High-Throughput Quality Inspection of Solar Cells Using Deep Learning Under Consideration of Its Sustainability Impact

Thesis

submitted in partial fulfillment of the
requirements for the degree of

Master of Science
in
Industrial Ecology

Leiden, The Netherlands, August 31, 2022
Word count: 21,382 words

Marko Reinhard

TUD Student ID: 5423600 / LU Student ID: 2834472

1st Examiner:	Rudi Santbergen, PhD (Delft University of Technology)
2nd Examiner:	Carlos Felipe Blanco, PhD (Leiden University)
Technical Advisor:	Philipp Kunze, MSc (Fraunhofer ISE)

TU Delft Fraunhofer ISE Universiteit Leiden

Collaboration with External Research Institute

This project is carried out in partial fulfilment of the requirements for the degree of Master of Science in Industrial Ecology at TU Delft and Leiden University. It is part of a collaboration with the Fraunhofer Institute for Solar Energy Systems (ISE), Heidenhofstraße 2, 79110 Freiburg im Breisgau, Germany.

Confidentiality Clause

This master thesis contains confidential data of Fraunhofer ISE. This document may only be made available to the first and second reviewers and authorized members of the board of examiners until October 1, 2023. Any publication and duplication of this master thesis - even in part - is prohibited until then. An inspection of this work by third parties requires the expressed permission of the author and Fraunhofer ISE.

Declaration

I hereby declare that this thesis is a presentation of my original research work and that no other sources were used other than what is cited. I furthermore declare that wherever contributions of others are involved, this contribution is indicated, clearly acknowledged and due reference is given to the author and source. I also certify that all content without reference or citation contained in this thesis is original work. I acknowledge that any misappropriation of the previous declarations can be considered a case of academic fraud.

Leiden, The Netherlands, August 31, 2022

Location, Date

Signature of Author

Abstract

Photovoltaics (PV) are already an important part of the emerging global energy system. Rapidly decreasing production costs due to high throughput production have made this development possible. To meet market demands, it will remain important to further scale up global production. However, during the production of solar cells, several defects can occur. Current approaches in quality inspection based on electroluminescence (EL) are reaching their speed limits. The goal of this thesis project is to advance current solar cell quality inspection by using deep learning-based computer vision (CV) algorithms to detect production defects without human supervision at high speeds. The goal is to achieve this while reducing the necessary manual efforts to label (annotate) defects in the training data of such algorithms. Next to developing and evaluating algorithms, the second goal of this project is to investigate in which ways and to which extent the technical innovation can impact the sustainability performance of the production process. Multiple scenarios are investigated using a Life Cycle Assessment (LCA) model. The LCA results are used to estimate the potential large-scale impact of increasing solar cell production throughput.

The results show it is possible to detect defects in EL measurements taken at very high process speeds using convolutional neural networks (CNNs). This could enable much higher throughput rates than current state-of-the-art solutions. A slight reduction in detection accuracy currently has to be expected compared to non-moving measurements. From a sustainability perspective, no reasons were found that would advise against deploying high throughput quality inspection. The environmental impact of the developed algorithm due to energy consumption is neglectable on a per-cell level. However, it is emphasized that future projects aiming at the further scale-up of PV production should automatically integrate an end-of-life (EoL) perspective. This is required to mitigate possible negative impacts made possible by technological advancements that happen faster than the development of EoL schemes.

Contents

1 Introduction **3**

2 Fundamentals **6**
2.1 Crystalline Silicon (c-Si) Solar Cells and Their Future 6
2.2 Spatially Resolved Characterization of Solar Cells 6
2.3 Machine Prototype for Faster Measurements in Mass Production 9
2.4 Machine Learning and Convolutional Neural Networks 10
2.5 Interdisciplinary Nature of This Project 10

3 Related Work **12**
3.1 State of the Art Quality Inspection in Industry 12
3.2 CNN-Based Inspection of PV Systems 12
3.3 Innovation of Measurement Techniques 13
3.4 Sustainability Aspects 13
3.5 How This Project Differs From Related Work 13

4 Technical Approach **15**
4.1 Goal and Requirements Definition 15
4.2 Data and Machine Learning Pipeline 16
4.3 Smart Labeling 17
4.3.1 Overview 17
4.3.2 R_S Labels 19
4.3.3 j_0 Labels 22
4.4 Network Architecture 23

5 Technical Experiments **25**
5.1 General 25
5.1.1 Neural Network Settings 25
5.1.2 Evaluation Metrics 26
5.2 Experiment 1: Still, Non-Blurred 27
5.2.1 Goals 27
5.2.2 Measurements at Fraunhofer ISE 27
5.2.3 Dataset 28
5.2.4 Scenario Definition 29
5.3 Experiment 2: Artificial Motion Blur 29
5.3.1 Goals 29
5.3.2 Dataset 30
5.3.3 Estimation of Realistic Motion Blur 30
5.3.4 Scenario Definition 31

5.4 Experiment 3: Moving, Real-World Blurred . . . 32
5.4.1 Goals . . . 32
5.4.2 Measurements at Industrial Partner . . . 32
5.4.3 Dataset . . . 34
5.4.4 Scenario Definition . . . 34

6 Technical Results **36**
6.1 Experiment 1: Still, Non-Blurred . . . 36
6.1.1 Quantitative Results . . . 36
6.1.2 Qualitative Results . . . 39
6.2 Experiment 2: Artificial Motion Blur . . . 40
6.2.1 Quantitative Results . . . 40
6.2.2 Qualitative Results . . . 45
6.3 Experiment 3: Moving, Real-World Blurred . . . 51
6.3.1 Quantitative Results . . . 51
6.3.2 Qualitative Results . . . 54

7 Technical Discussion **62**
7.1 Detection Accuracy and Impact of Motion Blur . . . 62
7.2 Smart Labeling . . . 63
7.3 Dataset . . . 63

8 Sustainability Approach and Initial Assessment **65**
8.1 Framework Definition . . . 65
8.2 Initial Application of Framework Levels . . . 66
8.2.1 Algorithm Level . . . 67
8.2.2 Process Level . . . 69
8.2.3 Life Cycle Assessment Model . . . 69

9 Sustainability Assessment **70**
9.1 Rigorous Scenario . . . 70
9.1.1 Overview . . . 70
9.1.2 Methodology . . . 71
9.1.3 Results . . . 73
9.1.4 Discussion . . . 74
9.2 Efficiency Scenario . . . 75
9.2.1 Overview . . . 75
9.2.2 Methodology . . . 75
9.2.3 Results . . . 76
9.3 Additional Scenarios and LCA Connection . . . 77
9.3.1 Definition of Additional Scenarios . . . 77
9.3.2 Modeling Cleanroom Usage . . . 79
9.3.3 Factory Building Materials . . . 80

10 Integration of Sustainability Results **81**
10.1 Main Results of Assessment Scenarios 81
10.2 Process Level 83
10.3 Large-Scale Level 84

11 Sustainability Discussion **90**

12 Conclusion **93**
12.1 Technical 93
12.2 Sustainability 94

Bibliography **VI**

Appendix A - Code and Models **XIII**

Appendix B - LCA Unit Processes and Absolute Results **XIV**

List of Abbreviations

PV Photovoltaics

CNN Convolutional neural network

CV Computer vision

DL Deep learning

EL Electroluminescence

PL Photoluminescence

STC Standard test conditions

c-Si Crystalline silicon

ReLU Rectified Linear Unit

C-DCR Dark saturation current and series resistance

TOPCon Tunnel Oxide Passivating Contact

PERC Passivated Emitter and Rear Contact

HJT Heterojunction

R_S Series resistance

j_0 Dark saturation current density

ML Machine learning

AI Artificial intelligence

ROI Region of interest

RSP Rotational screen printing

FSP Flatbed screen printing

LCA Life cycle assessment

GHG Greenhouse gas

EoL End-of-life

SDG Sustainable Development Goal

ITRPV International Technology Roadmap for Photovoltaic

CAPEX Capital expenditure

COO Cost of ownership

OOE Overall equipment effectiveness

1 Introduction

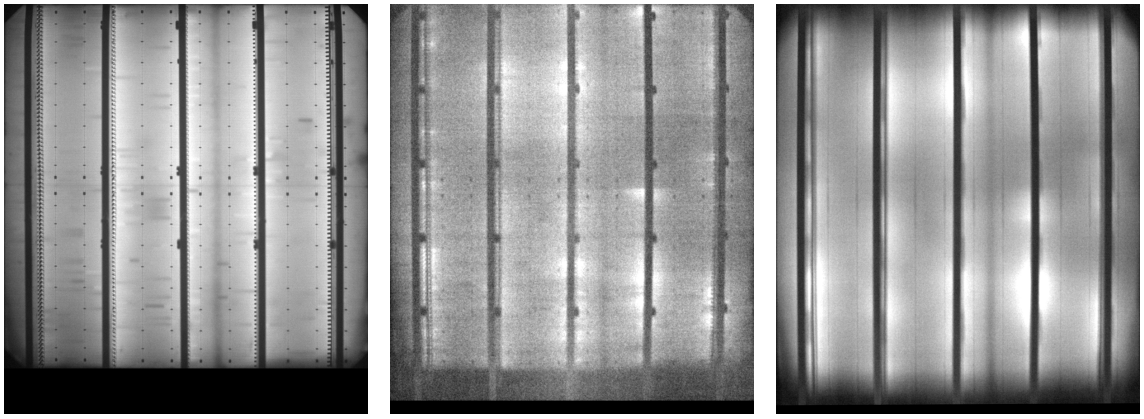
Increasing the production throughput of solar cells is a central lever to reduce production costs in photovoltaics (PV) and to meet global demand. Especially better machine and factory space utilization help lower the capital expenditure (CAPEX) per cell significantly [1], [2]. In mass production, solar cells have to be characterized at the end of the production line, i.e., their technical characteristics have to be determined. This is an important step in quality inspection. Currently, this is either done by measuring the current and voltage (I-V) of the cell or via electroluminescence (EL) imaging. Nowadays, EL imaging can be regarded as one of the standard means of final quality inspection of solar cells in mass production [3]. The International Technology Roadmap for Photovoltaic (ITRPV) (2021 Results) states a world market share of EL imaging of 50 % in 2021 and predicts a share of above 90 % for 2032 [4].

When scaling up the production throughput of solar cells, existing quality inspection systems such as EL imaging are too slow leading to the need of faster systems and advanced image evaluation. Within the current state-of-the-art, solar cells are stopped on the conveyor belt to take measurements such as EL images. This limits the production throughput. Prototypes for faster systems are in development performing measurements while the solar cells are in motion. Currently, EL images are often already interpreted using machine-learning techniques such as convolutional neural networks (CNNs). However, when images are taken in motion, motion blur occurs in the images which makes it hard to impossible for a human operator to find defects and process abnormalities. This can be compared to the effects occurring when taking an image of a quickly moving train. This leads to the main technical research question: *How well can electroluminescence images of TOPCon solar cells taken in motion at high production throughput be analyzed by deploying CNNs?*

The second major challenge of this project is to consider the sustainability impact in the early development phase of this technical development. Often in innovation quality is prioritized without considering the sustainability impact. In the mass production of solar cells, small process changes potentially have a large impact. Critical decisions such as which cells are sorted out during production and which are disposed are based on the output of defect detection algorithms whose sustainability impact has not sufficiently been studied until now. E.g., EL imaging could also become an important means of process feedback with the potential of stopping the production in case of systematic defects [5] decreasing the quality of all following cells. The main sustainability research question within this work is: *What is the prospective sustainability impact of novel quality inspection algorithms in high throughput PV production?*

The technical research question of this project is approached via the development and evaluation of a suitable defect detection algorithm that can detect production defects in moving measurements. The challenge is demonstrated in Figure 1.1. The image in Figure 1.1a is taken on a

standard measurement system without moving the cell. Dark roundish spots indicate possible process defects. In fast measurements, as shown in Figure 1.1b and Figure 1.1c even experts are not able to safely predict defects, either due to low contrast at short illumination times or due to motion blur at high illumination times. The algorithm trained within this project is capable of predicting the defects.



(a) EL image in non-moving case (reference)

(b) Moving EL image at 2 ms illumination time

(c) Moving EL image at 7 ms illumination time

Figure 1.1: Example of a measurement image (left) taken without moving the cell, a moving image with motion blur at low illumination time (middle), and a moving image with motion blur at high illumination time (right).

> “Without new approaches, we become ’blind’ in quality inspection when scaling up production throughput.”

The technical goal of the project is to set up an end-to-end pipeline from taking measurements of solar cells in the laboratory, automatic labeling (annotation) of defects, training state-of-the-art convolutional neural networks (CNNs), and evaluating the results. Around 500 TOPCon solar cells are measured in the laboratory at Fraunhofer ISE. Three experiments are designed aiming at training a defect detection algorithm that can predict the location of two types of defects in EL images: series resistance (R_S) defects and dark saturation current density (j_0) defects. In the first experiment, non-moving images such as those shown in Figure 1.1a are used to train a CNN serving as a reference for the following two experiments. In the second experiment, measurement images with simulated motion blur are used. In the third experiment, measurement images with real-world motion blur such as shown in Figure 1.1c taken at an industrial partner company are used.

To train the CNNs, a novel approach called “Smart Labeling” based on Kunze et al. [5] is developed to annotate defects in solar cells more quickly and reliably. In an industrial context, this new labeling approach would allow a faster and potentially more accurate deployment of a defect detection algorithm. It uses dark saturation current and series resistance (C-DCR) measurements as a reference technique to create spatially resolved labels for the defects. The CNNs are then trained with pairs of EL images and the smart labels as input coming from the three experiments,

respectively. The performance of the defect detection algorithms is both evaluated quantitatively by statistical scores and by a qualitative analysis focusing on specific defects in sample solar cells.

The sustainability research question is approached via the definition of a new framework that allows viewing the technical innovation from three different perspectives: The algorithm level, the process level, and the large-scale level. The framework is applied to the technical innovation of this project. Different scenarios to estimate the impact of the algorithms are developed. These scenarios are then tested for their prospective environmental impact using a life cycle assessment (LCA) model. The LCA model estimates the environmental impact of producing 1 kWh electricity using a 7 $\mathrm{kW_p}$ slanted roof PV installation in Europe in different scenarios.

This document provides a brief outline of the required technical background in Chapter 2. Most importantly, the used measurement techniques are introduced. In Chapter 3, an overview of current research in the field of solar cell quality inspection and the sustainability impact of PV is given. The technical research is grouped into four chapters, Chapter 4 describing the methodological approach, Chapter 5 defining the experiments, Chapter 6 providing the experimental results, and finally Chapter 7 discussing the technical results.

The sustainability research begins with the definition of the aforementioned framework to view the technical innovation from multiple perspectives in Chapter 8. In this chapter, the framework is initially applied to the case and the innovation is viewed from algorithm level, process level, and large-scale level in an initial assessment. In Chapter 9 different scenarios considering both the algorithm and process level are defined and tested using an LCA model. In Chapter 10 the findings from the initial assessment and the LCA scenarios are integrated and used to project the large-scale impact. Furthermore, the results are discussed and ideas for the future are presented. Finally, a combined conclusion states the main findings, implications, and limitations of this research in Chapter 12.

2 Fundamentals

2.1 Crystalline Silicon (c-Si) Solar Cells and Their Future

Solar energy is an abundantly available energy source. Photovoltaics make use of the photoelectric effect. In very simple terms, charge carriers are freed by solar energy in semiconductor materials, mostly silicon. The generated free charge carriers are separated, creating a difference in electric potential (voltage). This voltage can be used to drive electrical circuits. The crystalline silicon solar cell using the photoelectric effect has been invented in 1954. Since then, solar cells have become an important option for producing renewable electrical energy [6].

The computer vision (CV) algorithms that are developed and evaluated in this project will be applied to Tunnel Oxide Passivating Contact (TOPCon) cells. TOPCon is one of the newest crystalline silicon (c-Si) cell technologies that have the potential for mass production offering a higher conversion efficiency than current Passivated Emitter and Rear Contact (PERC) cells. It could replace PERC as the "actual 'workhorse' of the PV industry" [7, p. 10] because only evolutionary upgrades are required in the production process [8]. This work focuses on TOPCon solar cells in particular to make sure the findings generalize well to the future mass production of c-Si solar cells.

2.2 Spatially Resolved Characterization of Solar Cells

Taking measurements of solar cells to understand their electrical and optical properties is referred to as "characterization". In mass production, characterization is carried out at the end of the production line. To understand the measurements taken in this project, a piece of fundamental knowledge of the electrical characteristics of solar cells is required.

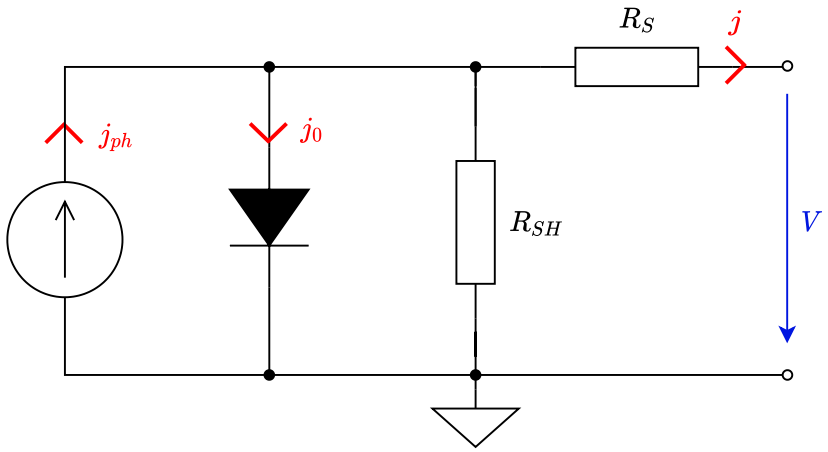


Figure 2.1: Single-diode equivalent circuit of a solar cell.

The current-voltage characteristics of an illuminated solar cell can be described with a simple equivalent circuit as shown in Figure 2.1. In this circuit, it is assumed that the solar cell behaves as an ideal diode [9]. The notation and its capitalization of currents or current densities differ in literature. This work assumes current densities j instead of currents I.

The current source providing the current density j_{ph} represents the photo-generated current density. The current density through the diode, j_0, can be regarded as a measure of recombination. A diode with larger recombination has a larger j_0 [10]. In case of recombination, free charge carriers recombine and are therefore no longer available to drive an external circuit connected at the terminals (right side of Figure 2.1) and the energy is lost. For this reason, it is desired to keep recombination effects in solar cells low.

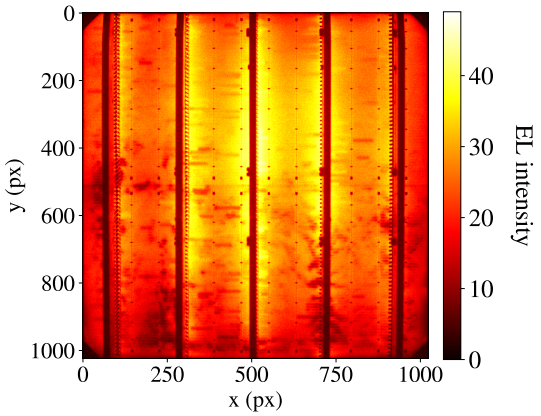
The series resistance R_S models the series resistance loss of the solar cell. It represents the resistance that the photo-generated current has to pass on its way to the terminals of the solar cell. The main contributors to R_S are the resistance of the metal electrodes connecting the semiconductor material of the solar cell to an external circuit and the contact resistance of this connection. Additionally, the bulk resistance of the junction of the solar cell contributes to R_S [9].

There are several imaging techniques for defect detection in solar cells. electroluminescence (EL) measurements make irregularities in the crystal structure of the cell visible. Additionally, EL measurements can be used to find process defects [11]. To take an EL measurement, the solar cell needs to be contacted. Via the contacts, an external current is applied to the cell. The cell then starts irradiating light due to radiative recombination of the provided charge carriers that is captured by a camera. An example EL image is given in Figure 2.2a. The most important variable parameters of EL measurements are the illumination time T_L, the camera gain G, and the current I applied to the solar cell. Especially these parameters will become relevant later in this project.

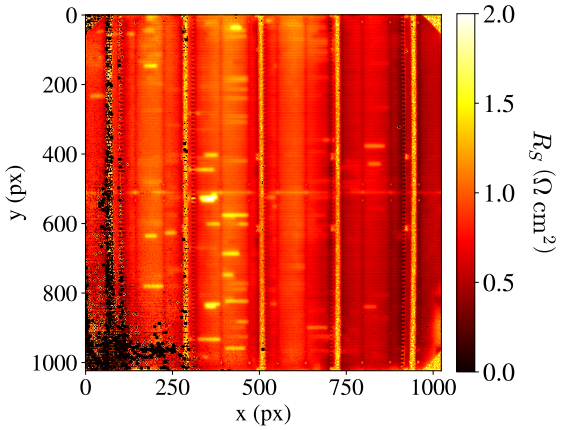
Photoluminescence (PL) measurements are mainly used to find material defects [12]. To take a PL measurement, excitation light is applied to the cell. This light generates electron-hole pairs in the semiconductor material that recombine. The recombination leads to the emission of photons which can be captured by a camera.

In dark saturation current and series resistance (C-DCR) measurements, EL and PL are combined. A C-DCR measurement consists of two images. One image shows the spatial distribution of the dark saturation current density (j_0), whereas the second image contains the spatial distribution of the series resistance (R_S) [13]. Example C-DCR images are given in Figure 2.2b and 2.2c. Note that the pixel intensity values in C-DCR images have a physical unit. The j_0 image shows a spatial representation of the dark saturation current density in the unit $\frac{\mathrm{A}}{\mathrm{cm}^2}$. As j_0 can be regarded as a measure of recombination and recombination leads to losses in solar cells, i.e., a high j_0 can be regarded as a defect. The R_S image shows a spatial representation of the series resistance. The intensity of every pixel has the unit $\Omega\ \mathrm{cm}^2$. As can be seen from the equivalent circuit in Figure

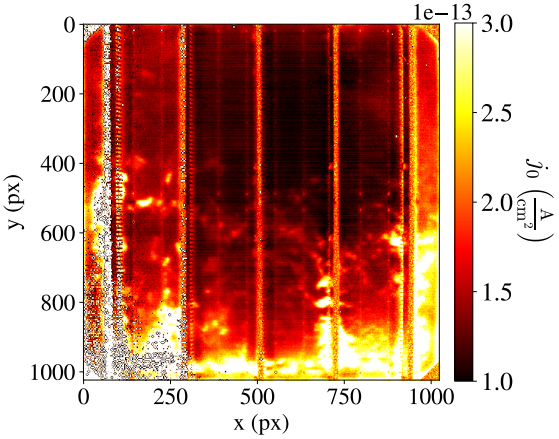
2.1, the R_S is on the main current path to the load. A higher series resistance leads to a higher voltage drop across R_S and therefore to higher losses.



(a) EL image



(b) C-DCR R_S image



(c) C-DCR j_0 image

Figure 2.2: Example of EL and C-DCR measurement images. Note that both EL and C-DCR images show five contact bars. In these regions, the measurements of both techniques are undefined and do not deliver meaningful measurement results.

Finger interruptions are common process-related defects in crystalline silicon solar cells that appear as bar-shaped dark areas in the EL image. Figure 2.2a and 2.2b both show some finger interruptions. Fingers are part of the metallization of the solar cell and are collecting and transporting currents in the solar cell. Damaged fingers lead to a reduction in cell performance [14]. Finger interruptions are well visible in R_S images from C-DCR measurements as can be seen as bright areas in Figure 2.2b.

Common material-related defects are dislocation structures, i.e., irregularities in the crystal structure of the cell leading to higher recombination. Process-related recombination losses often occur in the edge regions due to imperfections in the cut edges [15]. In the pre-industrial TOPCon cells used in this project these defects occur regularly. They can be well seen in the example j_0 image in Figure 2.2c. As a general rule a locally increased value in C-DCR images can indicate a defect.

2.3 Machine Prototype for Faster Measurements in Mass Production

In current state-of-the-art mass production of solar cells, a production throughput of around 5,000 wafers per hour on a single production line is achieved [16]. In the end-of-line characterization, the cells are currently stopped to contact the cells to the measurement device and to take EL measurements. In a typical production cycle, the steps *move-in of cells and contacting* and *decontacting and move-out of cells* take about 450 ms each [3] which are multiples of the measurement time itself. This is too slow when the production throughput on a single production line should be doubled or even quadrupled.

