From Physical to Virtual: Leveraging Drone Imagery to Automate Photovoltaic System Maintenance

Maximilian Lowin[1], Domenic Kellner[1], Tobias Kohl[1], Cristina Mihale-Wilson[1]

Abstract: Optimizing the maintenance of large-scale infrastructure can be a significant cost driver for small and medium-sized enterprises (SMEs). This paper presents a feasible approach to combine data from real-world physical structures collected through an automated maintenance process with cloud-based AI services to generate a meaningful virtual representation of such structures. We use photovoltaic systems as an exemplary physical structure and thermal imaging, collected through scheduled drone monitoring. With help of these unstructured data sources, we demonstrate our approach's applicability. Our solution artifact provides a lightweight AI application that is adoptable for other problem spaces, enabling an easier knowledge transfer from research to SMEs. By combining Cloud Computing with Machine Learning, the artifact identifies present and emerging damages of physical objects. It provides a virtual representation of the object's state and empowers a meaningful visualization.

Keywords: Digital Twin; Machine Learning; Predictive Maintenance; Photovoltaic System; Internet of Things, Process Automation, Neural Networks, Visualization

1 Introduction

Ineffective maintenance is a major cost driver [Mo02]. Failures of expensive infrastructure especially influences small and medium-sized enterprises (SMEs). Hence, optimizing the timing of maintenance cycles by predicting upcoming failures and thus minimize cost as well as resources is particularly important for SMEs. One emerging Predictive Maintenance method is based on Digital Twins (DT). The DT method has therefore seen adoption in a wide variety of different industry sectors [Sh18]. DT has a data-centric view, which enables the system itself to review its own contextual information and make operational decisions based on those data-driven analytics. With the advent of Internet of Things (IoT) devices and interconnected structures such as in smart cities, Predictive Maintenance based on DT will become ubiquitous. Although DT has many advantages, one of the main reported challenges by businesses concerning the adoption of DT´s in Predictive Maintenance processes is data availability [EBA20]. This is especially true for SMEs, since they cannot rely on large physical structures and surplus resources. Often businesses are legally required to provide regular cost intensive on-site inspections for their physical assets, so that Predictive Maintenance based on DT

[1] Goethe University Frankfurt, Theodor-W.-Adorno-Platz 4, D-60629 Frankfurt am Main, Germany, {lowin, kellner, tokohl, mihale-wilson}@wiwi.uni-frankfurt.de

would be more than useful. Against the background that large-scale physical structures are not trivial to realize as virtual entities, this paper aims to utilize object detection and the DT concept to create virtual representational models for real-world physical structures. We show how to achieve this based on the example of photovoltaic (PV) systems. Specifically, we present a prototype artifact that automatically utilizes thermal drone images and Artificial Intelligence (AI) to show a complete virtualized representation of the physical structures condition.

To evaluate our artifact directly in the application context of SMEs, we tested a pilot project on-site and evaluated its validity in collaboration with solar park operators and a medium-sized technical support company. We contribute to the automation of creating digital representations of physical systems to build more sustainable business models for SMEs. While our solution artifact provides a lightweight AI application focused on the digitalization of solar parks, SMEs can adapt this approach to other problem spaces and industries i.e. allowing them to implement the functionality needed in their business context. Thus, our approach improves and optimizes current processes by saving time and costs.

We structured the remainder of the paper as follows. The next section provides an overview of the existing state of knowledge on Predictive Maintenance and Digital Twins. Afterwards, we explain our produced artifact in more detail, covering data collection, object detection, and the DT visualization. Finally, after evaluating our artifact, we conclude with implications of our work and possible future research paths.

2 Related Work and Research Context

2.1 Design Science Research Methodology

We draw on the Design Science Research (DSR) Contribution Framework presented by [GH13]. The DSR Contribution Framework allows us to frame our work and prioritize specific steps required to achieve our research goal. Second, for implementation, we draw on the Design Science Research Methodology (DSRM) by [Pe07]. The DSRM provides helpful principles, practices, and procedures consistent with prior Information Systems literature, suitable to evaluate automation processes of SMEs. It starts with identifying the problem and defining the artifact's objectives. We incorporate the existing state of knowledge for the design and development as suggested by [GH13]. Altogether, our artifact encompasses three layers: a data collection layer, a processing layer, and a visualization layer. After presenting each of the three layers, we follow [Pe07] to evaluate and demonstrate the proposed artifact's value.

