*, 13(19):3930, 2021.
- [27] Gérard Brachet. Spot: the first operational remote sensing satellite. In *Satellite Remote Sensing for Resources Development*, pages 59–80. Springer, 1986.
- [28] Simon J Cantrell, Aparajithan Sampath, James C Vrabel, Paul Bresnahan, Cody Anderson, Minsu Kim, and Seonkyung Park. System characterization report on the pléiades neo imager. Technical report, US Geological Survey, 2023.
- [29] European Union. Copernicus data space ecosystem. <https://dataspace.copernicus.eu/>, 2025. Accessed: 2025-09-21.
- [30] Microsoft. Microsoft planetary computer. <https://planetarycomputer.microsoft.com/>, 2025. Accessed: 2025-09-21.
- [31] Google. Google earth engine. <https://earthengine.google.com/>, 2025. Accessed: 2025-09-21.
- [32] Yann LeCun, Bernhard Boser, John S Denker, Donnie Henderson, Richard E Howard, Wayne Hubbard, and Lawrence D Jackel. Back-propagation applied to handwritten zip code recognition. *Neural computation*, 1(4):541–551, 1989.
- [33] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 2012.

- 
- [34] Yanming Guo, Yu Liu, Ard Oerlemans, Songyang Lao, Song Wu, and Michael S Lew. Deep learning for visual understanding: A review. *Neurocomputing*, 187:27–48, 2016.
- [35] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [36] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [37] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 580–587, 2014.
- [38] Ross Girshick. Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 1440–1448, 2015.
- [39] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28, 2015.
- [40] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988, 2017.
- [41] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- [42] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected

- crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4):834–848, 2017.
- [43] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969, 2017.
- [44] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- [45] Rishi Bommasani. On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*, 2021.
- [46] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C Berg, Wan-Yen Lo, et al. Segment anything. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 4015–4026, 2023.
- [47] Jesper E Van Engelen and Holger H Hoos. A survey on semi-supervised learning. *Machine learning*, 109(2):373–440, 2020.
- [48] Longlong Jing and Yingli Tian. Self-supervised visual feature learning with deep neural networks: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 43(11):4037–4058, 2020.
- [49] Jean-Bastien Grill, Florian Strub, Florent Altché, Corentin Tallec, Pierre Richemond, Elena Buchatskaya, Carl Doersch, Bernardo Avila Pires, Zhaohan Guo, Mohammad Gheshlaghi Azar, et al. Bootstrap your own latent—a new approach to self-supervised learning. *Advances in neural information processing systems*, 33:21271–21284, 2020.
- [50] Mark Everingham, Luc Van Gool, Christopher KI Williams, John Winn, and Andrew Zisserman. The pascal visual object classes (voc) challenge. *International journal of computer vision*, 88(2):303–338, 2010.

- 
- [51] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer, 2014.
- [52] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1):1929–1958, 2014.
- [53] Anders Krogh and John Hertz. A simple weight decay can improve generalization. *Advances in neural information processing systems*, 4, 1991.
- [54] Lutz Prechelt. Early stopping-but when? In *Neural Networks: Tricks of the trade*, pages 55–69. Springer, 2002.
- [55] Antonio Torralba and Alexei A Efros. Unbiased look at dataset bias. In *CVPR 2011*, pages 1521–1528. IEEE, 2011.
- [56] Siddharth Samsi, Dan Zhao, Joseph McDonald, Baolin Li, Adam Michaleas, Michael Jones, William Bergeron, Jeremy Kepner, Devesh Tiwari, and Vijay Gadepally. From words to watts: Benchmarking the energy costs of large language model inference. In *2023 IEEE High Performance Extreme Computing Conference (HPEC)*, pages 1–9. IEEE, 2023.
- [57] Jordan M Malof, Rui Hou, Leslie M Collins, Kyle Bradbury, and Richard Newell. Automatic solar photovoltaic panel detection in satellite imagery. In *2015 International Conference on Renewable Energy Research and Applications (ICRERA)*, pages 1428–1431. IEEE, 2015.
- [58] Jordan M Malof, Kyle Bradbury, Leslie M Collins, and Richard G Newell. Automatic detection of solar photovoltaic arrays in high resolution aerial imagery. *Applied energy*, 183:229–240, 2016.