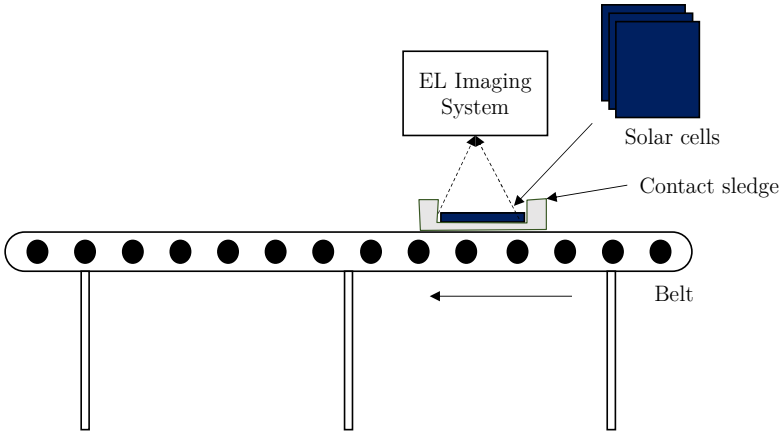


Figure 2.3: Concept sketch of existing prototype to test moving EL measurements.[1]

Therefore, the industrial partner company of this project has developed a hardware prototype of a measurement system that consists of a moving contact sledge as shown in Figure 2.3. The concept entails that the cells would be contacted and measured "on-the-fly" removing the previously mentioned time consuming contacting steps. At the current state of the development, a solar cell is inserted into the sledge manually one at a time. The sledge moves the cell below a camera that takes a measurement image at a given illumination time T_L while the cell is moving at a belt speed of $v_b = 1.9\frac{m}{s}$. This way, more than 16,000 cells per hour could be measured[2]. In the real-world experiments later in this project, T_L is varied to investigate different trade-offs between camera gain (amplification), cell current, and illumination time.

The moving sledge is contacted to the measurement system via a bar of sliding contacts. The sliding contacts are connected to contact bars that are pushed on the solar cell under test. The contact bars are used to connect the solar cell to the EL measurement device.

[1] Adopted with permission from Leslie Kurumundayil (Fraunhofer ISE). Internal Team Presentation.
[2] Johannes Greulich (Fraunhofer ISE). NextTec Equipment Data Input Sheet for Quality Inspection Processes. July 2nd, 2020.

2.4 Machine Learning and Convolutional Neural Networks

Deep Neural Networks (DNNs), a machine learning technique using Artifical Neural Networks (ANNs) [17], are an important technique used in this research project. In particular, CNNs, a type of DNNs, are used for the computer vision (CV) tasks. CNNs have various applications in computer vision, speech and natural language processing, and are one of the most extensively studied type of DNNs [18]. Generally, this thesis assumes some basic familiarity with the approach of deep learning. For reference, the book *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems* is recommended [19]. The book focuses on using the open-source library TensorFlow while this project uses PyTorch [20], but the core concepts are similar.

This project will apply a supervised learning approach, i.e., several models are trained on input data (EL images) that is labeled with the defects in a spatially resolved manner that the neural network should predict. A detailed overview of the applied network architecture is given in Section 4.4. The neural network models in this project are supposed to perform spatially resolved defect detection in EL images. This means every pixel should be assigned to one of the four classes No R_S defect, R_S defect, No j_0 defect, or j_0 defect. A pixel can be part of both a R_S defect and a j_0 defect. Using four classes allows the model to predict both j_0 and R_S defect for the same pixel as there are multiple combinations possible, such as R_S defect and No j_0 defect or R_S defect and j_0 defect [5]. Assigning each pixel a class is referred to as a semantic segmentation task in literature [21].

2.5 Interdisciplinary Nature of This Project

This project aims at combining the fields of Industrial Ecology, Machine Learning and Electrical Engineering to advance the quality inspection technologies used within the production of solar cells. The background knowledge of the target audience is diverse, and not all readers may be acquainted with the field of Industrial Ecology.

The field of Industrial Ecology is a discipline mainly involving scientists, engineers, and policy-makers. It shows “sympathy for multidisciplinary and interdisciplinary research and analysis" [22, p. 6]. There is no distinct definition of the field, especially due to the involvement of multiple disciplines [23]. However, the following statement by Graedel and Allenby describes the general concept of the field well:

> “Industrial ecology is the means by which humanity can deliberately and rationally approach and maintain sustainability, given continued economic, cultural, and technological evolution. The concept requires that an industrial ecosystem be viewed not in isolation from its surrounding system, but in concert with them. It is a systems view in which one seeks to optimize the total materials cycle from virgin material, to finished material, to component, to product, to obsolete product, and to ultimate disposal. Factors to be optimized are resources, energy and capital.”
>
> *Graedel and Allenby (2003) [24, p. 18]*

The main contribution of this project takes place on the technological side of the field. The advancement of defect detection techniques for solar cells aims at increasing eco-efficiency by reducing material waste and using machines more efficiently. At a greater level, it potentially contributes to make affordable and clean energy more accessible. However, this is to be verified in this research. An important element within the *Industrial Ecology toolbox* is the LCA methodology [25]. In this project, an LCA model is used to estimate the potential environmental impact of the developed defect detection algorithm in multiple scenarios.

3 Related Work

3.1 State of the Art Quality Inspection in Industry

EL imaging for end-of-line quality inspection of solar cells can be regarded as the current industry standard. As briefly mentioned in the introduction, the ITRPV (2021 results) states a world market share of EL imaging of 50 % in 2021 and predicts a share of above 90 % for 2032 [4]. In mass production, production rates of around 5,000 wafers an hour on a single production line can be achieved [16]. At these or even higher production rates, a human inspection of every EL image is hardly possible. Therefore, analyzing these images with machine learning approaches, most commonly CNNs, is already the standard. The most common approach applied in industry today is using human-made labels to train the algorithms, i.e., human experts need to annotate a set of existing images [3].

3.2 CNN-Based Inspection of PV Systems

While also being developed and deployed by industrial companies, CNN-based inspection of both solar cells and modules is extensively studied in research. Many approaches focus on finding defects in production on the cell level, whereas others focus on finding defects in modules either in production or in-field.

Su et al. [26], [27] propose defect detection in EL images of solar cells using their own CNN architecture with an emphasis on microcracks and finger interruptions. There is several other work focusing on detecting microcracks in solar cells using CNNs [28]–[32]. Other authors use EL images of PV modules instead of cell images to detect defects either on cell or module level [33]–[36]. Sovetkin et al. [37] investigated various encoder-decoder semantic segmentation models for defect detection on module level with an emphasis on identifying suitable network architectures. There is also work focusing not only on images taken in a production environment but also on in-field applications. Pierdicca et al. [38] employ a CNN for the thermographic inspection of ground-based PV systems. Another focus is to predict physical quantities from measurement images. E.g., Kovvali et al. [39] propose a network that can provide a spatial prediction of j_0 from photoluminescence (PL) images.

It should be mentioned that there are approaches to extract information about finger interruptions and other defects from EL images using classical/human-made image processing methods rather than CNNs. Some, like Tsai et al. [40], rely on threshold values. Others [41] rely on other observed features. CNNs have mostly replaced these approaches, however, for this research they are relevant to extract information from C-DCR reference measurements.

3.3 Innovation of Measurement Techniques

State-of-the-art CNN-based quality inspection heavily relies on innovative measurement techniques. To a certain extent, there is some competition between the deployment of CNNs and the advancement of measurement techniques. Dost et al. [42] propose a method to perform "Advanced Series Resistance Imaging" on solar cells to extract physical information directly from EL images. C-DCR as proposed by Glatthaar et al. [13] even provides information about the R_S and j_0 of the cell. However, the two mentioned approaches are currently not applied inline in mass production due to its limitations such as speed or required hardware.

3.4 Sustainability Aspects

Vinuesa et al. [43] assesses the potential of artificial intelligence (AI) on achieving the Sustainable Development Goals (SDGs) from a very high-level perspective. The multi-level perspective from Geels [44] is commonly applied to view technological transitions from multiple levels (micro, meso, macro) as a nested hierarchy.

Verlinden [45] has investigated which challenges the PV industry has to overcome to reach Terawatt scale manufacturing rates. Kafle et al. [8] explore technology options for cost-efficient industrial manufacturing of TOPCon cells. Aßmus et al. [46] have investigated the performance requirements of crack detection systems.

Müller et al. [47] have carried out a comparative life cycle assessment study of silicon PV modules, especially considering the manufacturing location. Blanco et al. [48] have focussed on evaluating technological developments in the field of PV during early research and development stages.

3.5 How This Project Differs From Related Work

As mentioned in the introduction, current EL inspection is approaching limitations when further increasing production throughput. Especially contacting the cells is time-consuming [3] as it takes longer than the measurement itself. Moving the cells while taking measurements could speed up the process which creates motion blur.

There is no known work investigating the impact of moving EL measurements at cell-level on defect detection performance. Wu et al. [49] proposed an algorithm to remove motion blur occurring at automated optical inspection. Some research at Fraunhofer ISE has focused on architectures to remove motion blur from EL images of solar cells but has not been published yet.

Additionally, most existing work relies on human labels as a reference for training and evaluating their algorithms. Kunze et al. [5] and Greulich et al. [50] consider expert labels poorly reproducible. Different experts produce different labels of the same defects [50]. Additionally, expert labeling is expensive and existing labels cannot easily be adapted to new cell production

processes. The approach chosen in this project can detect relevant R_S and j_0 defects without requiring additional hardware in production (such as PL equipment) with a well-established method (EL) in moving measurements using C-DCR as a reference to label defects for algorithm training. Using C-DCR as an "offline" reference has several advantages. C-DCR is not available as an inline technique, and it could certainly not be applied on moving measurements because it requires taking multiple images to calculate the R_S and j_0 image. C-DCR images allow for a separation of defect structures based on the two distinct R_S and j_0 images which is not easily possible in EL images. However, the reference measurement approach still requires some classical feature engineering as carried out by Tsai et al. [40].

For the sustainability perspective, there is no known approach to directly assess the impact of deploying an AI algorithm to an existing manufacturing process in the PV production chain. There is the need for an approach on the one hand allowing to go more into detail than high-level publications such as from Vinuesa et al. [43] or Verlinden [45] but allows for viewing a broader picture than purely economical and technical assessments as carried out by Aßmus et al. [46] on the other hand.

4 | Technical Approach

4.1 Goal and Requirements Definition

The overarching goal of this project is to investigate the feasibility of spatially resolved high-throughput solar cell quality inspection. As part of the investigation, it has to be assessed how well CNN-based semantic segmentation can cope with artificial and real-world motion blur in the application of defect detection for solar cells at high production throughputs. Another part of the project is to develop and apply a simplified labeling approach for solar cells that allow for faster deployment of defect detection algorithms. For all investigations, a dataset of measurements of TOPCon solar cells has to be created over the course of this research based on real cells. The same cells have to be measured at the industrial partner company which has developed a prototype to test fast EL measurements.

In this project, the absolute achieved detection performance of the defect detection algorithm is not of utmost importance. Due to the high number of new concepts and the explorative characteristic of the research, it is difficult to find baseline performance data. The performance of the defect detection algorithm on still (non-moving) images can serve as a baseline to rate the performance. Other algorithms applied in the industry already achieve high accuracies. Therefore, this work focuses on defect detection in blurry images rather conceptionally as proof of concept than optimizing for the highest accuracies.

A new labeling approach called *Smart Labeling* based on R_S and j_0 C-DCR reference measurements is developed and applied to TOPCon solar cells. The general concept of this labeling approach has only been used in one publication before [5]. It should be redeveloped and improved using a new approach to make it more robust against variations between different solar cell batches and process variations. Hence, finding the optimal neural network architecture is not the goal of the project either as several authors such as Sovetkin et al. (2020) have focused on finding suitable models for EL defect detection. In this project a slightly modified well established architecture is used.

The main requirements derived from the above goals are:

- Development of a functional Smart Labeling algorithm that can be applied to TOPCon, including data collection in the laboratory.
- Defect detection algorithm requirement 1: Separate R_S and j_0 defects in still EL images in a spatially resolved manner. Other defects could be added later, but not in this research project.
- Defect detection algorithm requirement 2: As requirement 1, but on images with artificial motion blur simulating moving measurements.

- Defect detection algorithm requirement 3: As requirement 1 but on images taken at high production throughput with motion blur.

From the above requirements, there are several sub-questions arising that are supposed to be answered by experiments:

- How do labels need to be changed in case of motion blur? Can they remain in the same position, as the features move within the image?
- Do we need to increase the size of the image region as the cell could blur out of the still image sizes?
- How does the chosen approach perform on real-world motion-blurred cell data?

4.2 Data and Machine Learning Pipeline

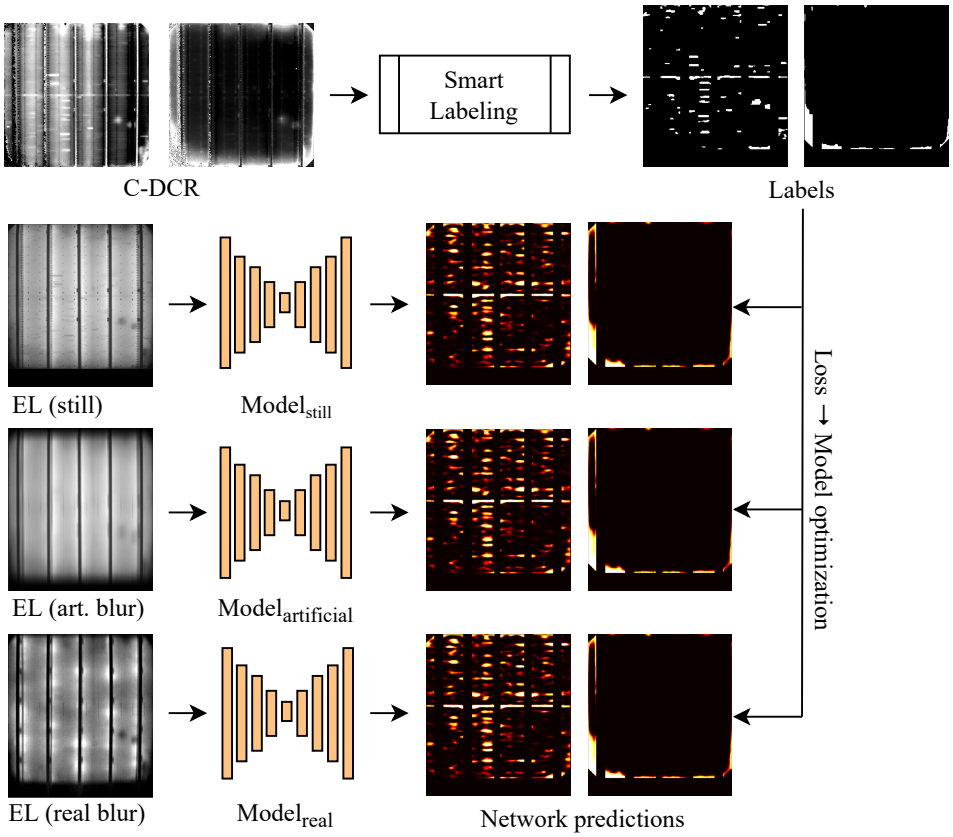


Figure 4.1: Overview of the data and machine learning (ML) pipeline including Smart Labeling based on C-DCR reference measurements. After applying Smart Labeling (first row) the neural network models are trained on three different datasets: Still EL images (second row), EL images with artificial motion blur (third row), and moving EL images taken on a novel prototype measurement device (fourth row).

This project uses an end-to-end supervised machine learning approach as shown in Figure 4.1. In the first row, it can be seen that based on the C-DCR images, the defect labels are calculated by Smart Labeling. The labels are shown as white areas. Following that, CNNs are trained within

three different experiments as shown in rows 2-4. The training is carried out with pairs of either still or blurry EL images and the corresponding labels.

Training the CNNs requires the acquisition of real-world data. Fraunhofer ISE develops new cell architectures and experiments with new production technologies. The experimental cells of future mass-produced cell technologies are usually produced in batches of 10-200 cells. These experimental cell batches are available for measurements. Therefore, this project requires the planning of those measurements and practical lab work (sorting suitable cells into carriers of 100 cells each, coordination with the measurement team at the cell testing machine), and defining suitable measurement parameters. The cell testing machine at Frauhofer ISE is unique as it is customized for the specific requirements of solar cell characterization in pre-mass-production research. The device integrates several measurement devices and can, among others, provide IV-curve measurements, EL, and C-DCR. A selection of cells is measured on the prototype introduced in Section 2.3 at the industrial partner company to create a dataset with EL images showing real motion blur.

After collecting and processing the data, several experiments with artificial neural networks are carried out as shown in Figure 4.1. In Experiment 1, several models are trained on pairs of still (non-moving) EL images and labels from the Smart Labeling algorithm. Several scenarios are executed to find the best-performing model. The best model from Experiment 1 is passed on to Experiment 2. In Experiment 2, the same labels and the same EL images are used to train a network, but while applying artificial motion blur representing different EL illumination times. The results of the experiments are used to evaluate the change in detection performance at different illumination times. Additionally, it is assessed how to adapt the labels and image dimensions to the motion blur. The best-performing models for each illumination time are then passed on to Experiment 3. In Experiment 3, real moving measurements taken at the industrial partner company are used. From this experiment, it can be derived how well the CNN adapts to real motion blur and which trade-offs should be considered in the settings of EL measurements.

4.3 Smart Labeling

4.3.1 Overview

The Smart Labeling approach intends to extract finger interruptions from R_S images and j_0 defects from j_0 images from C-DCR images. The C-DCR images provide a spatially resolved measurement image with physical quantities, i.e., each image pixel intensity represents a physical value. A threshold value can be applied from which a pixel is considered a “defective area”. Ideally, this leads to spatially resolved defect labels based on real physical quantities.

However, in the case of real C-DCR images of real experimental solar cell batches, several practical problems need to be considered. The experimental cell batches at Fraunhofer ISE vary significantly in their electrical and optical characteristics which inhibits the application of a single threshold. Furthermore, the C-DCR images can be inhomogeneous, i.e., pixels of similar physical

value in the real cell are assigned higher pixel intensities in one image region (bottom region) than in another (top region). Another challenge is that EL images and the C-DCR images need to be aligned. This means the cells in the images need a common coordinate system. This is important to ensure that the location of the generated labels and the EL images match. This alignment is carried out using the internal "alignment tool" developed at Fraunhofer ISE in earlier machine learning projects. It detects the edges of the solar cell in the measurement images, rectifies the image, and crops it to the cell dimensions.

Not the entire image region can be used for labeling and defect detection. The regions with contact bars, as can be seen in Figure 2.2 in Chapter 2, are excluded as the measurements are not defined at those positions. The cell is separated into six regions of interest (ROIs). This is shown in Figure 4.2. The separation for the separate processing of each ROI.

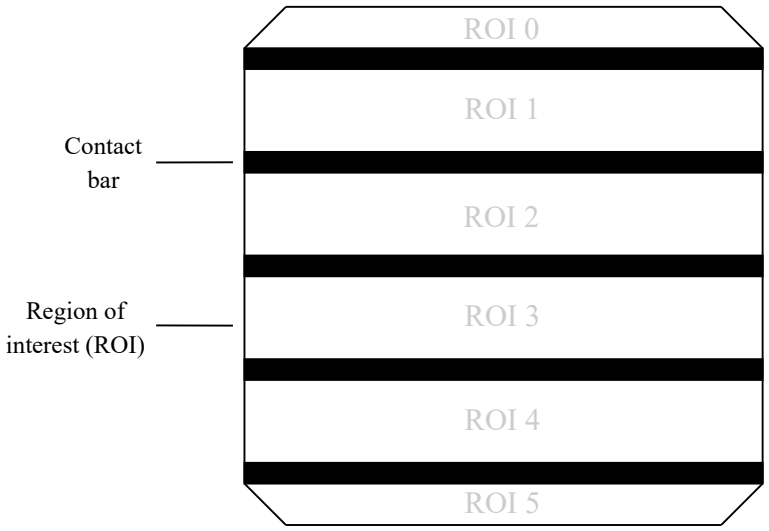


Figure 4.2: Concept of separating the solar cell in regions of interest (ROIs). The information at the contact bar locations cannot be used to label defects of the cell. The following figures will showcase the application of the respective algorithm to one ROI.

Figure 4.3 gives an overview of the developed Smart Labeling algorithm, where the left side shows the determination of the R_S and the right side of the j_0 relevant labels, respectively. The details of the flowchart are explained in the following sections.

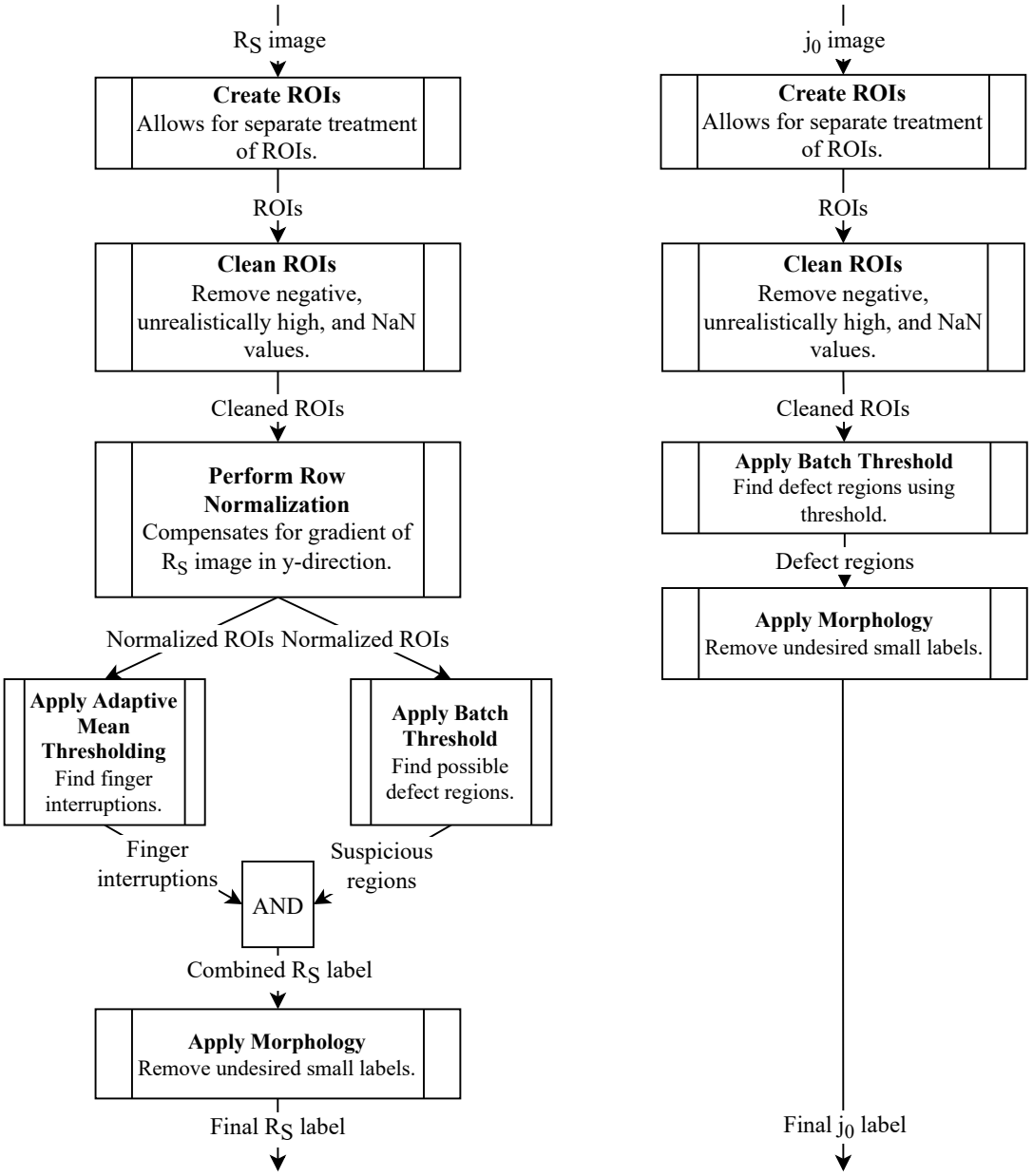


Figure 4.3: Detailed flowchart of Smart Labeling for both R_S and j_0 labels. While the R_S approach makes use of the geometric properties of finger interruptions in adaptive thresholding, the j_0 approach follows a simple global thresholding approach. Because the global thresholding is less robust against variations in the absolute pixel intensities $I(x, y)$, more care is required in the threshold definition for each cell batch.

4.3.2 R_S Labels

Note that the coordinate system in this section is rotated by 270° with respect to plots of other sections in this document. This leads to swapped x and y axes in this section but enables a more document-compatible presentation of figures that focus on one ROI.

Extracting finger interruptions from the R_S images of the C-DCR reference measurements comes

with challenges. Due to the limitations of the measurement technique and its post-processing, the image quality is rather low. Because of the camera properties and the contacting technique, the image shows brightness variations in the y direction while changes in the x direction can be attributed to finger interruptions.

To overcome the undesired variations in y direction, a technique called "row normalization" is developed. A sliding window of height H moves across each ROI with a step size s of $\frac{H}{2}$. The procedure is visualized in Figure 4.4. At each window position, the mean pixel intensity within the window is subtracted from each pixel inside the window. The result is copied to a new result image, i.e., the consecutive window positions are not affected by the subtraction. After applying row normalization, the global mean, i.e., the mean across all pixels of all ROIs, is added to every pixel of the resulting image to sustain the physical meaning of each pixel intensity.

Figure 4.4: Concept of row normalization. The sliding window of height H moves with step size s. The arithmetic mean of all pixel intensities within the sliding window is calculated. Then, the mean is subtracted from all pixels within the window.