2.2 Digital Twin

The DT represents the formation of a two-way relationship between a physical structure and a virtual model. The DT concept has mostly seen wide adoption in manufacturing industries as it establishes a relationship between the physical object and sometimes multiple virtual models. However, to bring the concept of DTs into various other industries a wide adoption by SMEs will be crucial.

The DT supports the creation of virtual representations allowing recording of activities throughout the whole life-cycle [Sc17]. Originally, the DT started as a Computer-Aided Design (CAD) description-based model of a product or so-called "Thing" with descriptive information added to it. However, the concept evolved to a more encompassing data-centric view by adding actionable components [GV17]. These components enable one of the DTs' core capabilities to predict an unexpected event's physical response before it occurs [Sc17]. Whereas CAD description-based models were mostly static representations, the DT allows for dynamic interactions and simulations, resulting in actionable knowledge becoming available before real physical events occur [GV17]. DTs are not just singular models but rather a set of linked data artifacts and simulation models. By being an evolving "living" model for a physical asset, the DT offers an effective way of real-time interaction and integration in a so-called cyber-physical system (CPS) [Sh18]. This actionable knowledge allows us to monitor so-called emergent behavior. Especially for automated maintenance based on drone imagery, recognizing emergent behavior is a focal point why a DT is particularly relevant [GV17].

Therefore, in our artifact, we use the DT to include all types of information about the physical structure, each of the structures' assembled units, and their operational conditions. Thereby the DT provides access to the critical information for diagnostics and helps to initiate the maintenance processes. [GV17]

2.3 Automated Maintenance Processes

Maintenance is a significant cost driver in almost all industries with tangible assets [WY07]. As advanced analytics methods like Predictive Maintenance (PdM) became popular in the late 1980s due to ineffective maintenance methods [Mo02a] they are now omnipresent in many industries. PdM uses a system's information to determine the system's condition and predicts the optimal maintenance schedule [Al14]. As such, data has become of ubiquitous significance for PdM success. While this information can originate from various sources such as IoT devices and sensors by sensing internal information like temperature, vibration, voltage, or current [Ha10] it is almost nonexistent for larger physical structures as tangible assets. Maintenance of these assets also often legally requires businesses to provide regular on-site inspections. These inspections bind many resources and are a significant cost driver.

To implement automated PdM, rich sensor data is a prerequisite that enables anomaly and potential failure detection through a combination of Machine Learning (ML) techniques and Cloud Computing. ML uses historical data to identify patterns in data and learn from them to formulate predictions and insights relevant to various stakeholders [Be20]. In practice, ML methods help to detect unusual behavior like anomalies by identifying data records that deviate from the non-anomalous distribution of the sensor data [CBK09]. To date, mostly industrial settings such as manufacturing or aviation dispose of the data necessary to optimize their PdM activities [Al14]. However, due to PdM's huge potential, transferring the concept to further applications can prove beneficial, too.

Besides pure forecasts for optimal maintenance windows, businesses are also interested in automating the maintenance process as much as possible. With current efforts focused on the automation of scheduling the maintenance process rather than automating the maintenance itself [Yü20], this work seeks to advance the PdM process automation by presenting an automated monitoring solution for PdM.

3 Artifact Description

We conduct this research in the context of PV systems. Implementing Predicted Maintenance and automated monitoring solutions on PV panels is beneficial since analysts can often predict specific failures at a certain time in advance [Be17]. Traditional maintenance approaches analyze each panel's output data, e.g., by incorporating the produced power or the present voltage at each panel [De18]. An essential downside of such traditional maintenance approaches is that it is very work-intensive, it requires modern measuring technology, and accurate information for each panel separately. Besides, it also requires expert knowledge on the technical set-up and the conversion of solar energy into electricity.

Newer maintenance approaches are more independent of the solar panel's technology and use infrared (IR) thermography to inspect solar panels. Based on their thermal characteristics [BS15], such approaches can identify faulty modules and panels. A single PV module consists of multiple cells that convert light into electricity. Faulty cells or panels heat up due to the inability to convert solar energy to power [Sp12]. Under infrared thermography, the heated spots appear like "hotspots". These so-called hotspots may occur due to cell failure, interconnection failure, partial shading, or mismatched cells [De12]. Undetected hotspots in a module may lead to a degradation of cell properties and damage the entire module. IR cameras can reveal modules and cells with anomalies [Sp12]. Because such anomaly detection does not require detailed information on the modules or advanced expert knowledge, thermography is a non-intrusive maintenance technique with high potential cost benefits.