- 
- [59] Jiafan Yu, Zhecheng Wang, Arun Majumdar, and Ram Rajagopal. Deep-solar: A machine learning framework to efficiently construct a solar deployment database in the united states. *Joule*, 2(12):2605–2617, 2018.
- [60] Kevin Mayer, Zhecheng Wang, Marie-Louise Arlt, Dirk Neumann, and Ram Rajagopal. Deepsolar for germany: A deep learning framework for pv system mapping from aerial imagery. In *2020 International Conference on Smart Energy Systems and Technologies (SEST)*, pages 1–6. IEEE, 2020.
- [61] Lucas Kruitwagen, KT Story, J Friedrich, L Byers, S Skillman, and Cameron Hepburn. A global inventory of photovoltaic solar energy generating units. *Nature*, 598(7882):604–610, 2021.
- [62] Hou Jiang, Ling Yao, Ning Lu, Jun Qin, Tang Liu, Yujun Liu, and Chenghu Zhou. Multi-resolution dataset for photovoltaic panel segmentation from satellite and aerial imagery. *Earth System Science Data Discussions*, 2021:1–17, 2021.
- [63] Zhiling Guo, Zhan Zhuang, Hongjun Tan, Zhengguang Liu, Peiran Li, Zhengyuan Lin, Wen-Long Shang, Haoran Zhang, and Jinyue Yan. Accurate and generalizable photovoltaic panel segmentation using deep learning for imbalanced datasets. *Renewable Energy*, 219:119471, 2023.
- [64] Jorge Felipe Gaviria, Gabriel Narváez, Camilo Guillen, Luis Felipe Giraldo, and Michael Bressan. Machine learning in photovoltaic systems: A review. *Renewable Energy*, 196:298–318, 2022.
- [65] Hongzhi Mao, Xie Chen, Yongqiang Luo, Jie Deng, Zhiyong Tian, Jinghua Yu, Yimin Xiao, and Jianhua Fan. Advances and prospects on estimating solar photovoltaic installation capacity and potential based on satellite and aerial images. *Renewable and Sustainable Energy Reviews*, 179:113276, 2023.
- [66] Jingbo Chen, Anzhi Yue, Chengyi Wang, Qingqing Huang, Jiansheng Chen, Yu Meng, and Dongxu He. Wind turbine extraction from high

- spatial resolution remote sensing images based on saliency detection. *Journal of Applied Remote Sensing*, 12(1):016041–016041, 2018.
- [67] Fereshteh Abedini, Mahdi Bahaghighat, and Misak S’hoyan. Wind turbine tower detection using feature descriptors and deep learning. *Facta Universitatis, Series: Electronics and Energetics*, 33(1):133–153, 2019.
- [68] Ting Zhang, Bo Tian, Dhritiraj Sengupta, Lei Zhang, and Yali Si. Global offshore wind turbine dataset. *Scientific Data*, 8(1):191, 2021.
- [69] Thorsten Hoeser, Stefanie Feuerstein, and Claudia Kuenzer. Deepowt: A global offshore wind turbine data set derived with deep learning from sentinel-1 data. *Earth System Science Data Discussions*, 2022:1–26, 2022.
- [70] Thorsten Hoeser and Claudia Kuenzer. Synteo: Synthetic dataset generation for earth observation and deep learning—demonstrated for offshore wind farm detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 189:163–184, 2022.
- [71] Tingting He, Yihua Hu, Fashuai Li, Yuwei Chen, Maoxin Zhang, Qiming Zheng, Yukan Jin, and He Ren. Mapping land-and offshore-based wind turbines in china in 2023 with sentinel-2 satellite data. *Renewable and Sustainable Energy Reviews*, 214:115566, 2025.
- [72] Caleb Robinson, Anthony Ortiz, Allen Kim, Rahul Dodhia, Andrew Zolli, Shivaprakash K Nagaraju, James Oakleaf, Joe Kiesecker, and Juan M Lavista Ferres. Global renewables watch: A temporal dataset of solar and wind energy derived from satellite imagery. *arXiv preprint arXiv:2503.14860*, 2025.