On the "row normalized" images, thresholding is performed to find the finger interruptions exceeding a now comparable threshold value. Global thresholding is one of the easiest forms of separating the foreground and background of an image, i.e. in defect and non-defect areas, respectively. It applies a global threshold T_{global} to each pixel of an image as described by Equation 4.1.

$$b(x,y) = \begin{cases} 0, & \text{if } I(x,y) \geq T_{global} \\ 1, & \text{otherwise} \end{cases} \tag{4.1}$$

With:

$b(x,y)$: Binarized image

$I(x,y)$: Intensity of pixel at location (x,y)

However, in this case, this global thresholding is not sufficient due to variations in contrast and intensity in the R_S images between two contacting bars. Here the series resistance increases with increasing distance from the contact units. Therefore, local thresholding is applied that uses a spatially dependent threshold $T(x,y)$. It can generally be expressed by Equation 4.2 [52].

$$b(x,y) = \begin{cases} 0, & \text{if } I(x,y) \geq T(x,y) \\ 1, & \text{otherwise} \end{cases} \tag{4.2}$$

Due to high variations of the R_S between different solar cell batches, the local threshold adopts independently of the cell type. Local adaptive thresholding is a common technique in image processing, e.g., in the digitization of documents to separate letters from the paper background [52]. There are several implementations available, such as in OpenCV [53]. A common approach is to calculate the local arithmetic mean $m(x, y)$ of the pixel intensities within a window of size $W \times W$ centered around the pixel $P(x, y)$ at position (x, y) of intensity $I(x, y)$. $m(x, y)$ is then used to determine the local threshold $T(x, y)$ as shown in Equation 4.3.

$$T(x, y) = m(x, y) - C \cdot m(x, y) = (1 - C) \cdot m(x, y) \tag{4.3}$$

The constant C can be used to control the sensitivity of the thresholding. A negative C will increase the threshold, i.e., the detection becomes more selective.

The adaptive thresholding implementation of OpenCV in the function cv.adaptiveThreshold only allows a rectangular window for the mean calculation [53]. Therefore, for the local thresholding to find finger interruptions, a customized version of local adaptive thresholding is developed. Figure 4.5 illustrates the concept of adaptive thresholding applied to the detection of finger interruptions in R_S images. The customized version allows a custom window dimension $W \times H$. The window size is chosen based on the assumption that fingers are higher than wide. This is to make sure that the window always contains a sufficiently large share of areas without finger interruptions as shown in Figure 4.5. In the actual implementation, $W = 65\ px$ and $H = 11\ px$ is used on R_S images with a resolution of 1024×1024 pixels.

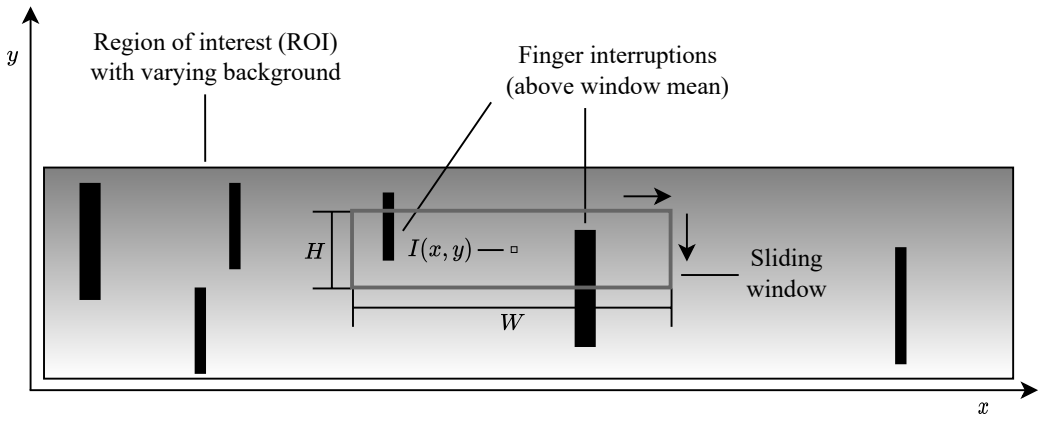


Figure 4.5: Concept of adaptive thresholding with rectangular window dimensions implemented to detect finger interruptions in R_S image. The center pixel $P(x, y)$ at position (x, y) of intensity $I(x, y)$ is well below the mean of the surrounding sliding window and would therefore be classified as "no finger interruption". When the window reaches finger interruptions as illustrated as black bars, the corresponding pixels would be classified as "finger interruption" as their intensity $I(x, y)$ lies well above the window mean.

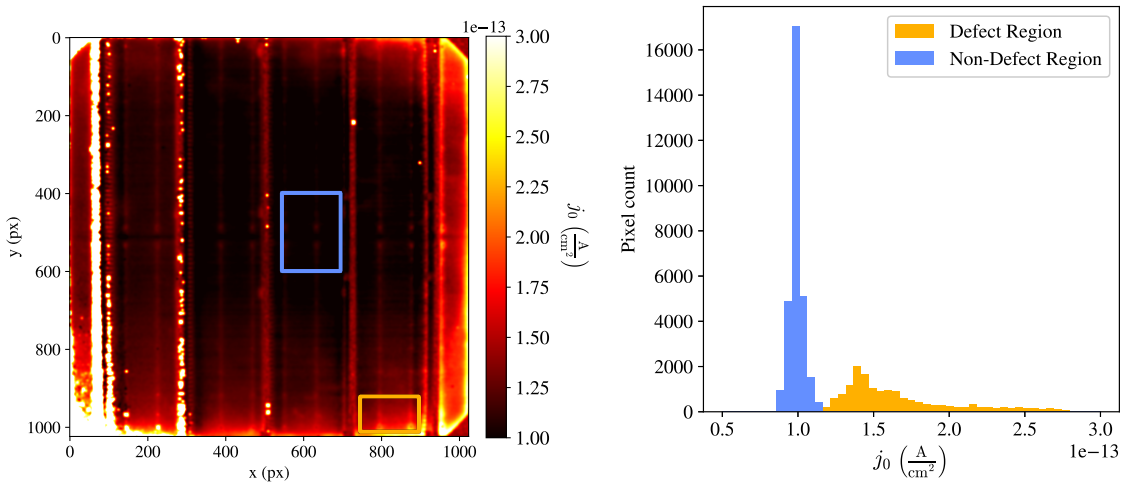
As shown in the flowchart in Figure 4.3 a global threshold is applied in parallel to the local adaptive thresholding. This threshold can either be defined for each batch of solar cells or one global threshold can be used. In total, this combined approach can be considered local adaptive thresholding with a global threshold as a boundary condition. The global threshold finds potential R_S

defects. From potential R_S defects real defects are filtered based on local adaptive thresholding.

4.3.3 j_0 Labels

The approach to extract labels from j_0 images is different compared to the R_S related one as it cannot rely on the geometric properties of the defect. However, this poses challenges as well. As shown on the right side in Figure 4.3, a batch threshold is applied. The arithmetic mean of pixel intensities varies significantly between the different TOPCon cell batches. This means, that one threshold valid for a batch of cells has to be defined. Several approaches to defining those thresholds automatically have been tested. Among others, defining a threshold based on statistical properties, e.g., a deviation from the mean pixel value across the batch, has been tested. However, a manual definition following a clearly defined approach has proven to be the most robust. This section presents this developed approach.

As shown in Figure 4.6, the batch threshold is defined based on a histogram plot for a defective and a non-defective region in a j_0 image of at least one cell from a particular batch. On the y-axis, the histogram shows the number of pixels with the pixel value within the respective bin that is shown on the x-axis. In many cases, it is possible to observe a clear separation between the value ranges of the defect and non-defect regions.



(a) j_0 image of cell 210015ISE0616 with defect region and non-defect region

(b) Analyzing histograms of defect region and non-defect region of cell 210015ISE0616

Figure 4.6: Example for j_0 threshold definition from a sample image of a batch. A defective region and a non-defective region are chosen. Then, the histograms of those regions are plotted. The threshold is chosen at a j_0 value at which the two histograms begin to overlap only slightly. In this example, this would be around $1.2\text{e}-12\ \frac{A}{cm^2}$.

For every sample of each batch, the corresponding batch threshold is applied followed by a morphological operation on the resulting label mask. In particular, this includes the removal of very small defective areas that are smaller than a defined threshold. This operation is performed because very small defective areas can be caused by noise in the j_0 image.

4.4 Network Architecture

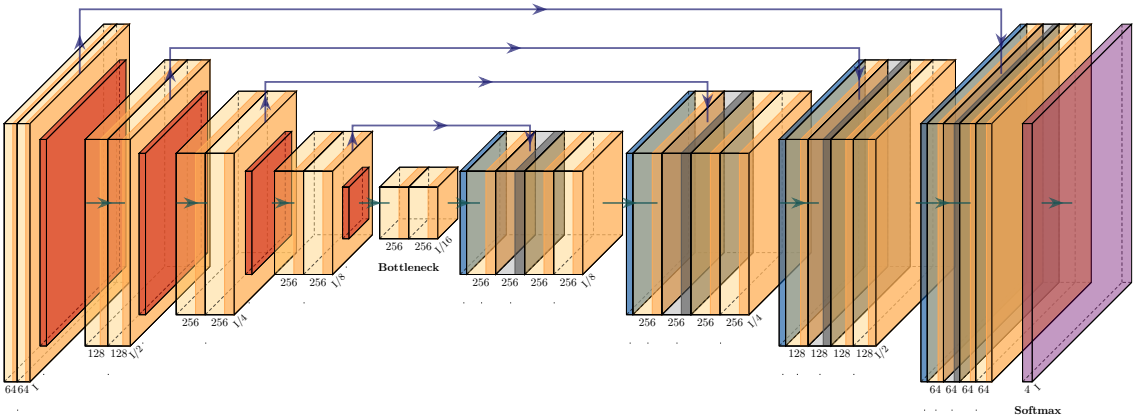


Figure 4.7: U-Net-based architecture applied in all experiments. Plotted using PlotNeuralNet [54]. Zoom in to read the dimensions of layers.

In all experiments, a modified U-Net architecture is used as an encoder and decoder network. The U-Net is a fully convolutional network and one of the most commonly applied models to perform semantic segmentation tasks [21]. Originating from biomedical image processing, it has successfully been applied in the field of PV before [5], [51].

The models architecture is visualized in Figure 4.7. It consists of a contracting path (left side) and an expansive path (right side) [21]. The contracting part compresses the images semantically with respect to the target variable, i.e., defect labels, while the expansive part assigns the extracted and compressed defect information to the corresponding pixels making use of skip connections. The input image of size $I \times I$ is fed into the contracting path. Each block in the contracting path consists of two 3×3 convolutional layers with a Rectified Linear Unit (ReLU) activation function. After each block, a 2×2 max pooling layer reduces the dimensions of the feature maps until the bottleneck is reached. At the *bottleneck*, the feature maps have a size of $I/16$. In the case of EL images with a resolution of 256x256 pixels, this equals a feature map size of 16×16 pixels. In contrast to the original U-Net architecture, the used architecture uses a maximum of 256 feature map channels whereas the original goes up to 1024 channels in the bottleneck. This reduces the number of trainable parameters and therefore model complexity.

After the bottleneck, the expansive path begins. The blocks in the expensive parts begin with a transposed convolution layer increasing the dimensions of the feature maps. The result of the transposed convolution is concatenated with the output of the corresponding block with the same feature map dimensions from the contracting path. These connections, indicated by blue arrows in Figure 4.7, are called skip connections. The concatenated feature maps are passed through two 3×3 convolutional layers followed by a ReLU activation function. With each block in the expansive path, the dimensions of the feature map double. This is repeated until the dimensions equal the input dimensions $I \times I$. In both the contracting path and the expansive path batch normalization [55] is applied after each block.

The output segmentation map has four channels for each class: No R_S defect, R_S defect, No j_0 defect, j_0 defect. Finally, a Softmax layer is applied that converts the real numbers from the output segmentation map into a probability map. Each channel of the probability map assigns each pixel a probability to belong to the respective class.

5 Technical Experiments

5.1 General

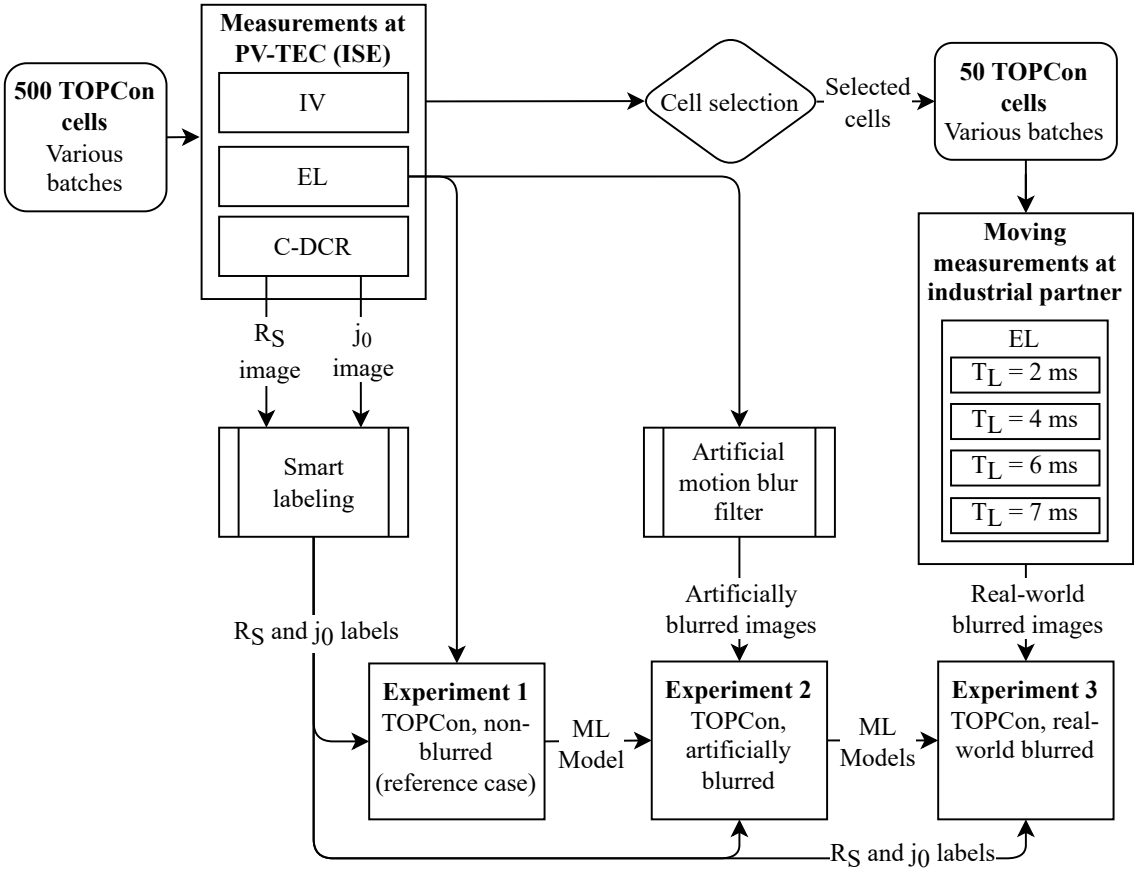


Figure 5.1: The full experiment workflow providing an overview. Includes measurements at ISE and consecutive Smart Labeling as well as measurements at industrial partner.

5.1.1 Neural Network Settings

This section states all neural network design choices and details that apply to all experiments. All networks are trained on an Nvidia TITAN Xp with 12 GiByte of memory. They are implemented in the open source machine learning framework PyTorch [20]. In all experiments, the cross entropy loss [56] function is used. It is a common choice for a multi-class classification task. The loss function calculates the difference between the output of the neural network and the target labels. The class balance parameter is used to take the imbalance of classes into account. The underlying assumption is that there are more non-defective pixels in the EL images than there are pixels that are classified as defective. If this is not considered in the loss function, the network may be incentified to predict no defect as this is correct in most cases.

The Adam optimization algorithm [57] is used to calculate the updated trainable neural network parameters based on the value of the loss function. One of the most important parameters of an optimization algorithm is the learning rate, i.e., the step size with which the trainable parameters (weights) of the neural networks are updated with. Weight decay, also called L_2 regularization, is a technique aiming at penalizing complexity by avoiding that the weights of the neural network grow to large [19]. The learning rate and weight decay settings will be varied in experiment 1 within a grid search. The results of Experiment 1 determine the learning rate and weight decay choices for the consecutive experiments.

5.1.2 Evaluation Metrics

Quantitative Evaluation

The performance of each model is evaluated on the validation set using the F_1 score. To understand the F_1 score, the concepts of precision and recall are required knowledge.
Precision can be regarded as a measure of how many of the positive predictions are actually correct (Equation 5.1).

$$Precision = \frac{TP}{TP + FP} \tag{5.1}$$

With:

TP : Number of true positive predictions

FP : Number of false positive predictions

Recall is a measure of how many actual positives are correctly identified by the algorithm (Equation 5.2).

$$Recall = \frac{TP}{TP + FN} \tag{5.2}$$

With:

FN : Number of false negative predictions

Applied to the task at hand, a model not predicting any defects in the solar cell where no defects are present has a high precision of 1. A model that finds all actual defects has a high recall of 1. However, this results in two separate scores but both of them should be maximized for ideal model performance. For this reason, the F_1 score exists. It is the harmonic mean of precision and recall, see Equation 5.3 [19], [58]. For the interested reader, a more detailed definition and investigation is provided by Sasaki (2007) [58].

$$F_1 = \frac{2 \cdot precision \cdot recall}{precision + recall} = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \tag{5.3}$$

The F_1 score is suitable to evaluate algorithms performing binary classification tasks. Therefore, the two defect classes R_S and j_0 are evaluated individually. To do so, the respective F_1 score for both binary classification tasks (R_S defect, no R_S defect as well as j_0 defect, no j_0 defect) is calculated.

In the investigation, the F_1 score is calculated individually for each cell. The reported F_1 score is the arithmetic mean of all F_1 scores across all cells of the validation set, and the test set, respectively. The same holds for any reported precision and recall values.

In model selection, the F_1 score on the validation set is used at a decision threshold of 0.5 the respect to the predicted defect probability. This means if the predicted probability for a pixel belonging to a defect class is above 0.5, it is classified as a defect. It could be reasonable to use the value of the loss function (the loss) as well, however, the F_1 score can be regarded as a more accurate measure of how well the model performs on the actual task. The loss curves are still plotted for the selected scenarios to ensure that the loss is also low at the chosen training epoch. Usually, a low loss function value also corresponds with a high F_1 score. This will be discussed in greater detail in Chapter 6 (Technical Results).

Qualitative Evaluation

For each experiment, the results for one full solar cell from the test set are analyzed qualitatively. For this, the EL image of the cell, the labels for both R_S and j_0 defects, and the corresponding predictions of the network are visualized. In the qualitative analysis, one of the best performing scenarios is used at a decision threshold of 0.7, i.e., a pixel is classified as a defect if the neural network predicts a probability for the defect class of 0.7 or higher. Depending on the experiment, the performance on R_S defects is analyzed more rigorously at different degrees of motion blur. This is done by investigating the performance of the algorithm in specified image regions.

5.2 Experiment 1: Still, Non-Blurred

5.2.1 Goals

In the first experiment, the dataset relevant for all following experiments is created, and the entire experiment workflow is established. This includes the first training of a neural network for defect detection on still, non-moving EL images. The hyperparameters for the neural network training determined in this experiment serve as a reference for all following experiments.

5.2.2 Measurements at Fraunhofer ISE

This project focuses on a very diverse set of TOPCon cells with a large variety of defects. The project includes taking measurements in the laboratory at Fraunhofer ISE as there is currently no training data available using TOPCon cells. In the laboratory of Fraunhofer ISE, cells can be produced under pre-industrial conditions in small batches, often 10-100 cells with varying process parameters in the different manufacturing processes. These cells are well-suitable to train a robust defect detection algorithm, as the defect variety is much higher than in an industrial production process where process parameters are much more constant.

All measurements to create the dataset for this experiment are carried out on the *PV-TEC cell tester*, a modern highly automized measurement system for solar cells. It can measure so-called

carriers with 100 solar cells once at a time.

The EL images are taken on a measurement system from *halm elektronik GmbH*. The images are taken at an illumination time T_L of 45 ms and a camera gain of 3. The current is set to 10 A. A higher illumination time with a constant gain and current results in brighter images because more light reaches the sensor, and a lower illumination time in turn results in darker images. The measurement configuration is chosen in a way that the images are not saturated even with the cells that appear the brightest in the measurements, i.e., the brightest cell determines the upper limit of the illumination time. This means that no pixels of the EL image exceed the maximum intensity value of 255 (at an 8-bit image).

The C-DCR images are taken on a measurement system from ISRA Vision. The device is customized for Fraunhofer ISE. The C-DCR settings are organized in so-called *recipes*. The measurements are taken with a recipe developed for Heterojunction (HJT) cells, i.e., it is not adapted to TOPCon cells. There is no TOPCon-specific recipe available at this point in time. The PL measurements as part of the C-DCR technique are taken at a flash power of 36 W, equal to 1 sun, which represents standard test conditions (STC).

In total, 509 TOPCon cells are measured. Some of the cells are among the first TOPCon cells ever produced at Fraunhofer ISE. The cell architecture differs among the batches, e.g., some have no busbars and some have multi-busbars. This leads to different optical characteristics, posing additional challenges to the defect detection algorithm.

5.2.3 Dataset

Some of the 509 measured cells show such severe defects that render them unsuitable for meaningful defect detection. For the neural network experiments, only cells meeting the selection criteria open circuit voltage $V_{oc} < 630\ mV$ and conversion efficiency $\eta > 17.5\,\%$ are chosen. Cells with a lower η show a high number of finger interruptions mostly due to errors in the metalization process, most likely in the screen printing process to produce the contacts. It is difficult to label these cells correctly because the cells show dark areas across the entire cell. If produced in industry, these cells would be considered defective in every region. It should be avoided that inconsistent labels impact the training process negatively. Within the dataset of 509 cells, around 380 cells fulfill the selection criteria.

The dataset of 380 cells is randomly sub-divided into training (70 %), validation (15 %), and test (15 %) subsets. The training dataset is used to train the neural network. The validation dataset is used to select the best-performing training configuration and model. Finally, the test dataset is used to rate the final performance of the selected models.

5.2.4 Scenario Definition

The goal of this experiment is to find suitable parameters for training networks. Several scenarios with different neural network parameters are defined. All scenarios are trained for 100 epochs. The batch size is 8. The network architecture described in Section 4.4 used. All measurement images are scale to a size of 256×256 px^2.

In the scenarios, the learning rate, weight decay, class balance, and the padding region parameter (explained in Experiment 2, Figure 5.2) are varied. Because the training of neural networks is computationally expensive, only a limited number of scenarios can be tested. Two different learning rates, two different settings for the class balance, and weight decay and no weight decay are tested in all combinations. This results in 16 scenarios in total.

Table 5.1: Defined scenarios for still measurements.

scenario id	lr	wd	class balance	padding region
0	0.00010	0.001	[0.2, 0.8]	True
1	0.00010	0.001	[0.2, 0.8]	False
2	0.00010	0.001	[0.05, 0.95]	True
3	0.00010	0.001	[0.05, 0.95]	False
4	0.00010	0.000	[0.2, 0.8]	True
5	0.00010	0.000	[0.2, 0.8]	False
6	0.00010	0.000	[0.05, 0.95]	True
7	0.00010	0.000	[0.05, 0.95]	False
8	0.00001	0.001	[0.2, 0.8]	True
9	0.00001	0.001	[0.2, 0.8]	False
10	0.00001	0.001	[0.05, 0.95]	True
11	0.00001	0.001	[0.05, 0.95]	False
12	0.00001	0.000	[0.2, 0.8]	True
13	0.00001	0.000	[0.2, 0.8]	False
14	0.00001	0.000	[0.05, 0.95]	True
15	0.00001	0.000	[0.05, 0.95]	False

5.3 Experiment 2: Artificial Motion Blur

5.3.1 Goals

The goal of this experiment is to investigate how the approach from Experiment 1 can cope with the presence of motion blur. In the initial phase of the project, there is no dataset with real motion blur available. Therefore, a technique to add realistic motion blur is developed in this experiment. Additionally, the models that are trained in this experiment are used in Experiment 3 to perform transfer learning. This is to investigate if better results on the data with real motion blur can be achieved if a model is trained with artificial motion blur.

Furthermore, the impact of adding a "padding region" is tested in this experiment. In Figure 5.2, a still EL measurement is shown on the left and a moving measurement is shown on the right. The size of the image region containing information increases when motion blur is present as the information blurs against the direction of movement, highlighted by the red rectangle. Therefore, additional space, called *padding region*, is added to the bottom of the image giving the models the opportunity to gather information from a wider region and prevent information loss.