Drones equipped with IR technology enhance the maintenance process due to their versatility. Regular on-site inspections are a legal necessity for solar park operators. Combining these inspections with drone monitoring would enable a faster and more detailed maintenance process. Drones can quickly inspect field plants and access roofs without further safeguarding inspectors [BS15]. We use both technologies and enhance them with Machine Learning to automate the process and improve its applicability and efficiency.

3.1 Concept Architecture

To present our artifact in a meaningful way, we split the description into separate layers (see Fig. 1), including the data collection, processing, and visualization layer: The **data collection layer** encompasses the initial thermal images as an unstructured data source. The **processing layer** incorporates cloud computing and ML to detect anomalies like reduced operational capacities in unstructured data. Finally, the **visualization layer** transforms the unstructured data into the virtual entity of the DT. Thus, the layer enables visualization of the asset structure as well as the condition and status of each assembled unit to provide helpful decision support.

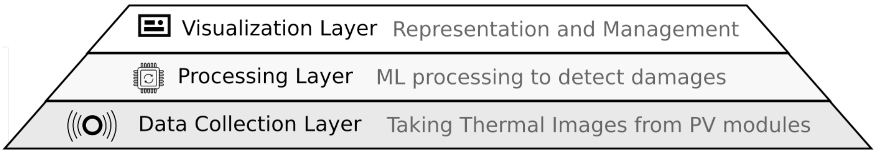

Fig. 1: Concept Architecture (Layer Structure)

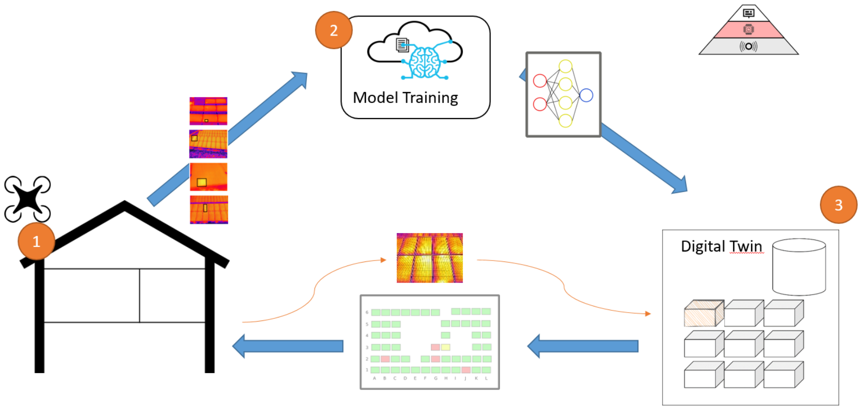

Fig. 2: Orchestration Architecture: AI-Modeling and Digital Twin Provisioning

For physical structures such as PV systems, the idea is to create a complete virtual model or DT of the respective structures. As with many areas of application, the most pressing issue is the availability of data. We gather the necessary data via an automated maintenance process based on thermal drone imagery to address this issue. Using such an external unstructured data source, we are not bound to utilize only certain PV module types or technologies. However, processing the unstructured data into a virtual entity poses various challenges: Firstly, we must incorporate the overall system's physical constructional composition and map each units' diagnostic condition on the observed model. Secondly, we have to observe and determine damages or patterns on each assembled PV unit. In the end, the visualization layer needs to highlight the assets condition based on the prognosticated health of all assembled units. Based on the thermal images, the processing layer creates a detailed prediction for each unit's operational condition and allows for actionable insights (e.g., corrective maintenance processes), see Fig. 2.

3.2 Data Collection & Processing Layer

In the **data collection layer,** we stored several thermal drone images from various solar panels. Mainly, we collected thermal images of panels produced by ‘Schott Solar’ and ‘CanadianSolar’ as well as inverters from ‘SMA’, ‘Fronius’, and ‘Emerson Control Techniques’. We collected 237 images showing up to 8 damages per picture. While some pictures show damaged cells, others display no failures or anomalies. The solar panels were mainly located in solar parks and are between 6 and 21 years old. Given the high density of PV panels, solar parks are the optimal location to build a reasonable data layer.