# Chapter 8

## Appendix

### 8.1 Curriculum Vitæ

Maximilian Kleebauer

---

#### Personal

|             |   |
|-------------|---|
| Occupation  | Research Associate at Fraunhofer IEE & University of Kassel |
| Nationality | German  |
| Residence   | Marburg   |
| Contact     | max_kleebauer@gmx.de  |

#### Education

|             |  |
|-------------|--|
| 2006 – 2009 | Professional education as an IT systems electronics technician at Siemens AG in Erlangen with degree: IT systems electronics technician specializing in system integration |
| 2012 – 2015 | Studied Physics (B.Sc.) at the Philipps University of Marburg  |

- 
- 2015 – 2017 Studied Physical Geography (B.Sc.) at the Philipps University of Marburg with degree B.Sc.
- 2017 – 2020 Studied Physical Geography with a specialization in Environmental Information Systems (M.Sc.) at the Philipps University of Marburg, graduating with degree M.Sc.

### **Work experience**

- 2009 - 2010 Technical support at Siemens AG - Industry Solution Division, Business Unit Metals Technologies with tasks installation of workstations for the field service, hardware and software support, repair processing, inventory issue, order processing, delivery tracking
- 2010 – 2011 Civil service at the Education Center for the Blind and Visually Impaired in Nuremberg with tasks of support and assistance in educational contexts
- 2013 – 2019 Student assistant in the Department of Studies and Teaching at the Philipps University of Marburg with tasks of administration, installation and support of accessibility, screen workstations for the blind and visually impaired
- 2018 – 2019 Independent contractor in the field of Geodata Analysis with tasks of accompanying construction planning, calculation and presentation of various influencing factors due to planned construction project
- 2019 Independent contractor in the research project Loewe - Nature 4.0 with tasks construction, calibration and maintenance of sensor units for the collection of biotic and abiotic environmental factors as well as data storage and quality control of the collected data

- Since 2021 Research associate in the field of Energy Meteorology and Geoinformation Systems at Fraunhofer Institute for Energy Economics and Energy System Technology, Kassel, with responsibility for industry project management, acquisition of third-party funding, and research on geoinformation systems, 3D city modelling, remote sensing, and machine learning methods
- Since 2022 Research associate at the Department of Energy Management and Operation of Electrical Grids at the University of Kassel, with responsibility for Horizon Europe project management, acquisition of research funding, and research on geoinformation systems, remote sensing, and machine learning methods
- Since 2024 PhD student at Philipps University of Marburg working in the field of Physical Geography with a focus on remote sensing and deep learning for renewable energy system recognition

## 8.2 Publications

- 2017 Bachelor thesis "Spatial modelling of landslides in the Franconian Alb and identification and transferability of their influencing factors".
- 2020 Master's thesis "Development of a method for detecting photovoltaic systems in high-resolution aerial images."
- 2021 Kleebauer, M., Horst, D., Reudenbach, C.: Semi-automatic generation of training samples for detecting renewable energy plants in high-resolution aerial images. *Remote Sensing*, 13(23), 4793.
- 2023 Kleebauer, M., Marz, C., Reudenbach, C., Braun, M.: Multi-resolution segmentation of solar photovoltaic systems using deep learning. *Remote Sensing*, 15(24), 5687.

- 2024 Kleebauer, M., Braun, A., Horst, D., Pape, C.: Enhancing wind turbine location accuracy: A deep learning-based object regression approach for validating wind turbine geo-coordinates. In *IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium*, IEEE.
- 2025 Botha, N., Coleman, T., Wessels, G., Kleebauer, M., Karamanski, S.: Power generation time series for solar energy generation: Modelling with ATlite in South Africa. *Solar*, 5(1), 8.
- 2025 Kleebauer, M., Hafdaoui, H., Bouchakour, S., Hückner, B., Lindenmeyer, M.: Globally scalable, QGIS-integrated workflow for solar photovoltaic system segmentation and capacity estimation: A case study in Algeria. In *IGARSS 2025 - 2025 IEEE International Geoscience and Remote Sensing Symposium*, Brisbane, Australia, IEEE.
- 2025 Kleebauer, M., Karamanski, S., Callies, D., Braun, M.: A Wind Turbines Dataset for South Africa: OpenStreetMap Data, Deep Learning Based Geo-Coordinate Correction and Capacity Analysis. *ISPRS International Journal of Geo-Information*, 14(6), 232

## 8.3 Eidesstattliche Erklärung

### ERKLÄRUNG

Ich erkläre an Eides statt, dass ich meine Dissertation

*Remote Sensing based Renewable Energy System Recognition using Deep Learning*

selbstständig und ohne unerlaubte Hilfe angefertigt und mich dabei keinerlei anderer als der von mir ausdrücklich bezeichneten Quellen und Hilfen bedient habe.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch an keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, 08.10.2025

.....

Maximilian Kleebauer