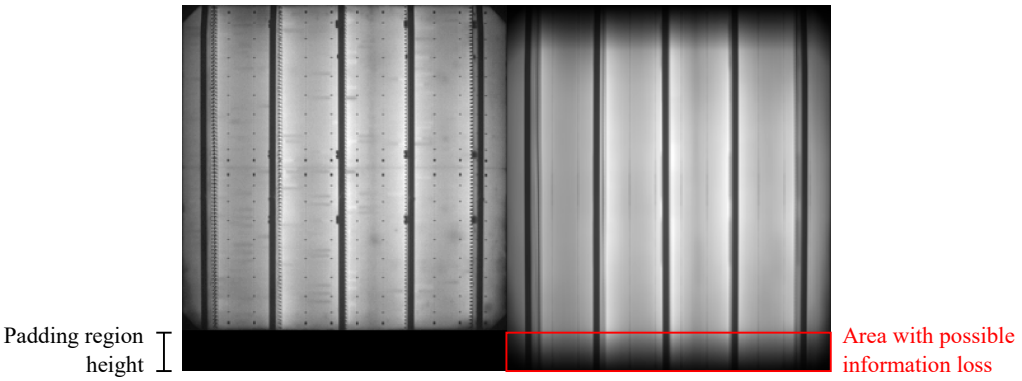


Figure 5.2: Illustration of the padding region approach. The red region in the blurred right image shows the area containing information that would be lost in case the image would be kept at its original dimensions.

As a reference, some scenarios have been trained with a padding region in Experiment 1. This allows isolating the impact of the padding region on the detection performance at different degrees of motion blur.

5.3.2 Dataset

This experiment uses the exact same dataset as Experiment 1 in Section 5.2. In the different scenarios of this experiment, motion blur corresponding to different EL illumination times is applied as presented in the following section.

5.3.3 Estimation of Realistic Motion Blur

Because only limited real-world data is available (only 50 cells can be measured due to high manual effort), it is important to understand the characteristics of the motion blur originating from the movement of the cell during the measurement.
A simplified mathematical model is used to generate samples with artificial motion blur. It can be described by Equation 5.4, where $*$ denotes the convolution operator. The blur kernel $h(x, y)$ is a two-dimensional matrix that is convolved with the input image $f(x, y)$ [59].

$$g(x, y) = f(x, y) * h(x, y) + n(x, y) \tag{5.4}$$

with:

$g(x, y)$: Blurred image

$f(x, y)$: Original (unblurred) image

$h(x, y)$: Blur kernel

$n(x, y)$: Additive noise

$$h(x,y) = \begin{cases} \frac{1}{l}, & \text{if } 0 \leq |x| \leq l \cdot cos(\theta), |y| = l \cdot sin(\theta) \\ 0, & \text{otherwise} \end{cases} \tag{5.5}$$

In this case, it holds that the blur angle is zero, i.e., $\theta = 0$. In this application, the model from Equation 5.4 is simplified by ignoring any further noise of the capturing process of the camera, i.e., $n(x, y)$ is not considered. Equation 5.6 can be used to estimate the approximate range l of motion blur in the image taken at a given illumination time T_L when the object is moving with the belt speed v_b [60].

$$l = v_b \cdot T_L \tag{5.6}$$

This physical length can then be converted into the number of affected pixels (equal to the kernel size K if the pixel size is known. The pixel size P stands for the physical distance a single pixel represents.

$$K = \frac{l}{P} \tag{5.7}$$

Combining Equations 5.6 and 5.7, the kernel size K in pixels can easily be estimated:

$$K = \frac{v_b \cdot T_L}{P} \tag{5.8}$$

To give an example, the maximum kernel size in these experiments can be estimated at the maximum illumination time of this experiment ($T_{L_{max}} = 7\ ms$):

$$K_{max} = \frac{1.9\ \frac{m}{s} \cdot 7\ ms}{160 \frac{\mu m}{px}} = 83.125\ px \tag{5.9}$$

Finally, some simplified experiments are carried out to compare real-world blurred images with artificial motion blur. The comparison is carried out both visually and based on the structural similarity index (SSIM) [61]. The analysis reveals that when the blur kernel $h(x, y)$ is only filled with '1s' from $y = \frac{K}{2}$ to K, the real motion blur can be estimated more accurately. The resulting blur is then uni-directional. Figure 5.2 shows the application of this motion blur kernel at $T_{L_{max}} = 7\ ms$.

5.3.4 Scenario Definition

In this experiment, mainly the padding region is switched on or off. Additionally, it is investigated whether using a pre-trained model from Experiment 1 or training the network from scratch yields better results.

Table 5.2: Defined scenarios for artificial measurements.

scenario id	pretrained switch	lr	class balance	T_L (s)	padding region
0	True	0.0001	[0.2, 0.8]	0.002	True
1	True	0.0001	[0.2, 0.8]	0.002	False
2	True	0.0001	[0.2, 0.8]	0.004	True
3	True	0.0001	[0.2, 0.8]	0.004	False
4	True	0.0001	[0.2, 0.8]	0.006	True
5	True	0.0001	[0.2, 0.8]	0.006	False
6	True	0.0001	[0.2, 0.8]	0.007	True
7	True	0.0001	[0.2, 0.8]	0.007	False
8	False	0.0001	[0.2, 0.8]	0.002	True
9	False	0.0001	[0.2, 0.8]	0.002	False
10	False	0.0001	[0.2, 0.8]	0.004	True
11	False	0.0001	[0.2, 0.8]	0.004	False
12	False	0.0001	[0.2, 0.8]	0.006	True
13	False	0.0001	[0.2, 0.8]	0.006	False
14	False	0.0001	[0.2, 0.8]	0.007	True
15	False	0.0001	[0.2, 0.8]	0.007	False

5.4 Experiment 3: Moving, Real-World Blurred

5.4.1 Goals

The goal of the third experiment is to investigate whether the developed algorithm also works on samples with real-world motion blur and which trade-offs should be considered in the settings of moving EL measurements. Finally, it is assessed if using models pre-trained on still images or pre-trained on images with artificial motion blur has a significant impact on detection performance.

5.4.2 Measurements at Industrial Partner

For the real-world measurements with motion blur, only 50 TOPCon cells can be measured because those measurements require a lot of manual work, because of manual contacting of the cell, changing measurement parameters, alignment of images, etc. Measuring one cell with four different illumination times takes around 5 minutes including contacting and decontacting. Considering other overhead, such as getting familiar with the measurement device prototype, 50 cells can realistically be measured per working day. Constraining the required time is important because visiting the industrial partner company requires a business trip including a hotel stay. The 50 cells are selected based on the measurements for Experiment 1.

Initial experiments at the industrial partner company revealed that very faulty cells appear too dark in EL images, especially in moving measurements. The partner company suggested using the grid resistance and the open circuit voltage V_{oc} as indicators for the cell selection. The grid resistance is a measure of the resistance of the metal grid pattern on the front of the solar cell. A high grid resistance can indicate a faulty metalization, i.e., the contacts are faulty. This both inhibits measuring the cell and using the cell in-field to produce electricity because the current flow is inhibited by the high grid resistance. The chosen indicators $R_{grid_{front}}$ and $R_{grid_{rear}}$ are used by the measurement system at Fraunhofer ISE but not industry standard. Every other indicator to ensure the resistance of the cell is not too high would be suitable as well.

To select around 50 cells, four cell categories are defined as shown in Table 5.3.

Table 5.3: Cell classes for measurements at industrial partner company.

Class 1

- $R_{grid_{front}} < 160\ \frac{\Omega}{m}$
- $R_{grid_{rear}} < 130\ \frac{\Omega}{m}$
- $V_{oc} \geq 690\ mV$
- **12 cells**

Class 2

- Similar R_{grid} as Class 1
- $687 \geq V_{oc} \geq 690\ mV$
- **25 cells**

Class 3

- Manual selection by visual inspection, less strict boundaries for R_{grid} and V_{oc}
- High diversity: Round defects, dark edge regions, distributed finger interruptions, cracks
- **6 cells**

Class 4

- Similar to Class 3, even more severe defects based on visual inspection
- **9 cells**

The four classes aim at covering a large variety of defects that occur across the entire dataset. The values are chosen by observing the distribution of the values of the three decision variables across the dataset. The selection criteria have been chosen in a way that ensures that 50 samples are available for the moving measurements, even though all cells cannot be considered "perfect" by industrial standards.

The measurements with the prototype at the industrial partner company should cover different measurement settings with different illumination times. This is to investigate the trade-off between different settings. On the one hand, it must be considered that the lower the illumination time, the darker the image. The darker the image, the harder it becomes to separate useful information, i.e. the signal, from the noise in the image, i.e. the signal-to-noise ratio decreases. On the other hand, the higher the illumination time, the higher the motion blur that occurs due to the movement of the cell. Therefore, each of the 50 cells is measured with four different measurement settings, so-called recipes, as defined in Table 5.4.

Table 5.4: Measurement recipes for moving measurements at industrial partner.

Recipe	T_L (ms)	G	I (A)
1	2	4	20
2	4	2	20
3	6	3	10
4	7	2	10

T_L : EL illumination time
G : EL camera gain (multiplier)
I : Current applied to cell

5.4.3 Dataset

From the 50 measured cells, only 38 cells are available for the neural network experiments. This is because some cells have been damaged when they have been measured again at Fraunhofer ISE for new reference measurements. The dataset of 38 cells is divided into training (60 %), validation (20 %), and test (20 %) subsets. The shares differ from the previous experiments because of the low number of samples. Using validation and test shares of 20 %, at least 7-8 samples are available to judge the model's performance. This is a rather low number for machine learning standards. However, it has to be considered in semantic segmentation that every pixel is subject to a classification task. Even with 8 samples, at a resolution of 256×256 pixels, there are still $8 \cdot 256 \cdot 256 = 524,288$ classification tasks carried out that are evaluated by the quantitative metrics.

5.4.4 Scenario Definition

In scenarios 0-3, a pre-trained model is loaded that is trained on artificial motion blur assuming the same illumination time of the real images as shown in Table 5.5. Furthermore, in scenarios 4-7, the best performing pre-trained model trained on still images is loaded. This is to evaluate if training images with artificial motion blur on a higher number of samples has a positive impact on the model's performance. All other network parameters are the same as the parameters of the best-performing model from Experiment 1.

Table 5.5: Defined scenarios for real measurements.

scenario id	pretrained switch	lr	class balance	T_L (s)	top padding
0	Artificial	0.0001	[0.2, 0.8]	0.002	True
1	Artificial	0.0001	[0.2, 0.8]	0.004	True
2	Artificial	0.0001	[0.2, 0.8]	0.006	True
3	Artificial	0.0001	[0.2, 0.8]	0.007	True
4	Still	0.0001	[0.2, 0.8]	0.002	True
5	Still	0.0001	[0.2, 0.8]	0.004	True
6	Still	0.0001	[0.2, 0.8]	0.006	True
7	Still	0.0001	[0.2, 0.8]	0.007	True
8	False	0.0001	[0.2, 0.8]	0.002	True
9	False	0.0001	[0.2, 0.8]	0.004	True
10	False	0.0001	[0.2, 0.8]	0.006	True
11	False	0.0001	[0.2, 0.8]	0.007	True

6 Technical Results

6.1 Experiment 1: Still, Non-Blurred

6.1.1 Quantitative Results

Validation Set Performance and Model Selection

Table 6.1: Results of the still measurement scenarios on the validation set. These results are used to select two of the best performing models, one with and one without padding region. The two chosen scenarios are highlighted in colors corresponding to the scenario plots in Figures 6.1, 6.2, and 6.3.

scenario id	padding region	best F_1 R_S	best epoch R_S	best F_1 j_0	best epoch j_0
0	True	0.691	45	0.845	43
1	False	0.692	54	0.848	50
2	True	0.612	99	0.818	46
3	False	0.597	98	0.815	99
4	True	0.683	44	0.847	50
5	False	0.694	46	0.840	49
6	True	0.598	46	0.817	58
7	False	0.587	52	0.804	58
8	True	0.667	99	0.826	99
9	False	0.670	86	0.836	99
10	True	0.558	85	0.773	99
11	False	0.535	99	0.750	99
12	True	0.663	73	0.837	99
13	False	0.669	78	0.833	99
14	True	0.538	53	0.741	88
15	False	0.550	81	0.758	99

Based on the validation set, two models have been selected for testing and further use in the following experiments. As can be seen from Figure 6.1, the class balance parameters has the most significant influence on the F_1 score for R_S defects. Scenario 5 using a class balance of [0.2, 0.8] is the best performing model without padding region achieving an F_1 score for R_S defects of 0.694. Changing the class balance in scenario 7 to [0.05, 0.95] alone reduces the F_1 score to 0.598. The impact of the class balance parameter on the R_S detection performance can be seen well in Figure 6.1. Scenario 4 is the best performing scenario with padding region using a class balance of [0.2, 0.8]. Scenario 6 used exactly the same settings except for changing the class balance to [0.05, 0.95], also significantly worsening the performance. The learning rate has the second most significant influence with 0.0001 performing better than 0.00001. Only in very few cases weight

decay has a positive influence.

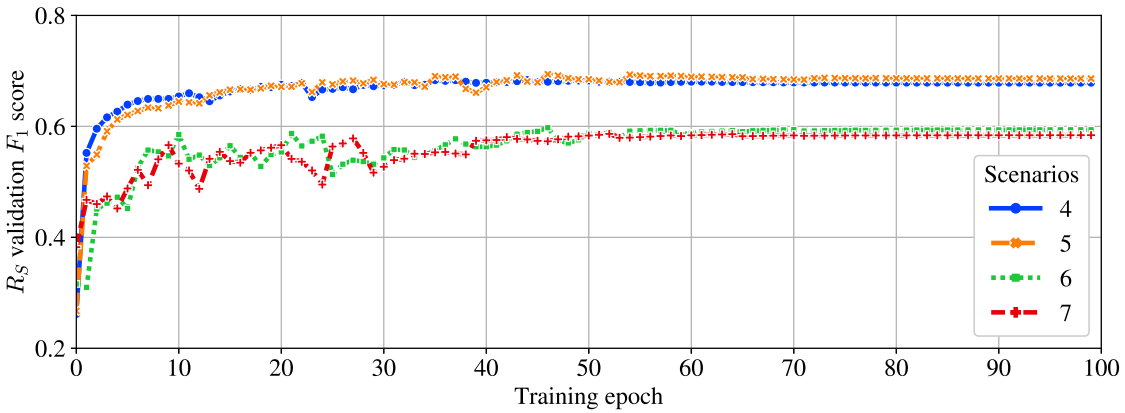


Figure 6.1: Validation accuracy for R_S defects after each training epoch. Plot compares scenarios 4 and 5 (best results with and without padding region) and scenarios 6 and 7 (only changed class balance to [0.05, 0.95]). It can be seen that the class balance parameter has a significant influence on the R_S F_1 score.

The detection of j_0 defects is much less sensitive towards parameter settings as shown in Figure 6.2. All scenarios show a very similar performance.

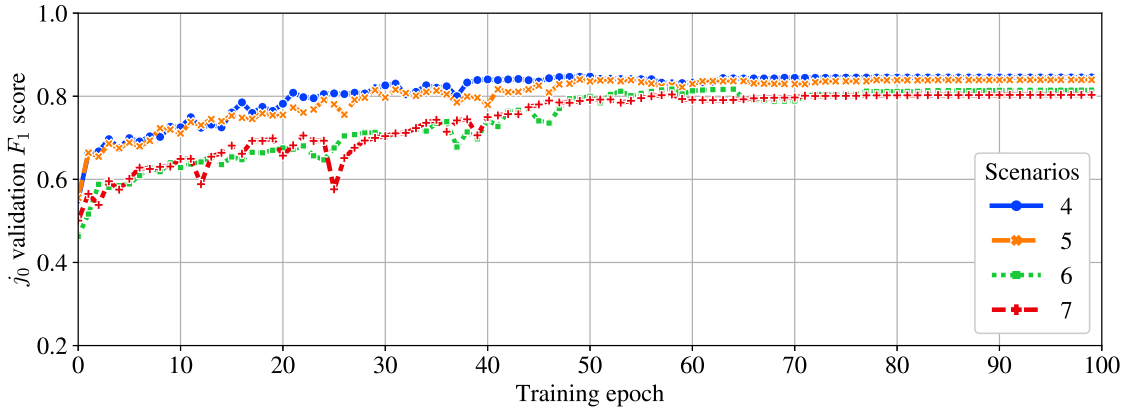


Figure 6.2: Validation accuracy for j_0 defects after each training epoch. The plot compares scenarios 4 and 5 (best results with and without padding region) and scenarios 6 and 7 (only changed class balance to [0.05, 0.95]). The influence on the j_0 F_1 score is rather low.

The models are sensitive toward overfitting. Figure 6.3 shows that the value of the combined loss function on the validation set increases for all plotted scenarios after around the 50$^{\text{th}}$ epoch. This is because the models learn the training dataset too well (overfitting), negatively impacting generalizability.

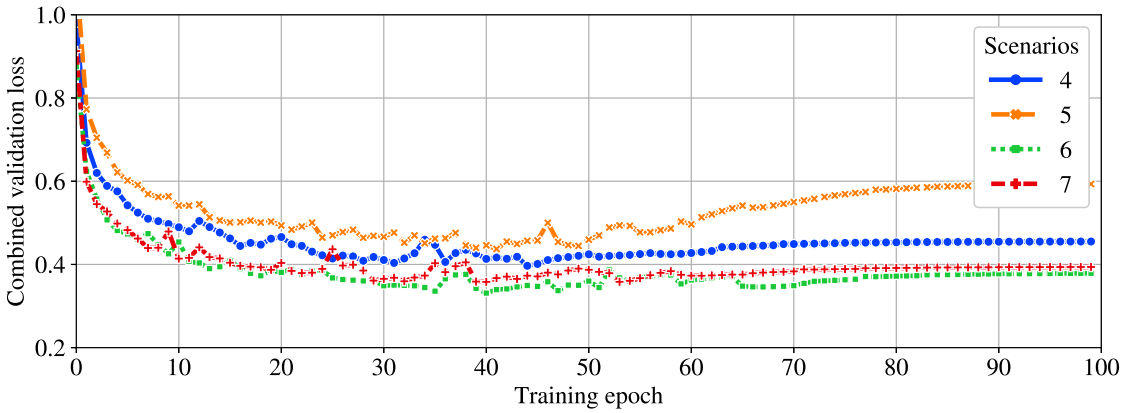


Figure 6.3: Combined validation loss for j_0 defects after each training epoch. The plot compares scenarios 4 and 5 (best results with and without padding region) and scenarios 6 and 7 (only changed class balance to [0.05, 0.95]). Especially for scenario 5, the validation loss increases again after the 50$^{\text{th}}$ epoch which indicates overfitting.

Because detecting R_S defects can be considered the more challenging task, the models are selected based on their performance on R_S defects. In particular, the models from scenarios 4 and 5 are selected based on their good performance on the R_S defects. Due to their performance on the validation set, scenario 4 (with padding region) and scenario 5 (without padding region) are selected for evaluation on the test set as well as for transfer learning in later experiments.

Even though scenario 0 with weight decay performs slightly better by a very small margin than scenario 4 without weight decay, scenario 4 is chosen to limit the complexity of parameters in consecutive experiments. For both scenarios, the model from the training epoch performing best on R_S defects is selected.

Test Set Performance

Table 6.2: Results on the test set for selected scenarios of experiment 1. P: Precision, R: Recall, F1: F_1 score

scenario id	T_L (s)	class threshold	P R_S	R R_S	F1 R_S	P j_0	R j_0	F1 j_0
4	0	0.3	0.493	0.921	0.636	0.743	0.969	0.826
5	0	0.3	0.522	0.886	0.651	0.718	0.947	0.798
4	0	0.5	0.593	0.852	0.695	0.799	0.940	0.850
5	0	0.5	0.622	0.805	0.695	0.787	0.906	0.823
4	0	0.7	0.698	0.739	0.715	0.853	0.891	0.857
5	0	0.7	0.727	0.687	0.698	0.855	0.832	0.823
4	0	0.9	0.850	0.486	0.608	0.920	0.751	0.810
5	0	0.9	0.877	0.435	0.565	0.938	0.614	0.715

In the evaluation on the test set, the analysis is carried out at different decision thresholds. The meaning of the decision threshold is explained in Section 5.1.2. From Table 6.2 it can be seen that

the F_1 score for the R_S defect is highest on either a decision threshold of 0.5 or 0.7. A decision threshold of 0.7 is used for the discussion of qualitative results in all consecutive experiments and analyses.

6.1.2 Qualitative Results

Analysis of Whole Cells

Figure 6.4 and Figure 6.5 show that the network from scenario 4 is capable of predicting defects in still EL images of solar cells. The upper left image is the EL image as it is passed to the network. In both figures, the two black images in the upper row show the labels for R_S and j_0 defects. White areas represent defects. The two black images in the bottom row show the defects as predicted by the network. The images show the performance on two cells from different batches selected for a qualitative analysis. In both cases, the j_0 prediction looks very similar to the corresponding label. Severe R_S defects that are clearly visible in the EL images are reliably detected as well.

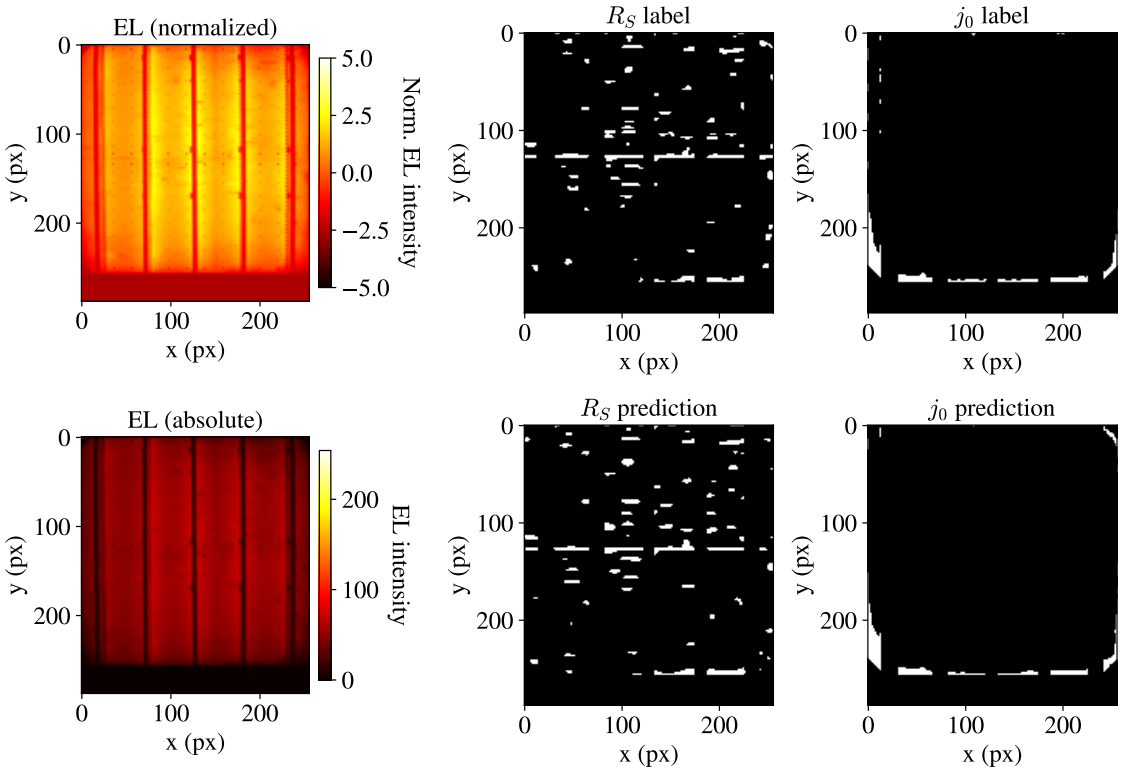


Figure 6.4: Performance of algorithm in scenario 4 on cell 210015ISE0748 at a decision threshold of 0.7.