In the **processing layer**, we detect where damages in a thermal drone image exactly appear by employing state-of-the-art Computer Vision (CV), i.e., object detection based on pre-trained deep learning models. Pre-trained models enable a lightweight knowledge transfer from research to SMEs through out-of-the-box models. To be more precise, we use YOLO_v5 (You Only Look Once - Version 5), a real-time object detection algorithm. We fine-tune the algorithm by optimizing the weights in the last layers. As a short explanation, the YOLO v5 algorithm is one of the most efficient object detection algorithms in CV. It builds on a Convolutional Neural Network (CNN), capable of real-time predictions with high accuracy in most cases. First, it applies the CNN to the complete input image before it divides the given image into different regions where YOLO_v5 predicts bounding boxes and probabilities. Afterward, these probabilities are weighted to identify the region where the object, i.e., damage, appears [Re16]. We compared the YOLO v5 algorithm to other models (e.g., EfficientDet and MobileNet) but YOLO v5 yields the best accuracy scores and processing performance - which is critical when conducting automated maintenance. We use 111 thermal drone images (including several damages) with a pre-processed size of 416x416. The training set entails 78 thermal drone images. The validation set comprises 21 images. Finally, the test set entails 12 images. We trained the model for 3,000 epochs using a batch size of 32. In Fig. 3, we present our models’ output results, i.e., annotated thermal images, and we can observe that it can capture certain solar panel damages. We are currently only using one class/object called 'broken', which indicates whether a panel is damaged or not. Future work can extend the number of classes to distinguish between different kinds of damages solar panels can exhibit. For example, a typical damage is a defect bypass diode, where a whole substring of the panel sees a loss in efficiency (see left panel Fig. 3). Besides, only single cells in the panel can be damaged, too (see right panel Fig. 3). Additionally, other relevant damages include cell cracks and defective junction boxes.

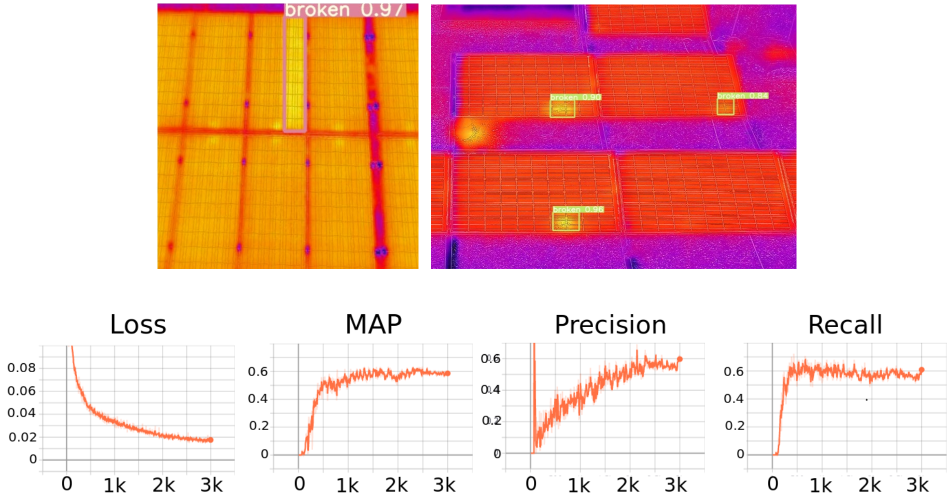

Fig. 3: Exemplary Output and Performance Metrics of YOLOv5 during Training

3.3 Visualization Layer

A powerful visualization solution supports technicians and decision-makers to comprehensively monitor several solar plants and PV modules. However, it is necessary to assign the faulty solar panels to their DT representation to visualize the ML phase results. To this end, we leverage the fact that state-of-the-art drones tag and store the pictures they take together with additional meta-information like the time and location of each image. In our example, the location consists of the drone's exact location and the camera's exact location as three-dimensional GPS coordinates, including the height of the drone relative to the departure point. Additionally, it stores the camera angle from the GPS coordinate as pitch, yaw, and roll. We can compute each solar panel's approximate position on a picture by applying linear algebra based on this information (see Fig. 4).

Finally, we can visualize the solar panels' DT in a central dashboard. Such a dashboard (see Fig. 4) provides a high-level overview of all modules' status and their need for action. Possible information includes basic data like the manufacturer of the modules and year of manufacture, performance measures like the energy produced, manually recorded notes like damages, and damages detected automatically. The dashboard can enrich this information with additional details and explanations, e.g., the remaining lifetime of a module, information regarding the type of damage (like cell crack or defect bypass diode), why the ML model classified a module as damaged, and recommendations for actions. Thus, it is a helpful extension for various stakeholder groups. On the one hand, various related industries such as the housing industry can adopt this dashboard to their existing Computer-Aided Facility Management systems to extend their infrastructure documentation. On the other hand, solar park operators and technical support companies can focus on the predictive components to improve their maintenance process.

Fig. 4: From a Digital Twin Representation to a Comprehensive Dashboard

3.4 Artifact Evaluation

To establish convincing evidence that our artifact provides a practical approach for modeling a virtual entity of the physical structure, we evaluated the artifact's performance in terms of validity and utility in a real-world setting. We observed several solar park sites for validity and used the artifact to create fitting virtual entities for the physical structures. Overall, we observed three different solar parks and 12 residential roofs equipped with PV modules. We collected 237 thermal images and identified 166 faulty cells. Overall, the tests on PV solar parks corroborate the efficiency and value of the proposed artifact. When looking at our ML model's performance metrics, we focus on the Mean Absolute Precision (MAP), Precision, and Recall, as these are the most reliable metrics in CV tasks. The MAP is 63%, Precision 58%, and Recall 63% on the test set. Although our models' results promise a good performance, training the ML model based on a different algorithm or more input data might be beneficial. However, different CV algorithms require significantly higher computational power and longer training times. Thus, they might be unsuitable for fast and efficient deployment and the hardware stack of most SMEs. Moreover, when comparing our model's results to EfficientDet and MobileNet, YOLOv5 still outperformed both lightweight models in terms of MAP (EfficientDet: 28.26%; MobileNet: 18.33%). Notably, the presented approach's identified damages might be incomplete or include so-called false positives (i.e., working modules wrongly classified as damaged). Against this background, our artifact is a lightweight support solution to significantly reduce expert knowledge needs, even though it does not replace expert knowledge altogether.

Overall, our artifact presents both advantages and disadvantages. One shortcoming of the described artifact lies in the frequency of data collection. The data is not real-time, i.e., permanent measurable, but instead based on automated maintenance cycles that might only occur several times a year. However, this is sufficient for most SMEs, as most of them currently have less control over such an infrastructure. On the beneficial side, our artifact can partly automate the full process by using autonomous flight capabilities of drones based on a predefined route combined with automated scheduling. Since flight scheduling depends on multiple factors, e.g., on the maintenance need and external conditions like wind, sunlight, and clouds, ML can predict an optimal timing. Due to legal restrictions, fully automated drone-based maintenance is currently not possible but by combining the drone flight with the required on-site inspections, we can decrease the necessary timeframe as well as staff requirements. Our proposed semi-automated process can decrease maintenance efforts, reduce costs, and increase structured information available for further analysis and decision-making.

To assess whether the artifact also fulfills necessary utility criteria in the problem space, we contacted the responsible technical operations managers and solar park operators of several sites. By showcasing the performance and discussing the results, we established a confidence level in our solution. The transformation of the whole process into separate services enables the transformation of PV systems into virtual entities. Thus, stakeholders can select the relevant services they want to use, increasing the presented system's usefulness. Since most PV system operators currently do not rely on a digital representation of their modules, they demand such a solution. Furthermore, our artifact provides a lightweight AI application that could be further adopted for other problem spaces. SMEs could adopt this approach utilizing pre-trained image recognition models to build specific solutions in their area of expertise. This solution lowers the barrier of entry for digitalization and implementing smart maintenance processes enabling sustainable business models also for SMEs.

4 Conclusion

Particular context areas and industries still lack the availability of structured qualitative data to analyze the condition of physical assets. Modern technologies like cloud computing, Digital Twins, and Machine Learning can help SMEs to tackle this task by integrating a virtual entity for physical structures. We demonstrated the applicability of such a transformation process on the example of PV modules. PV modules are rarely digitalized, and their maintenance is very work-intensive. With a drone equipped with a thermal camera, we reveal an improvement of current maintenance processes by using Computer Vision. Thus, we can transform solar panels and their unstructured data consisting of thermal images into their DTs. This information is useful for further processing, like predicting maintenance needs and automatic scheduling.

We contributed to the theory of processing drone imagery to identify the respective DTs based on location data. Additionally, we provide valuable insights into the process of optimizing PV maintenance. Last, our results serve as input for SMEs decision-makers on how to collect and evaluate PV modules' respective physical conditions and improve current processes.