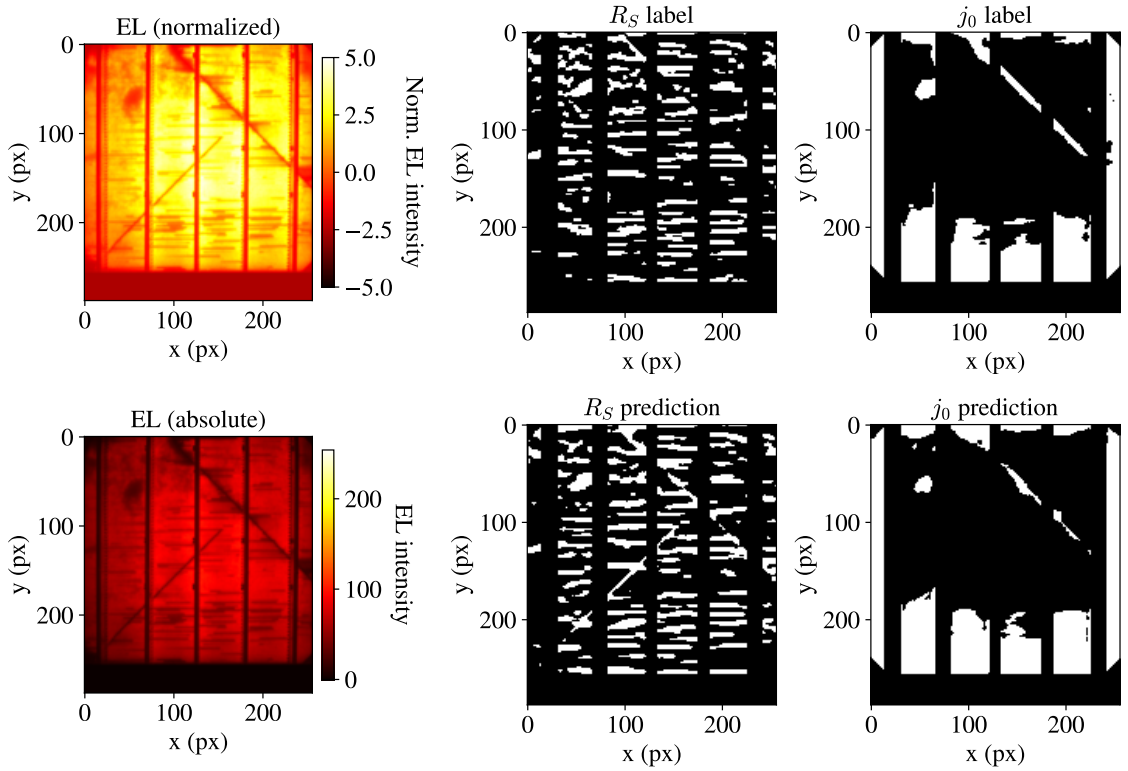


Figure 6.5: Performance of algorithm in scenario 4 on cell 210266ISEn337 at a decision threshold of 0.7. This cell shows a crack which is correctly predicted as j_0 and partially as R_S defect by the algorithm.

6.2 Experiment 2: Artificial Motion Blur

6.2.1 Quantitative Results

Validation Set Performance and Model Selection

Similar to Experiment 1, the scenarios with the highest F_1 score for R_S defects on the validation set is selected for both the qualitative analysis and to transform transfer learning in consecutive experiments. For each illumination time, one scenario is selected based on Table 6.3.

Table 6.3: Results of the artificial motion blur scenarios on the validation set. All results are based on a decision threshold of 0.5. These results are used to select the best-performing models for each illumination time. The four chosen scenarios are highlighted in colors corresponding to the scenario plots in Figures 6.6, 6.6, and 6.8.

scenario id	T_L **(ms)**	**padding region**	**pre-trained**	**best** F_1 R_S	**best epoch** R_S	**best** F_1 j_0	**best epoch** j_0
0	0.002	True	True	0.646	18	0.838	27
1	0.002	False	True	0.641	23	0.842	47
2	0.004	True	True	0.615	32	0.839	26
3	0.004	False	True	0.604	30	0.835	48
4	0.006	True	True	0.614	24	0.846	26
5	0.006	False	True	0.584	33	0.844	31
6	0.007	True	True	0.607	29	0.842	31
7	0.007	False	True	0.572	25	0.841	29
8	0.002	True	False	0.638	48	0.841	49
9	0.002	False	False	0.630	38	0.839	49
10	0.004	True	False	0.592	49	0.814	49
11	0.004	False	False	0.586	48	0.827	49
12	0.006	True	False	0.589	49	0.825	49
13	0.006	False	False	0.552	46	0.827	49
14	0.007	True	False	0.574	49	0.830	45
15	0.007	False	False	0.563	47	0.830	49

On the validation set, all scenarios using a padding region perform slightly better. The effect of pretraining is higher for R_S than for j_0 as can, for example, be seen when comparing scenarios 2 and 10. Generally, the results show that the F_1 score of the R_S defect detection decreases with increasing illumination time (see Figure 6.6) while the j_0 detection performance remains almost constant (see Figure 6.7).

Using models that are pretrained on still images can improve the results and training speed for all illumination times. Similar to the results from Experiment 1 the performance on j_0 defects is higher and the impact of the motion blur on detecting these defects is lower.

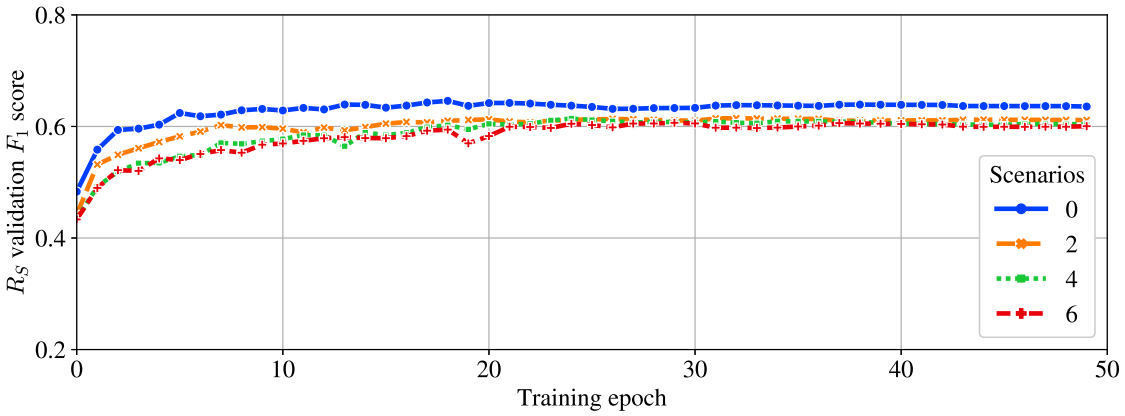


Figure 6.6: Validation accuracy for R_S defects after each training epoch on artificially blurred data set. Plot compares all chosen scenarios (one for each illumination time). Scenario 0 (2 ms), scenario 2 (4 ms), scenario 4 (6 ms), scenario 6 (7 ms). It can be seen that the F_1 score reduces most significantly from 2 ms to 4 ms.

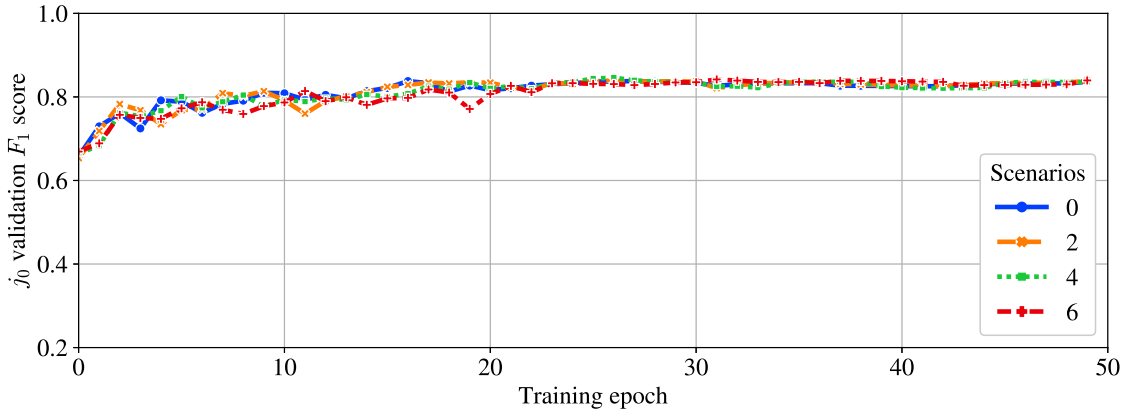


Figure 6.7: Validation accuracy for j_0 defects after each training epoch on artificially blurred data set. The plot compares all chosen scenarios (one for each illumination time). Scenario 0 (2 ms), scenario 2 (4 ms), scenario 4 (6 ms), scenario 6 (7 ms). It can be seen that the influence of motion blur on the j_0 F_1 score is rather low compared to the influence on R_S in Figure 6.6.

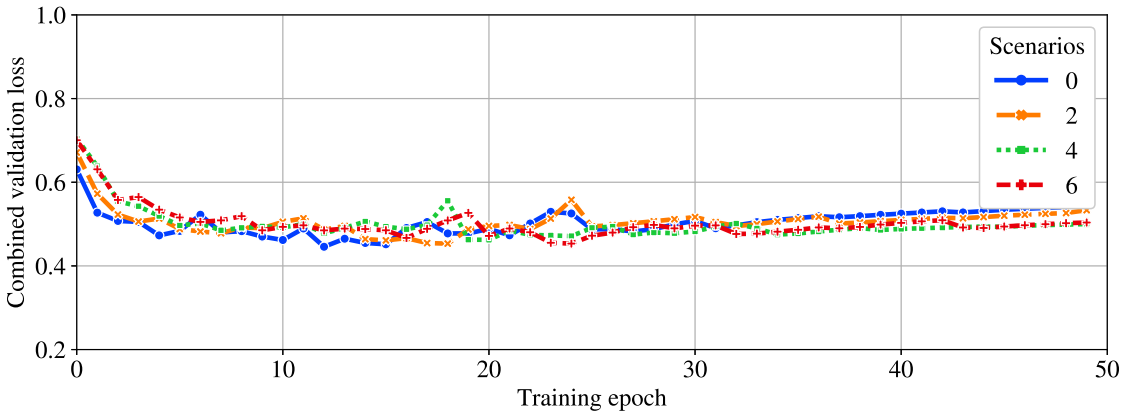


Figure 6.8: Combined validation loss for j_0 defects after each training epoch. The plot compares all chosen scenarios (one for each illumination time). Scenario 0 (2 ms), scenario 2 (4 ms), scenario 4 (6 ms), scenario 6 (7 ms).

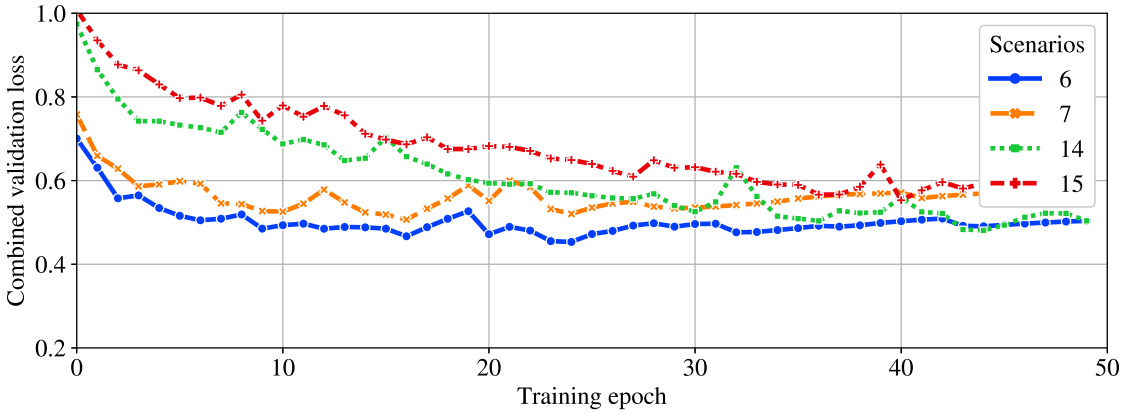


Figure 6.9: Combined validation loss for j_0 defects after each training epoch. The plot compares all scenarios trained with artificial motion blur simulating an illumination time of 7 ms. The loss of the scenarios without pre-trained models (scenarios 14 and 15) decreases significantly slower. The scenario using a pre-trained model and a padding region (scenario 6) achieves the lowest loss (i.e., best performance).

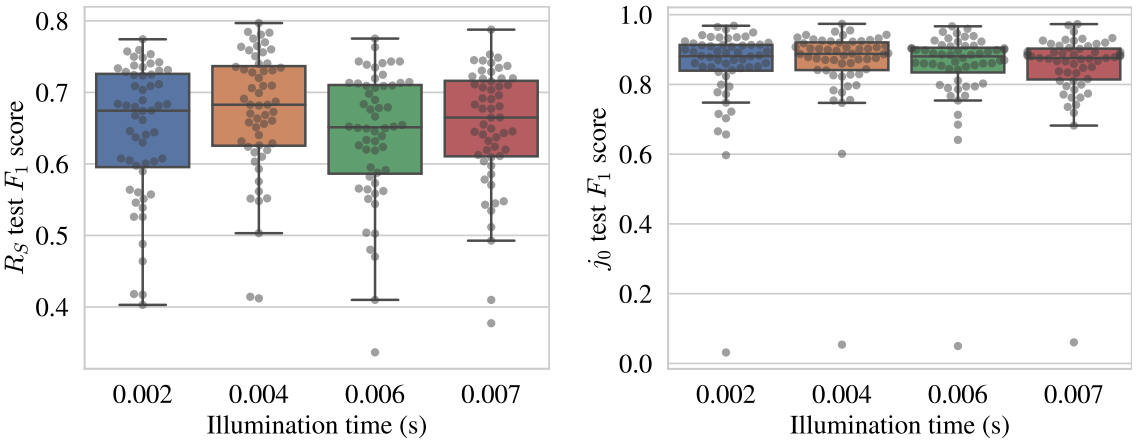
Test Set Performance

The evaluation on the test set is carried out for different decision thresholds between 0.3 and 0.9. As the validation set results in Table 6.3 are based on a decision threshold of 0.5, this decision threshold has to be used when comparing validation set performance and test performance from Table 6.4. The test set results at a decision threshold of 0.5 are slightly better than the validation set results. The share of the number of cells from each cell batch differs between the validation and test dataset, i.e., more “challenging” cells could be included in the validation set.

Table 6.4: Results on the test set for selected scenarios of experiment 2. P: Precision, R: Recall, F1: F_1 score

scenario id	T_L (s)	class threshold	P R_S	R R_S	F1 R_S	P j_0	R j_0	F1 j_0
0	0.002	0.3	0.489	0.862	0.620	0.684	0.971	0.787
2	0.004	0.3	0.466	0.883	0.606	0.741	0.970	0.826
4	0.006	0.3	0.442	0.860	0.580	0.706	0.965	0.801
6	0.007	0.3	0.435	0.878	0.578	0.733	0.960	0.818
0	0.002	0.5	0.604	0.750	0.665	0.749	0.951	0.823
2	0.004	0.5	0.568	0.798	0.661	0.797	0.943	0.850
4	0.006	0.5	0.552	0.759	0.635	0.766	0.942	0.832
6	0.007	0.5	0.540	0.785	0.636	0.790	0.928	0.841
0	0.002	0.7	0.720	0.596	0.646	0.816	0.913	0.847
2	0.004	0.7	0.675	0.677	0.673	0.852	0.897	0.860
4	0.006	0.7	0.668	0.622	0.639	0.828	0.900	0.849
6	0.007	0.7	0.652	0.657	0.651	0.844	0.876	0.847
0	0.002	0.9	0.884	0.313	0.453	0.904	0.780	0.821
2	0.004	0.9	0.832	0.429	0.559	0.922	0.773	0.825
4	0.006	0.9	0.834	0.362	0.497	0.909	0.766	0.815
6	0.007	0.9	0.819	0.401	0.531	0.916	0.733	0.797

Figure 6.10 shows the distribution of F_1 scores for both defect classes at each simulated illumination time on the test set with artificial motion blur. Each dot represents one cell in the test set. The standard deviation for j_0 defects is lower than for R_S defects indicating a more reliable prediction. For both classes, the mean F_1 score is highest at $T_L = 4\ ms$ but does not differ significantly between the four illumination times.



(a) R_S defects

(b) j_0 defects

Figure 6.10: Boxplots showing the distribution of F_1 scores for both defect classes at each illumination time at a decision threshold of 0.7.

6.2.2 Qualitative Results

Analysis of Whole Cells

The same cells as in Experiment 1 are selected for a qualitative inspection. Artificial motion blur simulating $T_L = 4\ ms$ is applied to the cells before they are passed to the model from scenario 2. Compared to Figure 6.5 and Figure 6.11 from Experiment 1, no significant degradation in performance can be observed for both cells in Figur. Again the labels are shown in the first row while the model's predictions are visualized below them.

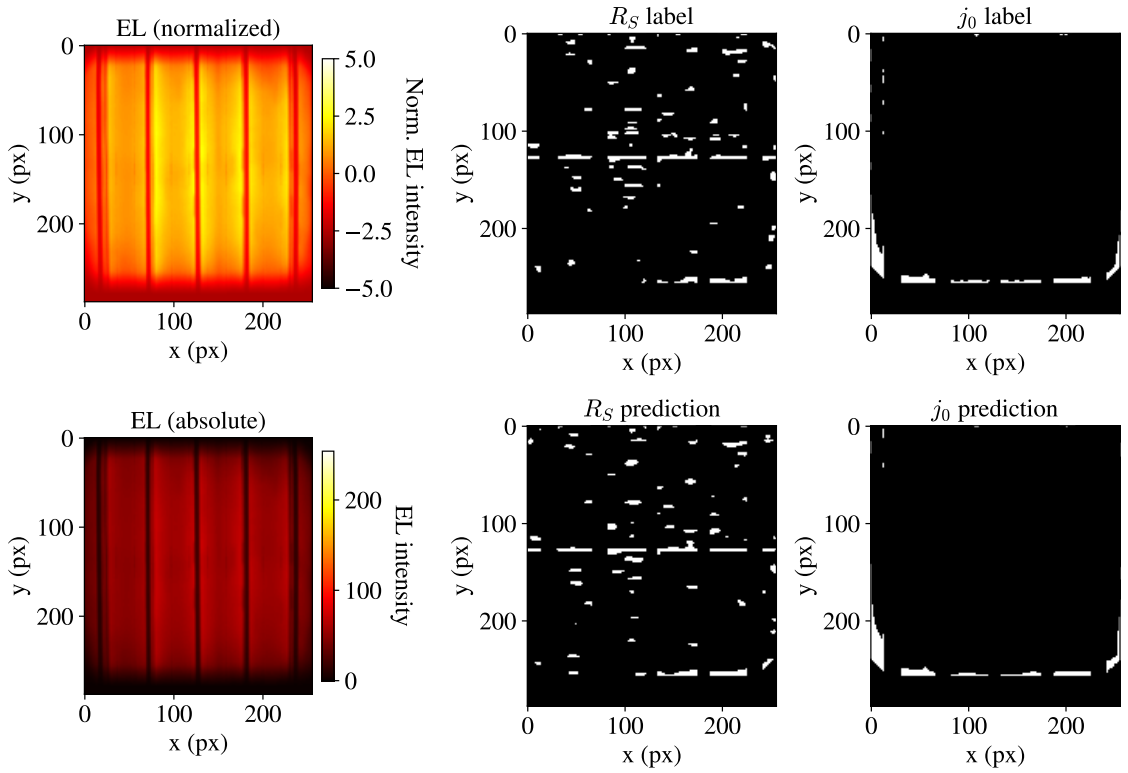


Figure 6.11: Performance of algorithm in scenario 2 ($T_L = 4\ ms$) on cell 210015ISE0748 at a decision threshold of 0.7.

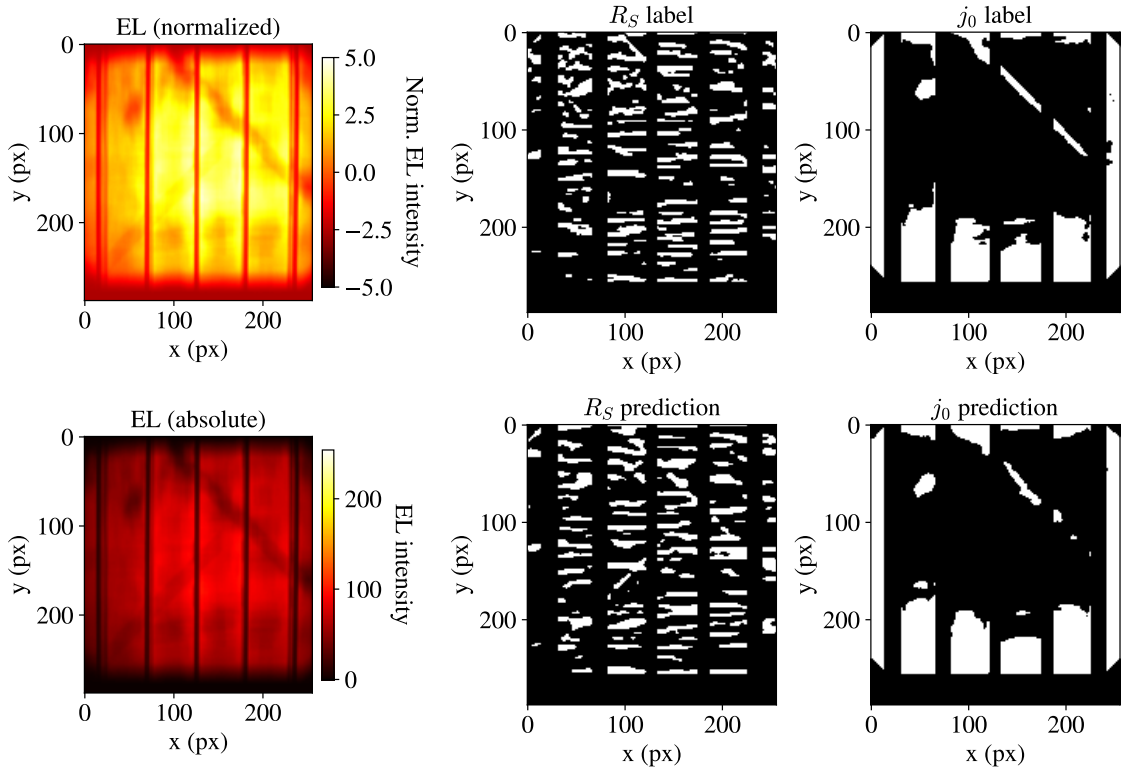


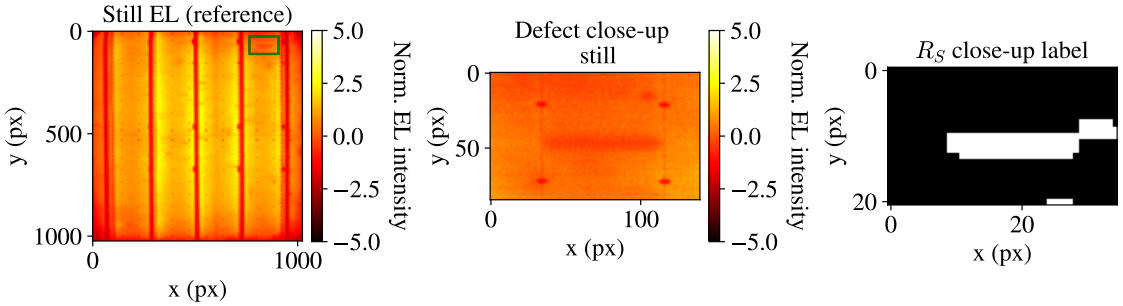
Figure 6.12: Performance of algorithm in scenario 2 ($T_L = 4\ ms$) on cell 210266ISEn337 at a decision threshold of 0.7. This cell shows a crack that is correctly predicted as a j_0 defect by the algorithm.

Detailed Analysis of R_S Defects

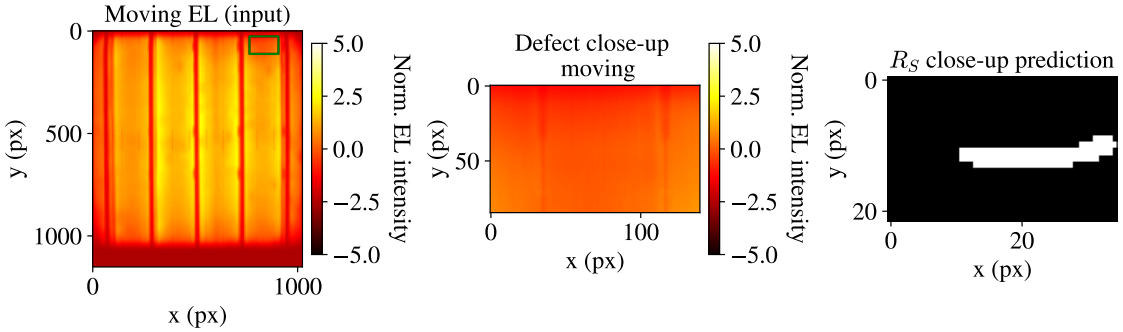
The insights gained from the quantitative results alone are not sufficient to understand the characteristics of its predictions in different scenarios in detail. The F_1 scores are similar. Additionally, they are computed based on the smart labels. The Smart Labeling algorithm uses very sensitive settings, therefore it labeled also very minor finger interruptions hardly visible in EL images. Hence, the shape of the predictions does not necessarily match the labels precisely even though the prediction can still be sufficiently accurate.

Two cells are selected to test the preliminary result that little degradation in performance can be observed when applying artificial motion blur. For both cells, a clearly visible finger interruption in the still EL image is identified by a human expert. The defect region is then marked in both the still and the EL images with motion blur. The images with motion blur from each illumination time are then fed into the network trained with images taken at the same illumination time. Due to the similarity of the 6 and 7 ms results, the 6 ms pictures are not plotted. An example of this analysis can be seen in Figure 6.15. The first row shows the full still EL image of the cell (left), a close-up image of the defect region (middle), and the label at this position (right). The other rows follow the same scheme but with moving images taken at 2 ms, 4 ms, and 7 ms illumination time. In these rows, the right plot shows the neural network prediction at the defect region instead of the label.

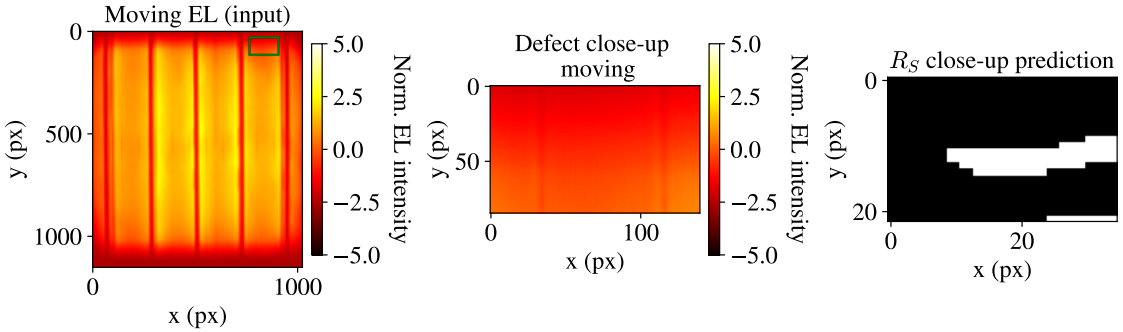
Figure 6.15 and Figure 6.14 reveal that the impact of the motion blur on the detection of finger interruptions is insignificant reflecting the results from the quantitative results. The predicted shape slightly changes, however, no general rule of how the shape changes can be observed. Overall, it can be said that the labels can remain in the same position and do not need to be changed in case of motion blur.



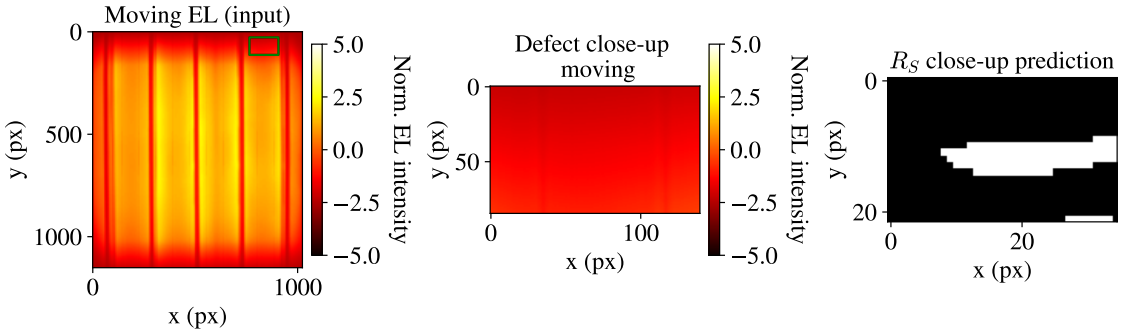
(a) Still image reference and label from Smart Labeling



(b) Image with artificial motion blur at $T_L = 2\ ms$ and prediction using model from scenario 0

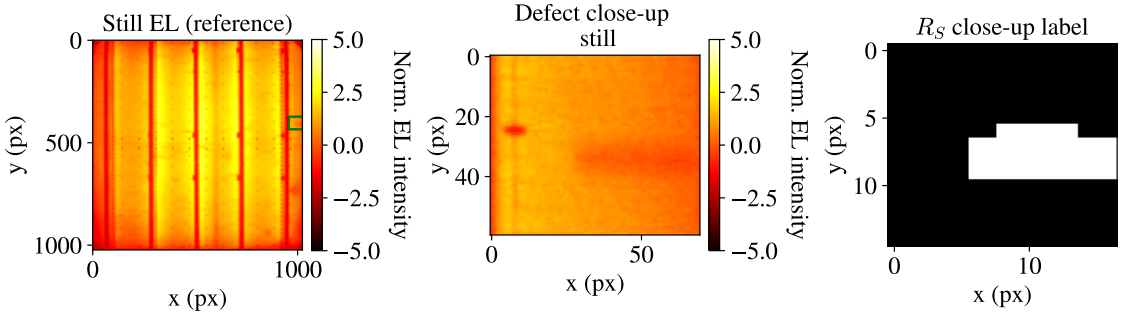


(c) Image with artificial motion blur at $T_L = 4\ ms$ and prediction using model from scenario 2

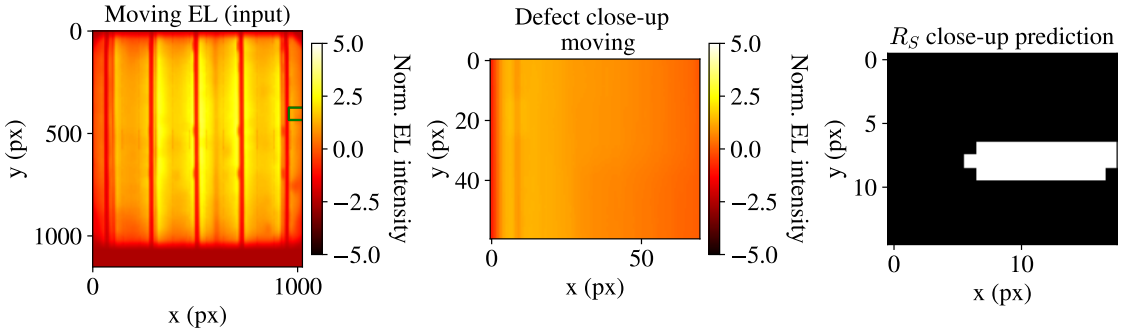


(d) Image with artificial motion blur at $T_L = 7\ ms$ and prediction using model from scenario 6

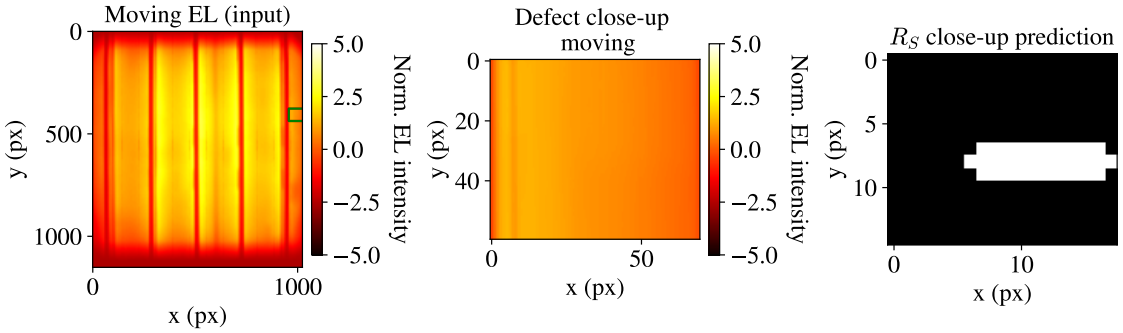
Figure 6.13: Finger interruption detection in images with artificial motion blur at different illumination times including still EL image as a reference for comparison (cell 210015ISE0748). Beware that the close-up images of the defect region are cropped from the original EL image at 1024x1024 pixels resolution while the network is only provided with a downscaled input of 256x256 pixels.



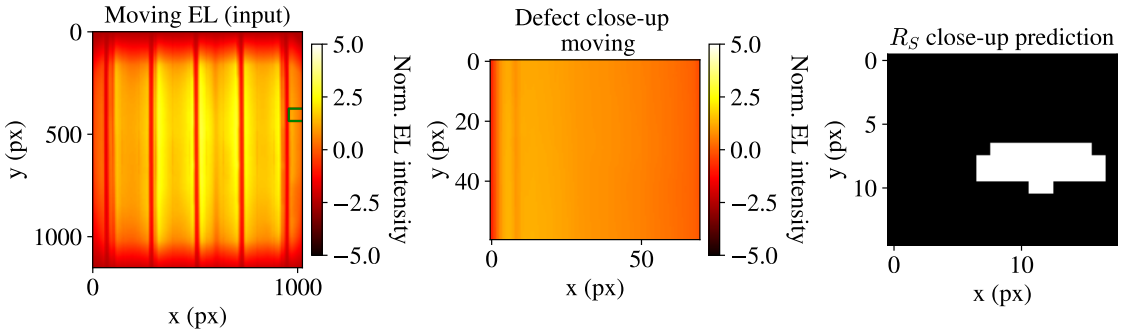
(a) Still image reference and label from Smart Labeling



(b) Image with artificial motion blur at $T_L = 2\ ms$ and prediction using model from scenario 0



(c) Image with artificial motion blur at $T_L = 4\ ms$ and prediction using model from scenario 2

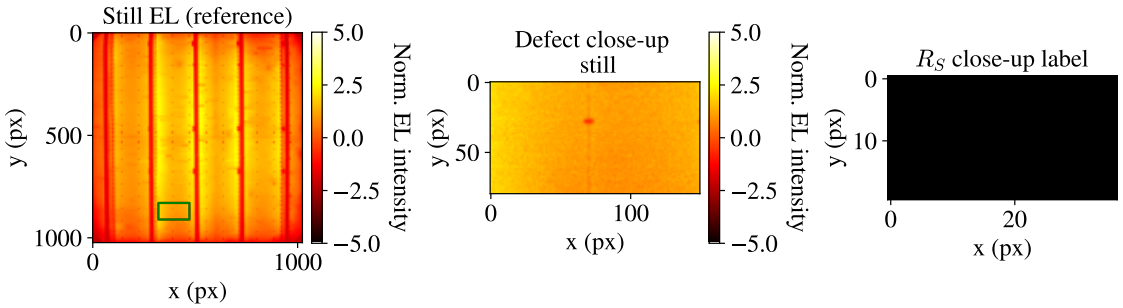


(d) Image with artificial motion blur at $T_L = 7\ ms$ and prediction using model from scenario 6

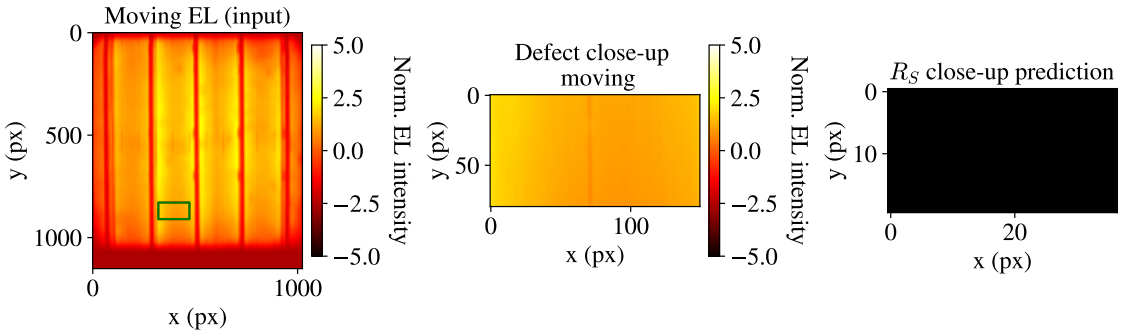
Figure 6.14: Finger interruption detection in images with artificial motion blur at different illumination times including still EL image as a reference for comparison (cell 210015ISE0819). Beware that the close-up images of the defect region are cropped from the original EL image at 1024x1024 pixels resolution while the network is only provided with a downscaled input of 256x256 pixels.

As a second analysis, a region with no defects is analyzed to gain some visual understanding of false positive predictions. In Figure 6.15 there is one false-positive prediction at $T_L = 4\ ms$

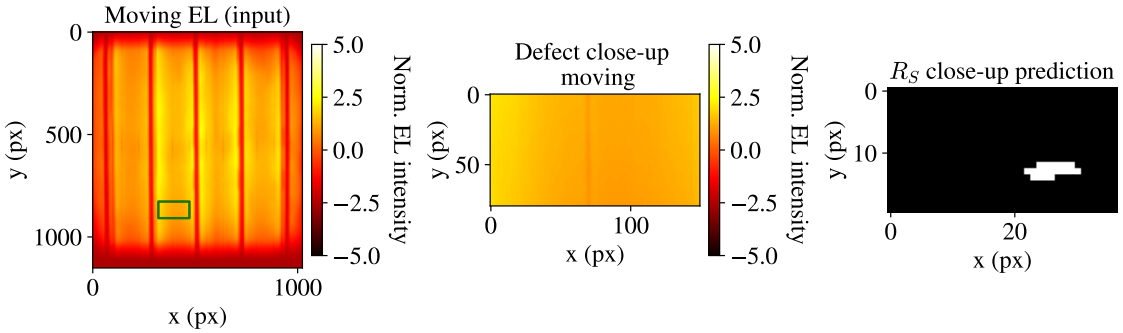
while all other models do not predict any defects in the same region.



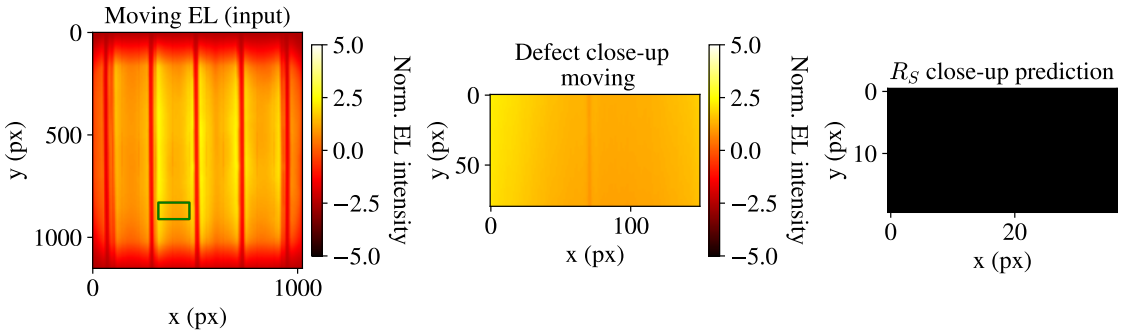
(a) Still image reference and label from Smart Labeling



(b) Image with artificial motion blur at $T_L = 2\ ms$ and prediction using model from scenario 0



(c) Image with artificial motion blur at $T_L = 4\ ms$ and prediction using model from scenario 2



(d) Image with artificial motion blur at $T_L = 7\ ms$ and prediction using model from scenario 6

Figure 6.15: Example of false-positive detections at $T_L = 4\ ms$ (cell 210015ISE0748). At other illumination times, there is no false positive prediction.

6.3 Experiment 3: Moving, Real-World Blurred

6.3.1 Quantitative Results

Validation Set Performance and Model Selection

Table 6.5: Results of the real measurement scenarios on the validation set. These results are used to select the best-performing models for each illumination time. The four selected scenarios are highlighted in light blue.

scenario id	T_L (ms)	padding region	pre-trained	best F_1 R_S	best epoch R_S	best F_1 j_0	best epoch j_0
0	0.002	True	Artificial	0.579	11	0.755	49
1	0.004	True	Artificial	0.602	22	0.791	22
2	0.006	True	Artificial	0.546	19	0.755	49
3	0.007	True	Artificial	0.527	11	0.783	20
4	0.002	True	Still	0.577	9	0.783	23
5	0.004	True	Still	0.596	8	0.785	30
6	0.006	True	Still	0.549	21	0.786	22
7	0.007	True	Still	0.519	22	0.784	42
8	0.002	True	False	0.549	44	0.738	48
9	0.004	True	False	0.569	46	0.703	46
10	0.006	True	False	0.483	49	0.722	49
11	0.007	True	False	0.486	49	0.728	47

For every illumination time, one best-performing model is selected. Figure 6.16 shows the model selection for the models at $T_L = 7\ ms$. Scenario 3 using a pre-trained model trained on artificial motion blur performs best but only by a small margin. Not using any pre-trained model is not recommended as it produces worse results. However, considering the small size of the training dataset with real motion blur (22 samples) it is remarkable which accuracy is achieved without pretraining. From the validation set results, it can be seen that pre-trained models help to achieve better results. The difference between using models trained on artificial motion blur and still images is low especially for j_0 defects.

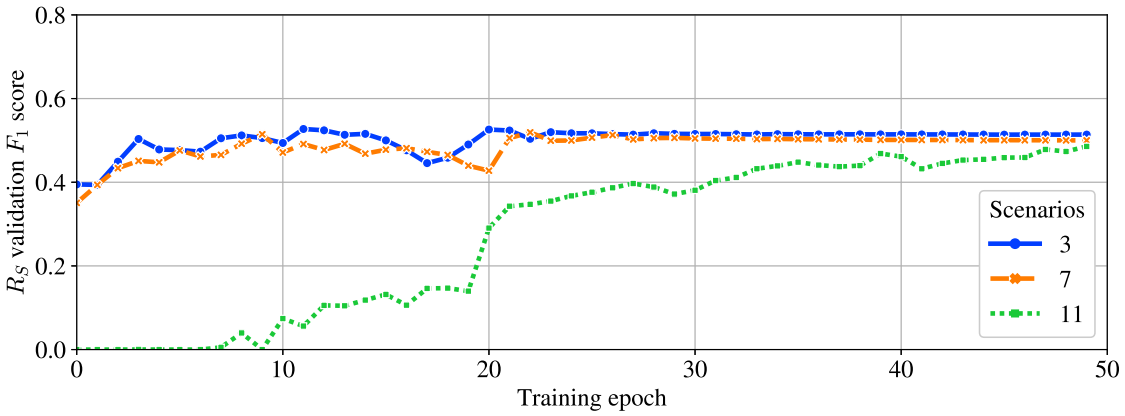


Figure 6.16: Validation accuracy for R_S defects comparing all scenarios trained on images at $T_L = 7\ ms$. While the pre-trained models trained on artificial motion blur (scenario 3) and still (scenario 7) lead to a similar performance on the moving images, the model without pretraining (scenario 11) does not reach the same performance after 30 epochs.

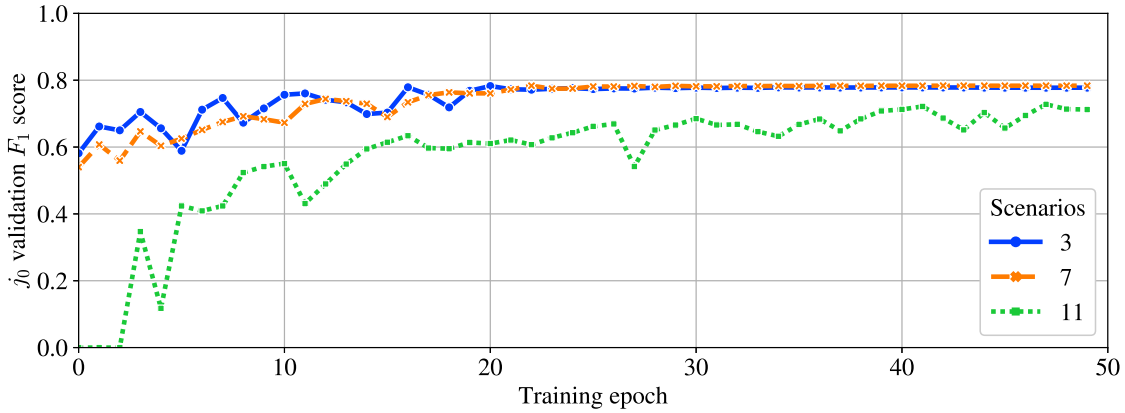


Figure 6.17: Validation accuracy for j_0 defects comparing all scenarios trained on images at $T_L = 7\ ms$. Similar to R_S defects, the pre-trained models perform significantly better.

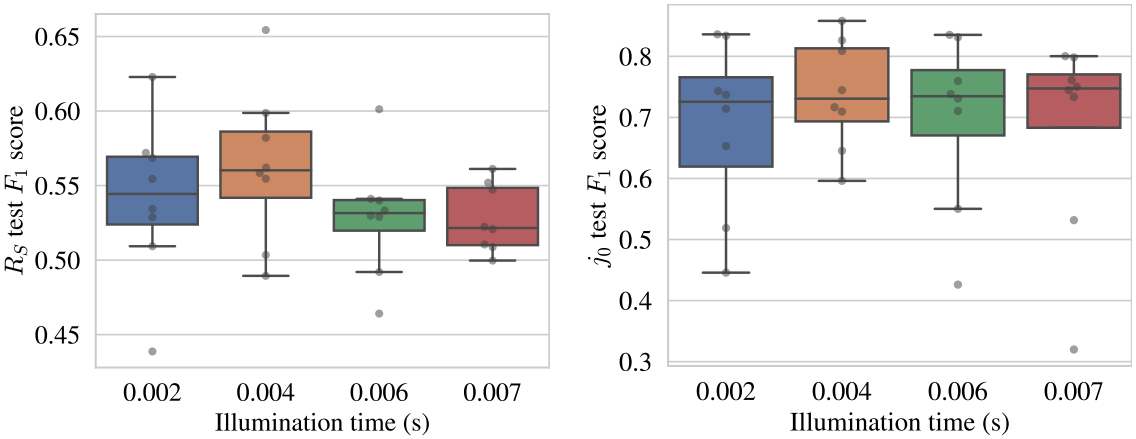
Test Set Performance

The models selected based on the validation set performance are evaluated using the test set at different decision thresholds. The results are presented in Table 6.6. The results between the validation set and test set do not differ significantly. As for the results in Experiments 1 and 2, a decision threshold of 0.7 yields the highest F_1 scores.

Table 6.6: Results on the test set for selected scenarios of experiment 3. P: Precision, R: Recall, F1: F_1 score

scenario id	T_L (s)	class threshold	P R_S	R R_S	F1 R_S	P j_0	R j_0	F1 j_0
0	0.002	0.3	0.435	0.662	0.523	0.452	0.960	0.596
1	0.004	0.3	0.519	0.635	0.569	0.577	0.926	0.695
6	0.006	0.3	0.380	0.641	0.476	0.479	0.941	0.612
7	0.007	0.3	0.394	0.625	0.483	0.470	0.947	0.601
0	0.002	0.5	0.527	0.579	0.550	0.521	0.921	0.643
1	0.004	0.5	0.597	0.566	0.579	0.653	0.873	0.729
6	0.006	0.5	0.468	0.573	0.514	0.553	0.903	0.662
7	0.007	0.5	0.504	0.546	0.524	0.547	0.899	0.646
0	0.002	0.7	0.620	0.484	0.541	0.606	0.863	0.685
1	0.004	0.7	0.677	0.486	0.563	0.723	0.794	0.738
6	0.006	0.7	0.569	0.497	0.529	0.640	0.838	0.698
7	0.007	0.7	0.626	0.458	0.528	0.639	0.826	0.680
0	0.002	0.9	0.779	0.325	0.457	0.734	0.702	0.688
1	0.004	0.9	0.809	0.336	0.473	0.851	0.638	0.710
6	0.006	0.9	0.749	0.351	0.476	0.797	0.673	0.702
7	0.007	0.9	0.802	0.310	0.446	0.792	0.659	0.677

Figure 6.18 shows the distribution of F_1 scores for both defect classes at each illumination time. It can be seen that the average score is highest for $T_L = 4\ ms$. Each dot corresponds to one solar cell in the test set. The low number of samples in the test set (8 samples) however limits the ability to draw statistical conclusions. Therefore, the qualitative analysis carried out in the next section is key to understanding the detection performance of the algorithm at different measurement settings in detail.



(a) R_S defects

(b) j_0 defects

Figure 6.18: Boxplots showing the distribution of F_1 scores for both defect classes at each illumination time at a decision threshold of 0.7.

6.3.2 Qualitative Results

Analysis of Whole Cells

Figure 6.19 shows convincing results for the j_0 defects. Severe R_S defects (finger interruptions) are found as well. However, the network also predicts finger interruptions where no interruptions are labeled. Given that the Smart Labeling is generally very sensitive towards finger interruptions, e.g. sometimes even interruptions hardly visible in the EL image are labeled, it can be assumed that the network is predicting some non-existing interruptions.