While our demonstration focuses on PV maintenance, a more abstract representation would also enable an operationalization of the artifact into other context areas. This approach would thus allow for significant knowledge transfer between current research and SMEs and enables sustainable business models through automation. Potential context areas include all those with larger physical structures and a lack of digitalization. Possible applications are virtual identities for buildings, public infrastructure, or traffic surveillance. Such a knowledge transfer would also significantly increase the value and validity of the researched approach.

Further research can either transfer our approach to a different context area or improve certain steps of our approach. Due to the modularity of possible services, future work could segment issues and research questions and address them individually without needing to address the whole implementation.

Acknowledgements

This work was funded by the BMWi through the ForeSight project. Additionally, we like to thank Oberhessische Versorgungsbetriebe AG (OVAG), RMS Rhein Main Solarpark GmbH, and GreenVesting GmbH & Co. KG for their time and valuable support. The authors state no conflict of interest.

Bibliography

[Al14] Alestra, S., Brand, C., Erofeev, P., Bordry, C., Burnaev, E., Papanov, A.: Rare Event Prediction Techniques in Application to Predictive Maintenance of Aircraft. In: Proceedings of ITaS conference, pp. 32–37, 2014.

[Be17] Betti, A., Trovato, M.L.L., Leonardi, F.S., Leotta, G., Ruffini, F., Lanzetta, C.: Predictive Maintenance in Photovoltaic Plants with a Big Data Approach. 33rd European Photovoltaic Solar Energy Conference and Exhibition; pp. 1895-1900, 2017.

[Be20] Bell, J.: Machine Learning: Hands-On for Developers and Technical Professionals. John Wiley & Sons, 2020.

[BS15] Buerhop-Lutz, C., Scheuerpflug, H.: Inspecting PV-plants using aerial, drone-mounted infrared thermography system. 2015.

[CBK09] Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: A survey. ACM Comput. Surv. 41, article 15, pp. 1-58, 2009.

[De12] Denio, H.: Aerial solar Thermography and condition monitoring of photovoltaic systems. In: 2012 38th IEEE Photovoltaic Specialists Conference, pp. 613–618, 2012.

[De18] De Benedetti, M., Leonardi, F., Messina, F., Santoro, C., Vasilakos, A.: Anomaly detection and predictive maintenance for photovoltaic systems. Neurocomputing 310, pp. 59–68, 2018.

[EBA20] Errandonea, I., Beltrán, S., Arrizabalaga, S.: Digital Twin for maintenance: A literature review. Computers in Industry 123, article 103316, 2020.

[GH13] Gregor, S., Hevner, A.: Positioning and Presenting Design Science Research for Maximum Impact. MIS Quarterly 37, pp. 337–356, 2013.

[GV17] Grieves, M., Vickers, J.: Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. In: Kahlen, F.-J., Flumerfelt, S., and Alves, A. (eds.) Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches, pp.85–113. Springer International Publishing, Cham, 2017.

[Ha10] Hashemian, H.M.: State-of-the-art predictive maintenance techniques. IEEE Transactions on Instrumentation and measurement 60, pp. 226–236, 2010.

[Mo02a] Mobley, R.K.: An introduction to predictive maintenance. Elsevier, 2002.

[Pe07] Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A design science research methodology for information systems research. Journal of management information systems 24, pp. 45–77, 2007.

[Re16] Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You Only Look Once: Unified, Real-Time Object Detection. Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.

[Sc17] Schleich, B., Anwer, N., Mathieu, L., Wartzack, S.: Shaping the digital twin for design and production engineering. CIRP Annals 66, pp. 141–144, 2017.

[Sh18] Shubenkova, K., Valiev, A., Shepelev, V., Tsiulin, S., Reinau, K.H.: Possibility of Digital Twins Technology for Improving Efficiency of the Branded Service System. In: 2018 Global Smart Industry Conference (GloSIC), pp. 1–7, 2018.

[Sp12] Spagnolo, G.S., Vecchio, P.D., Makary, G., Papalillo, D., Martocchia, A.: A review of IR thermography applied to PV systems. In: 2012 11th International Conference on Environment and Electrical Engineering, pp. 879–884, 2012.

[WY07] Widodo, A., Yang, B.-S.: Support vector machine in machine condition monitoring and fault diagnosis. Mechanical systems and signal processing 21, pp. 2560–2574, 2007.

[Yü20] Yürüşen, N.Y., Rowley, P.N., Watson, S.J., Melero, J.J.: Automated wind turbine maintenance scheduling. Reliability Engineering & System Safety 200, article 106965, 2020.