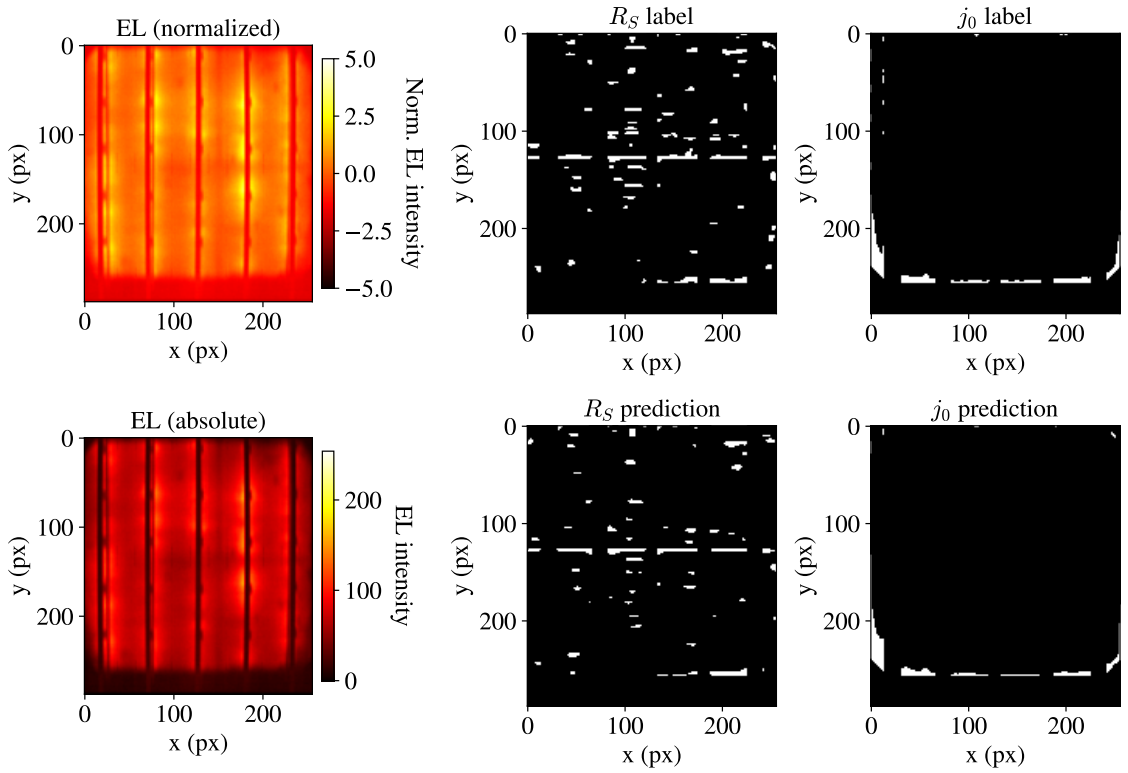
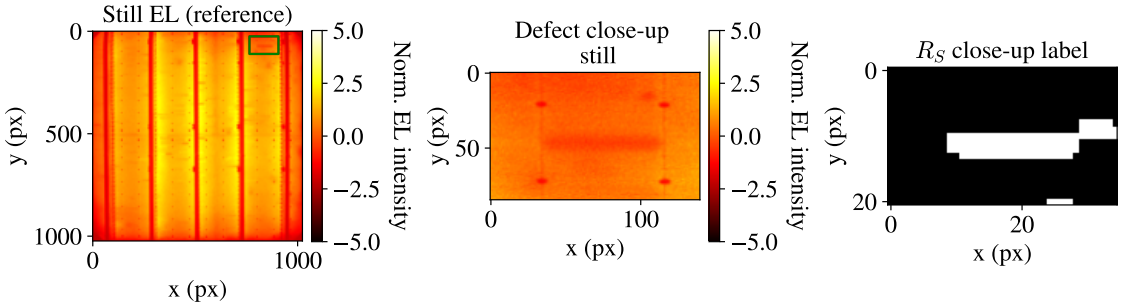


Figure 6.19: Performance of algorithm in scenario 1 ($T_L = 4\ ms$) on cell 210015ISE0748 at a decision threshold of 0.7.

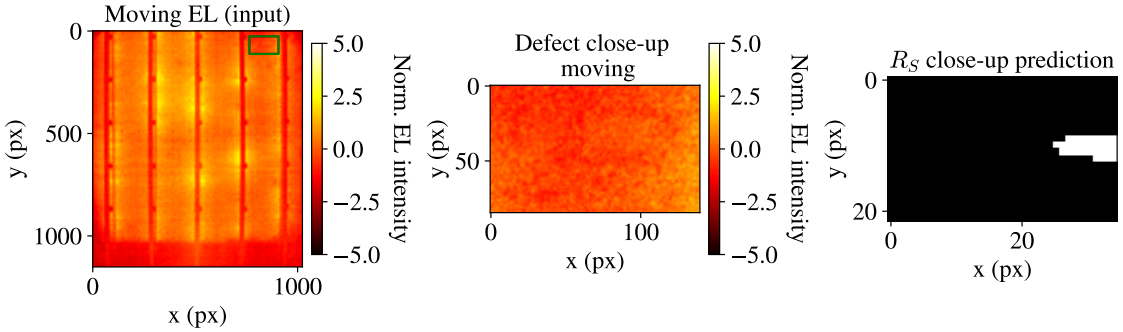
Detailed Analysis of R_S Defects

To cope with the mentioned limitations of the quantitative analysis, the algorithm performance is evaluated in detail on different defects. At first, three different cells from the test set are selected and regions with finger interruptions are chosen. Compared to Experiment 2, the qualitative analysis of this experiment is more elaborate because the results differ between different illumination times.

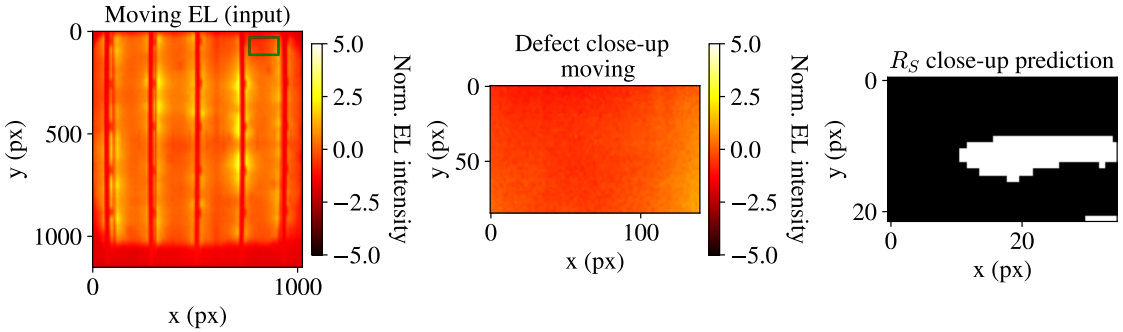
In Figure 6.20, it can be seen that a clearly visible finger interruption is successfully found at $T_L = 4\ ms$. At the other illumination times, the detection is less accurate but does not fail completely as part of the defective regions are still detected.



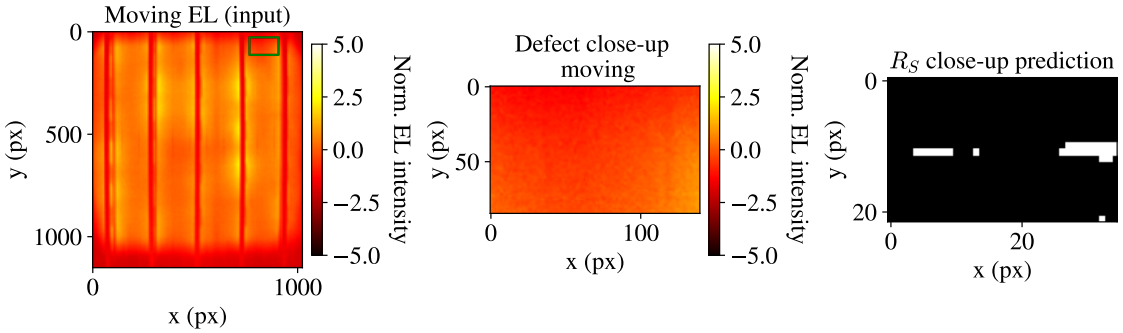
(a) Still image reference and label from Smart Labeling



(b) Moving image at $T_L = 2\ ms$ and prediction using model from scenario 0



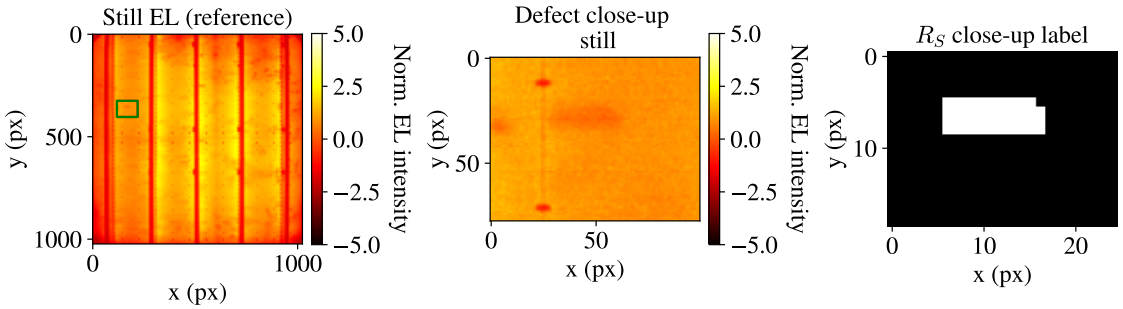
(c) Moving image at $T_L = 4\ ms$ and prediction using model from scenario 1



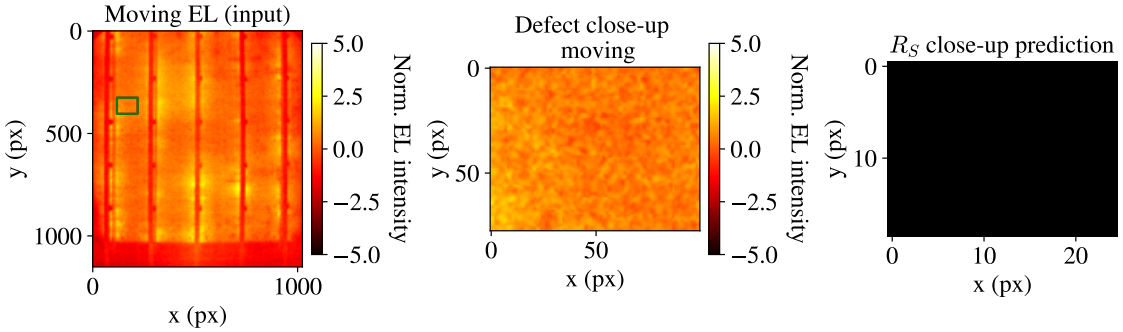
(d) Moving image at $T_L = 7\ ms$ and prediction using model from scenario 3

Figure 6.20: Finger interruption detection in moving images at different illumination times including still EL image as a reference for comparison (cell 210015ISE0748). Beware that the close-up images of the defect region are cropped from the original EL image at 1024x1024 pixels resolution while the network is only provided with a downscaled input of 256x256 pixels.

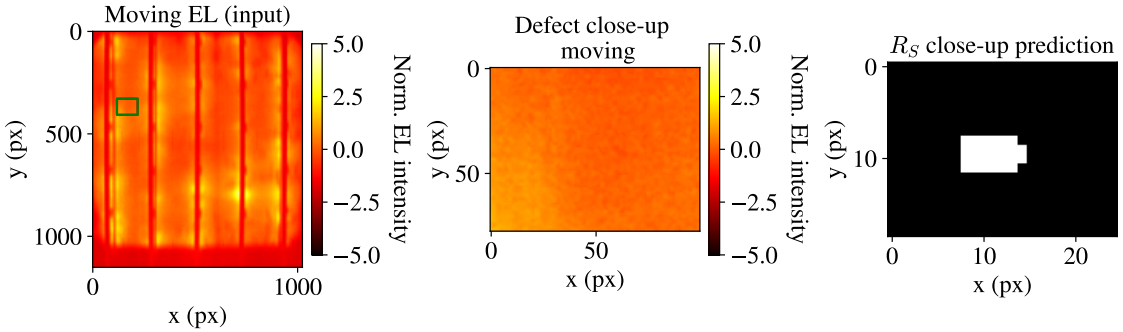
In Figure 6.21, the prediction works best at $T_L = 4\ ms$ as well. At the other illumination times, the network is totally unable to correctly predict the finger interruption.



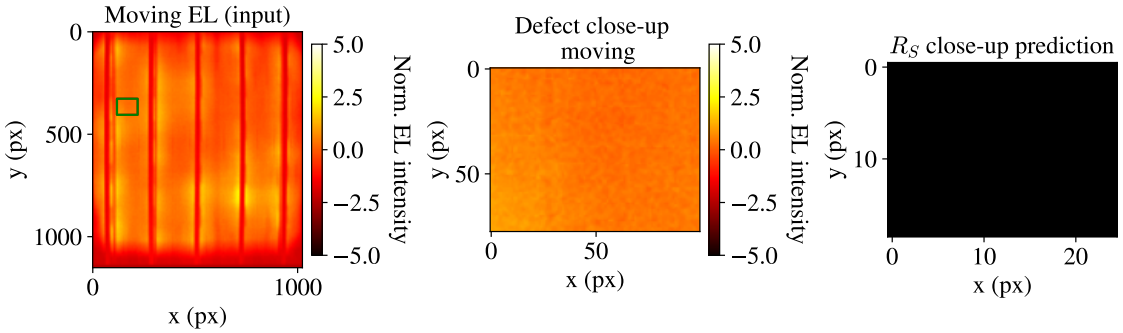
(a) Still image reference and label from Smart Labeling



(b) Moving image at $T_L = 2\ ms$ and prediction using model from scenario 0



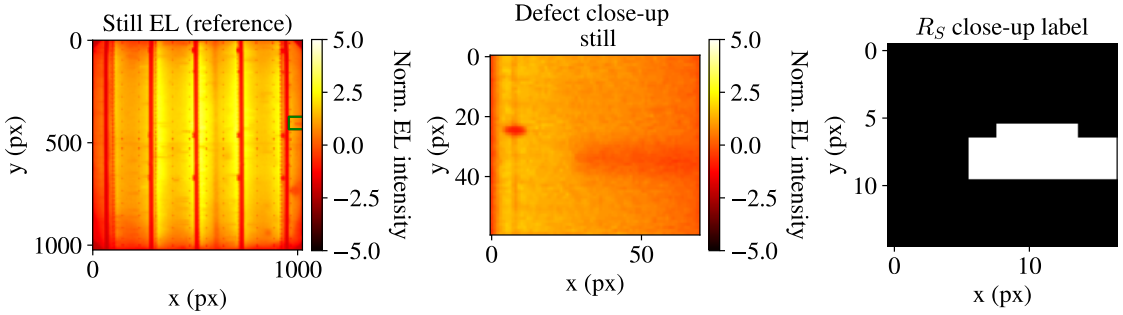
(c) Moving image at $T_L = 4\ ms$ and prediction using model from scenario 1



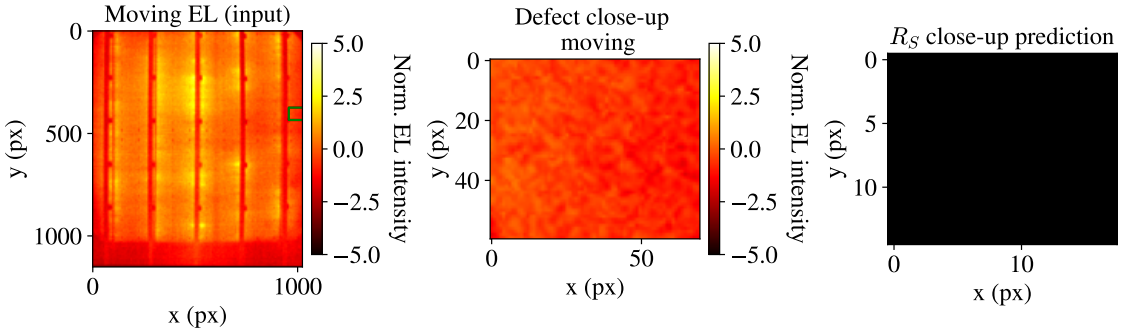
(d) Moving image at $T_L = 7\ ms$ and prediction using model from scenario 3

Figure 6.21: Finger interruption detection in moving images at different illumination times including still EL image as a reference for comparison (cell 210015ISE0709). Beware that the close-up images of the defect region are cropped from the original EL image at 1024x1024 pixels resolution while the network is only provided with a downscaled input of 256x256 pixels.

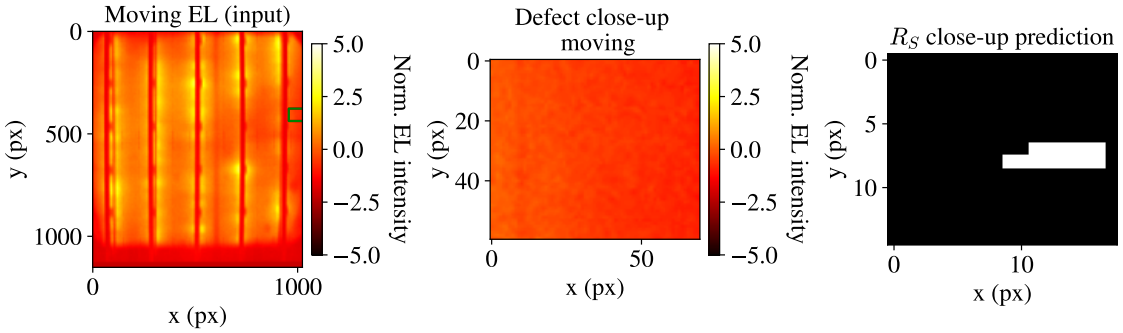
In the third cell, the network at 4 ms is the only one capable of detecting a strong finger interruption while the others do not find the defect (see Figure 6.22).



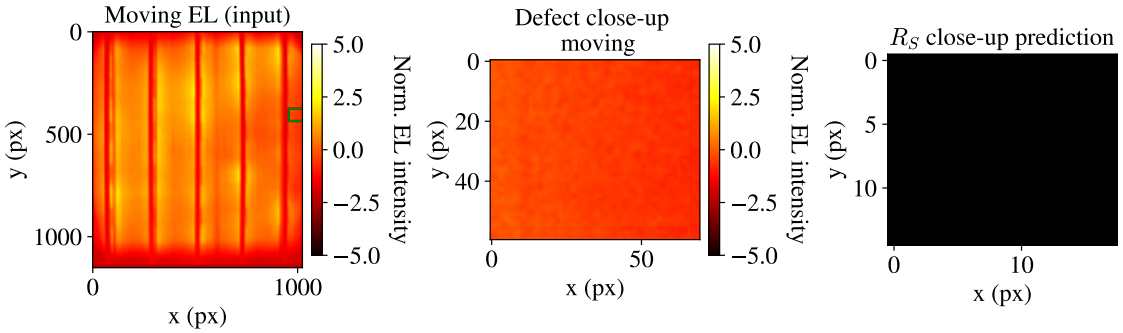
(a) Still image reference and label from Smart Labeling



(b) Moving image at $T_L = 2\ ms$ and prediction using model from scenario 0



(c) Moving image at $T_L = 4\ ms$ and prediction using model from scenario 1

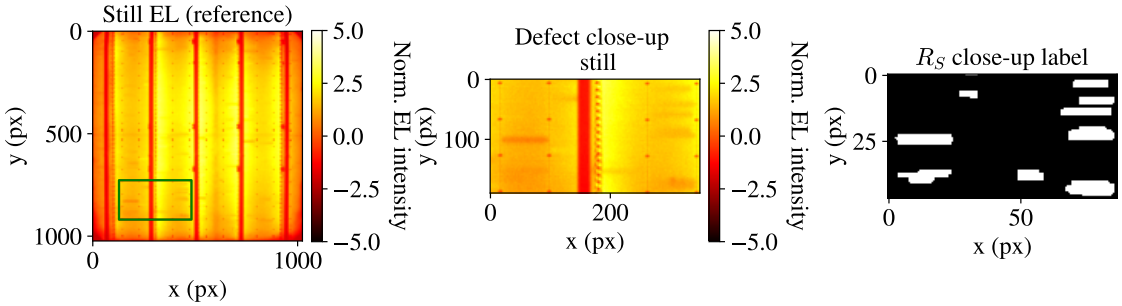


(d) Moving image at $T_L = 7\ ms$ and prediction using model from scenario 3

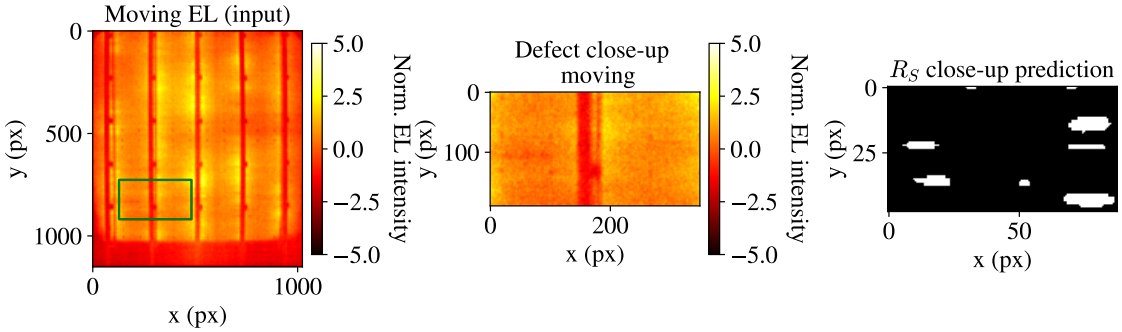
Figure 6.22: Finger interruption detection in moving images at different illumination times including still EL image as a reference for comparison (cell 210015ISE0819). Beware that the close-up images of the defect region are cropped from the original EL image at 1024x1024 pixels resolution while the network is only provided with a downscaled input of 256x256 pixels.

In Figure 6.23 the detection of finger interruption of different severities is investigated. The selected region is bigger than in the previous qualitative analyses. The prediction at $T_L = 7\ ms$

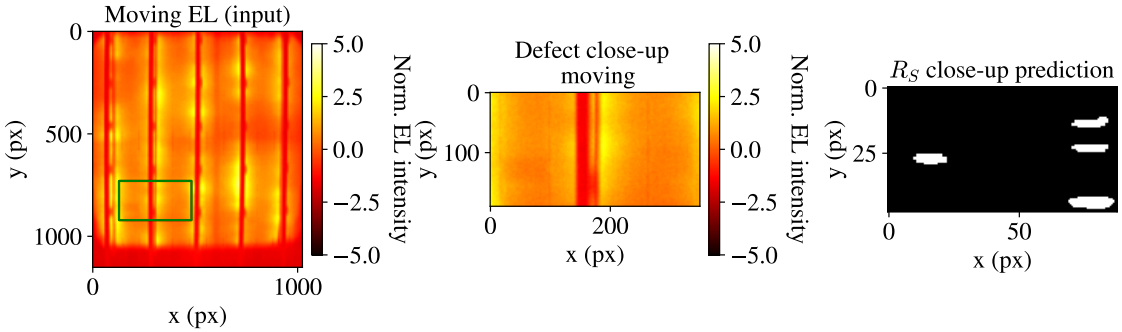
is clearly the worst. It overlooks the strong interruption in the left side of the selected region that is visible in the middle of Figure 6.23. At 2 ms, a higher number of small defect regions are predicted while the prediction at 4 ms looks cleaner. Whether the prediction at $T_L = 2\ ms$ or at $T_L = 4\ ms$ is better is therefore difficult to judge.



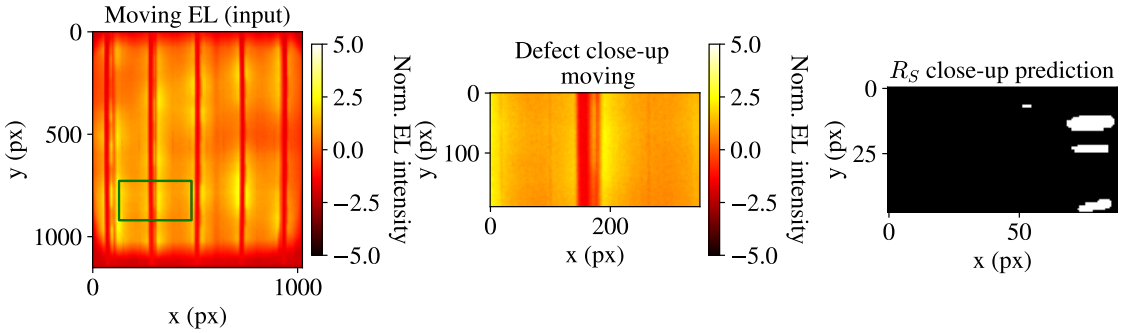
(a) Still image reference and label from Smart Labeling



(b) Moving image at $T_L = 2\ ms$ and prediction using model from scenario 0



(c) Moving image at $T_L = 4\ ms$ and prediction using model from scenario 1

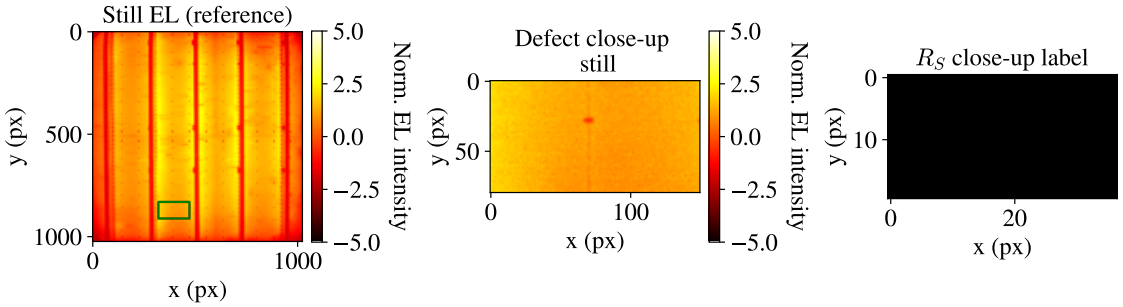


(d) Moving image at $T_L = 7\ ms$ and prediction using model from scenario 3

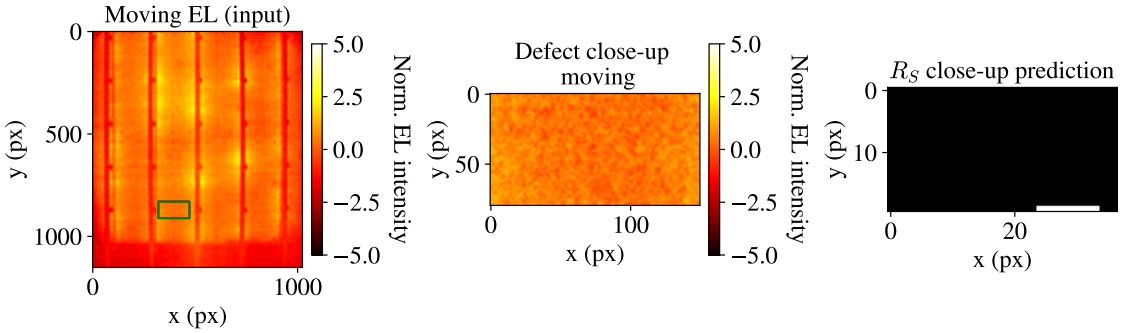
Figure 6.23: Finger interruption detection in moving images at different illumination times including still EL image as a reference for comparison (cell 210015ISE0691). Beware that the close-up images of the defect region are cropped from the original EL image at 1024x1024 pixels resolution while the network is only provided with a downscaled input of 256x256 pixels.

Lastly, an image region that contains no defects is chosen in cell 210015ISE0748 to investigate the false positive behavior of the networks. In Figure 6.24 it can be seen that the networks at

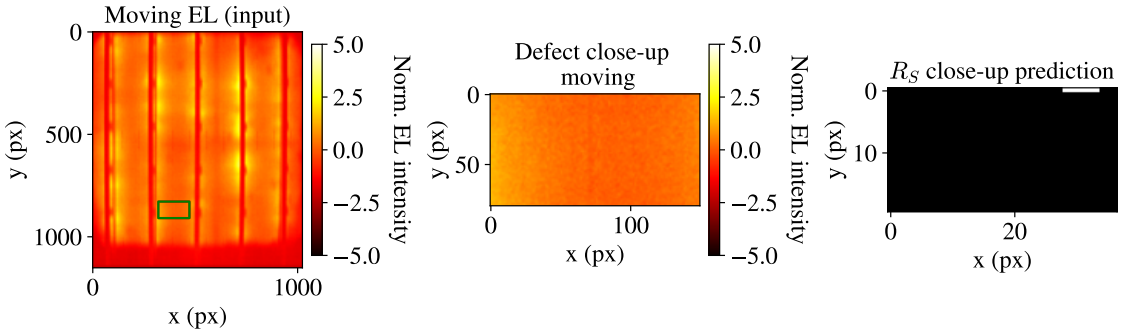
$T_L = 2\ ms$ or $T_L = 4\ ms$ predict thin rectangular bar-shaped at different locations. The false-positive prediction at $T_L = 2\ ms$ is much bigger than at $T_L = 4\ ms$.



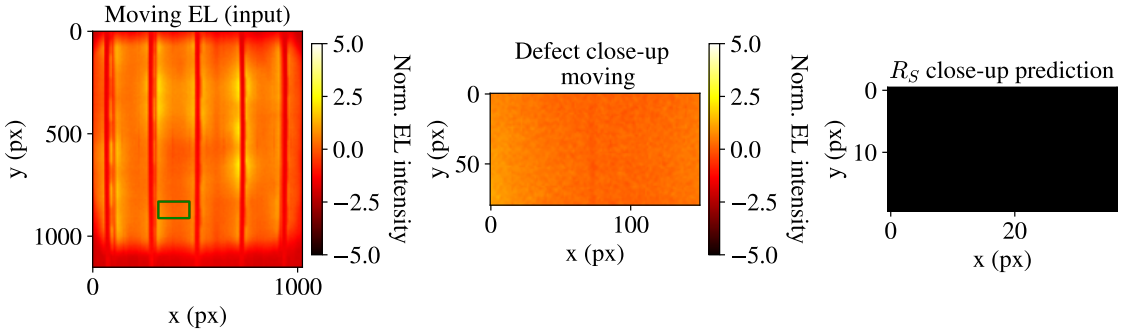
(a) Still image reference and label from Smart Labeling



(b) Moving image at $T_L = 2\ ms$ and prediction using model from scenario 0



(c) Moving image at $T_L = 4\ ms$ and prediction using model from scenario 1



(d) Moving image at $T_L = 7\ ms$ and prediction using model from scenario 3

Figure 6.24: Finger interruption detection in moving images at different illumination times including still EL image as a reference for comparison (cell 210015ISE0748). Beware that the close-up images of the defect region are cropped from the original EL image at 1024x1024 pixels resolution while the network is only provided with a downscaled input of 256x256 pixels.

Generally, it can be said that the prediction at $T_L = 4\ ms$ is the most reliable: In all presented examples, it has been capable of detecting clearly visible interruptions. Low-contrast interruptions are generally not found very well. It can be argued that some labeled interruptions are not visible enough in the EL images, especially not in moving images with a low signal-to-noise ratio, but they are most visible in the C-DCR reference used for labeling.

Generalizability to Other Cells

In an independent evaluation, a TOPCon cell provided by the measurement device prototype manufacturer is measured at all illumination times. The cell is manufactured in China under industrial conditions. The cell is considered non-defective by industrial standards. It has not been included in the training. A moving EL image of this cell taken at $T_L = 4\ ms$ is passed to the model of scenario 1. The algorithm response to the unknown sample can be seen in Figure 6.25. The algorithm still predicts defects even though the cell is considered defect-free. This behavior is expected because the sample is out of the data distribution that the network has learned.

However, measuring the industrial TOPCon cell still provides important insights regarding the industrial adaptability of the algorithm. The moving image in the upper left image in Figure 6.25 has been normalized using the pixel intensity mean and standard deviation of the moving dataset from Experiment 3 at $T_L = 4\ ms$ consisting of experimental cells from Fraunhofer ISE. Comparing the industrial cell to an experimental ISE cell, e.g. see Figure 6.19, it can be seen that the industrial cell is much brighter at the same illumination time. This indicates that other measurement parameters can be chosen for industrial-grade cells, e.g. shorter illumination time or lower gain, leading to higher quality images that are easier to be processed by the CNN.

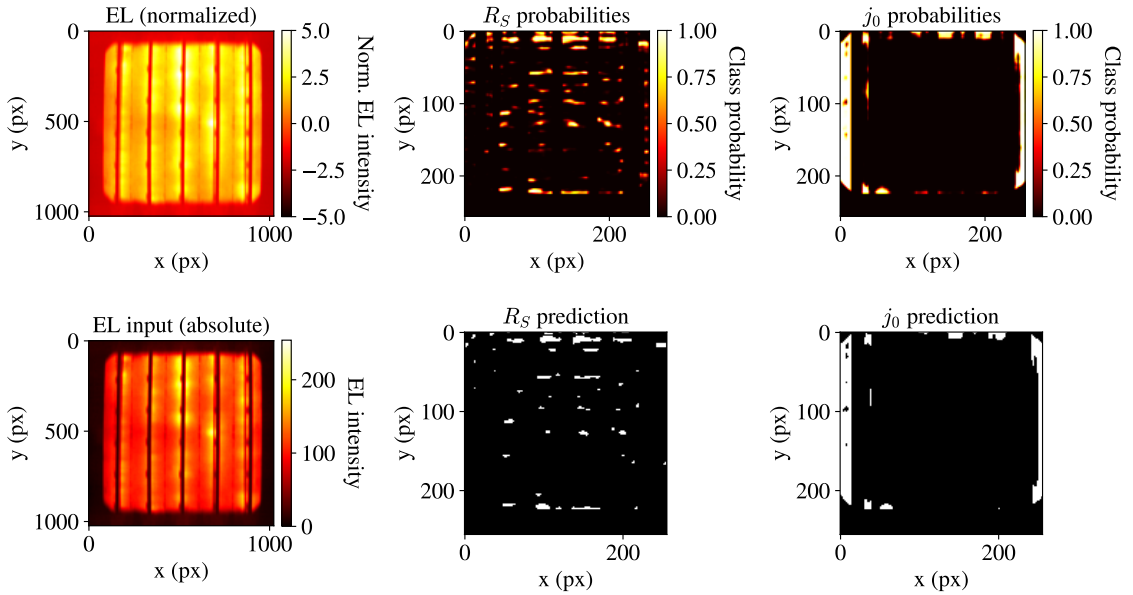


Figure 6.25: Performance of algorithm in scenario 1 ($T_L = 4\ ms$) on non-defective reference cell 5 at a decision threshold of 0.7.

7 Technical Discussion

7.1 Detection Accuracy and Impact of Motion Blur

Experiment 2 revealed that the impact of the artificial motion blur on the detection accuracy is rather low, i.e., CNNs can deal well with artificial linear motion blur. Experiment 3 also shows that the difference between using a pre-trained model with artificial motion blur or a model trained on still images is relatively low. Therefore, it becomes obvious that the approach used to estimate the motion blur is not comprehensive enough even though it looks realistic to a human observer. The camera noise should be estimated as well to create a realistic estimate of moving EL images.

Ideally, Experiment 3 would not require any additional training. This would require estimating the characteristics of motion blur realistically enough in order to train a network with realistic artificial samples. The dataset gathered in Experiment 3 can be used for a more realistic noise estimation to create a new dataset with artificial motion blur for future experiments. Aligning the moving images, i.e., cropping out only those regions where the cell is visible, from Experiment 3 with the labels currently requires high manual effort because the blurry edges cannot be detected reliably by the applied alignment algorithm.

For images with real motion blur, it can generally be said that the prediction at $T_L = 4\ ms$ is the most reliable. This indicates the required careful trade-off between motion blur and camera gain. Images taken at $T_L = 2\ ms$ at a high camera gain of 4 do not contain enough information to be reconstructed by the applied approach. The focus in finding the right parameters for moving EL images should not only be put on avoiding or reducing motion blur itself even though it strikes the human observer the most on the first sine.

While the size of j_0 defect areas is often predicted accurately, the exact size of finger interruptions is hard to predict in moving images with the applied approach. The size of predicted finger interruption regions tends to reduce in moving images. This could have implications when calculating quantities, e.g., power loss, based on the size of the affected region. The reason why j_0 defect areas are less impacted by motion blur can be explained by the geometric properties of the defect: While the finger interruptions causing the R_S defects have a bar shape perpendicular to the direction of movement, the j_0 defects generally form a larger closed area. Hence, the information about the thin bar-shaped finger is more likely to be impacted by linear motion blur.

Clear advice is given that images with motion blur should not be cropped too tightly. This is shown by the padding region experiments. The experimental results from Experiments 1 and 2 indicate that the hypothesis that the regions around the cell affected by motion blur contain important information that the CNN can recover is right. Another reason why the detection perfor-

mance is better with a padding region that has been considered is that the area of non-defective regions increases while the area containing defect regions remains constant. The assumption would be that it is easy to predict that the padding region contains no defects. But Experiment 1 (without motion blur) included scenarios with and without padding region. On still images without motion blur, no significant difference with and without a padding region could be observed. This indicates the positive impact of a padding region on the F_1 score for images with motion blur.

Finally, it should be mentioned that comparing the performance on the real-world dataset and the artificially blurred dataset is slightly "unfair" due to the very different sizes of the training and validation dataset, as well as the necessarily different settings of the EL images.

7.2 Smart Labeling

The refined Smart Labeling approach opens many opportunities for the simplified deployment of neural networks for quality inspection as proposed by Kunze et al. (2020) [5]. The improved approach in this work is more robust against variations of cells and therefore strengthens the big advantage of Smart Labeling to adopt neural networks for quality inspection easily to new production processes. Human labeling effort can be reduced and the physical meaning of the reference measurement leads to a clear definition of labels defined. In this work, the labels have been only evaluated qualitatively. For further research, it would be recommended to statistically evaluate the advantage over human labels.

The clear physical definition of j_0 defects based on physical thresholds seems to have a big advantage in detection accuracy whereas the more feature-based R_S labeling has a lower correlation with the EL image. The C-DCR measurement technique has not fully been adopted to TOPCon yet, especially resulting in inhomogeneous R_S images. Therefore, it was not possible to rely only on physical quantities for R_S defects.

Fine-tuning C-DCR to work more reliably with TOPCon would be key to generating more reliable R_S labels. This may include testing new measurement recipes in the laboratory. The developed approach using dynamic adaptive thresholding for R_S defect labeling is very sensitive. It labels even some defects that can hardly be seen in the EL image. In this case, the neural network can also not detect the defect and the loss function punishes the network for something that even the best algorithm could detect. The statistical evaluation using the F_1 score is negatively impacted by this as well. The qualitative analyses however have proven to be a good means of coping with this weakness by focusing on finger interruptions that are clearly visible in the still EL image.

7.3 Dataset

The dataset is rather imbalanced as it does not contain cells without any faults due to the experimental stage of the cells. Because the used cells stem from scientific experiments, they differ much more from each other than on an industrial production line. It can be expected that the

detection accuracy would increase if only one cell type is used in training. Additionally, the results could improve significantly when using industrial cells that generally appear brighter in EL images as they have a higher open-circuit voltage and conversion efficiency. With brighter industrial cells and therefore lower camera gain and lower illumination time (not higher than 4 ms), it can be assumed that defects clearly visible in EL images can be detected reliably by CNNs.

8 | Sustainability Approach and Initial Assessment

In this chapter, a general framework to assess the possible impacts of the technical innovation in the field of PV quality inspection is defined. The algorithm and process level of the new framework are directly applied to gain an overview of possible sustainability implications and to allow for the definition of experimental scenarios in Chapter 9.

8.1 Framework Definition

On the one hand, the view for any large-scale consequences of developed technical innovations can get lost easily in research with a strong technical focus. On the other hand, research that takes more of a systems perspective on technology, environment, and society easily overlooks important details or cannot even consider details by design.

To cope with these challenges and to leverage the full potential of combining technical research and the field of sustainability, a framework to assess the possible impacts of the technical innovation is defined. It makes use of three different perspectives: The algorithm level, the process level, and the large-scale level as presented in Figure 8.1. The concept of approaching a sustainability challenge from different perspectives or levels certainly draws inspiration from the multi-level perspective of Geels [44], however, the developed framework is tailored towards the technical development at hand. The framework provides general research questions at every level. It aims at simplicity to reduce the burdens for researchers to make use of it. It should not require too much specific knowledge in the field of sustainability. By embracing simplicity, it can be leveraged as an easy bridge between technical research and sustainability.

The innovation developed in this project can impact sustainability on all three levels. The development of quality inspection algorithms always has a potential real-world impact. Often, decisions such as sorting or quality grading are directly or indirectly influenced by algorithm decisions. The algorithm level provides the opportunity to assess those impacts.

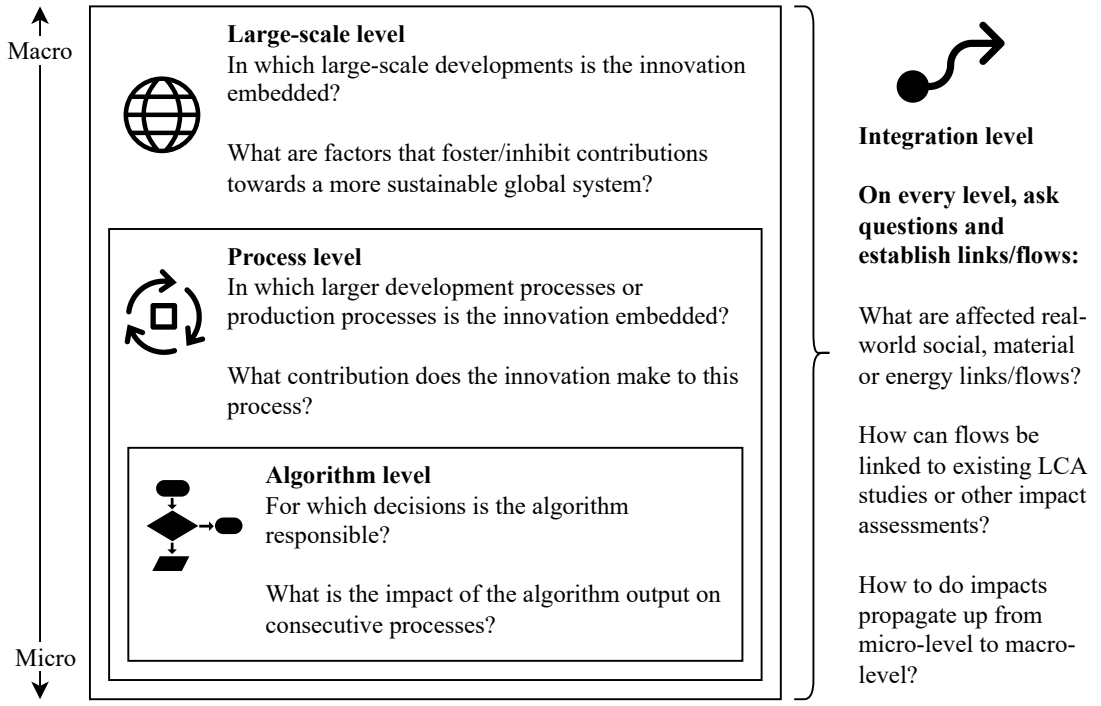


Figure 8.1: The designed framework that consists of three levels.

The process level looks at the larger process that is supposed to be changed by the application of an AI algorithm. For example, this could be improving the overall quality of production output or increasing production throughput. At the large-scale level, it can be estimated what impact the process-level improvements have at a larger scale.

8.2 Initial Application of Framework Levels

As shown in Figure 8.2, the algorithm development and evaluation is the building block of this project. In the initial assessment in this section, factors that have to be considered in the further analysis are identified at each level and used for scenario definition in Chapter 9.

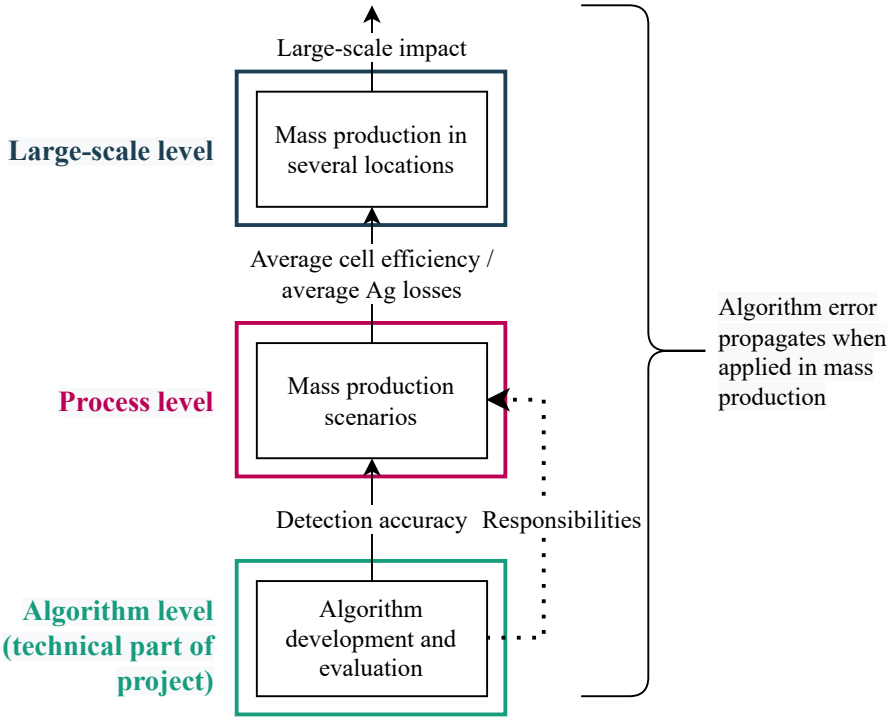


Figure 8.2: General research flow of the sustainability assessment: Knowledge gained from the development and evaluation of the algorithm is used to project the potential impact at a large scale.

8.2.1 Algorithm Level

At first, the greater context of solar cell quality inspection is studied. The knowledge and network gained from the technical research are beneficial to gain further process insights. Reviewing literature and the personal exchange with researchers at Fraunhofer ISE have revealed these two main algorithm use cases:

I. EL images can be used for end-of-line binning, i.e., to sort solar cells into different quality classes [3], [62]. Currently, this is often done by measuring the current-voltage (I-V) characteristics. EL images can reveal defects that I-V measurements do not show (e.g., microcracks) [3].

II. An important future vision is a highly connected PV production line (smart factory). The concept is to constantly monitor each process along the solar cell manufacturing chain and store digital twins of the cell [4], [63]. The major motivation is to find systematic process-related defects faster and to optimize machine utilization by means of predictive maintenance.

There are multiple decision consequences arising from use case I. If cells are binned incorrectly based on EL images, it could lead to material losses. A good cell could be discarded and is therefore lost for electricity production. As cells are usually combined into modules as shown in Figure 8.3, the error chain increases. In a module, a single bad cell can limit the overall module efficiency. Cells within a module should have similar efficiency to avoid mismatch loss [62]. Furthermore,

a module should not contain any damaged cells. If a bad cell is not detected and it is installed in a module, this could either lead to a fully damaged module or to increased mismatch loss.

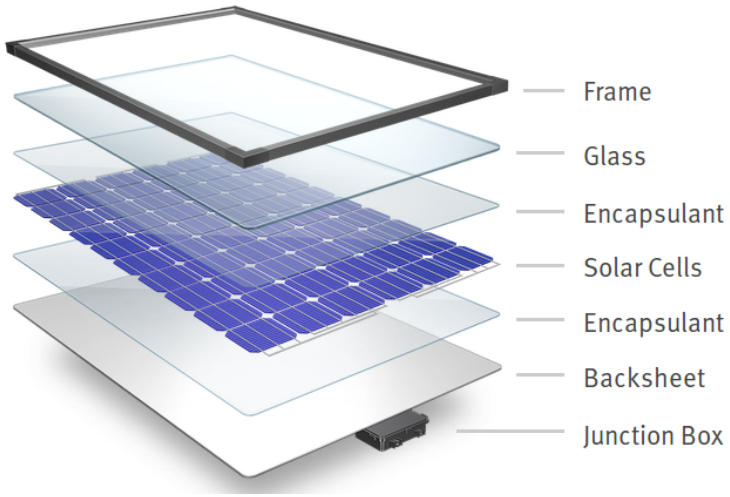


Figure 8.3: Structure of a photovoltaic module. *Image source: Wealth Daily* [64]

The developed algorithm can contribute to the "smart factory" vision in use case II mainly by its spatially resolved defect detection. In addition to the knowledge that a defect exists, the algorithm also predicts the location. Spatially resolved defect detection comes with the opportunity of finding systematic process defects faster. For example, defect detections can be superimposed to visualize the frequency of a defect per pixel [5]. If for example, a defect always occurs in the upper right corner, it can be concluded it is a systematic defect. This allows for faster troubleshooting in production and finding the root cause faster, leading to fewer defective cells being produced.

In use case II, the direct decision consequence is that a production line is stopped to resolve the cause of a systematic defect. A false positive detection would lead to an undesired loss in production yield as the line is stopped for no reason. In case the production is not stopped while a systematic error is present, faulty cells are manufactured. Depending on the severity, these cells either have a lower efficiency or the cell has to be dispensed. In both cases, this increases the environmental impact of PV electricity production. Either the efficiency is reduced and less energy is produced over the lifetime, or the good cells have to share the burden of production impact with the bad cells.

To demonstrate the advantages of spatially resolved defect detection, the following two scenarios are developed to test the impact of the algorithm in use case II, i.e., the case of detection of systematic production defects:

1. **Rigorous scenario:** Every systematic defect is considered so severe that it leads to the disposal of the produced cells. Using silver consumption as the main indicator, this scenario is used to showcase that the absolute material losses in the production line might increase even if the amount of silver per cell is significantly reduced.

2. **Efficiency scenario:** Every systematic defect only leads to cells with lower conversion efficiency η. This scenario is used to show the possible benefits of a smart factory to increase average cell conversion efficiency.

Both scenarios are evaluated at different production throughput rates. The baseline throughput is 5,500 wafers per hour (wph) on one production line, the conservative prospective throughput is 12,300 wph, and the progressive prospective throughput is 24,300 wph on a single production line.

8.2.2 Process Level

As mentioned in the introduction, the research on fast quality inspection is part of the bigger NextTec project aiming at identifying the next-generation production technologies with highly increased throughput. This does not only require improvements in quality inspection and characterization. Almost all manufacturing processes such as Emitter Diffusion and Oxidation have to be improved to meet the target throughput.[1] The most significant economic lever provided by the innovation on the process level is reduced CAPEX per cell. Machinery may become slightly more expensive due to increased complexity, but overall it can be assumed that the CAPEX remains almost constant at doubled production rates.

Reducing factory space utilization is an important sustainability lever, especially due to reduced cleanroom usage. This is assessed in greater detail in Section 9.3.2 in the Sustainability Assessment chapter.

8.2.3 Life Cycle Assessment Model

The impact of the aforementioned scenarios is tested using an existing LCA model that has been developed within the SiTaSol project [65]. This project explored pathways to increase the conversion efficiency of c-Si solar cells to 30 % by combining them with III-V top absorbers [48], [66]. The LCA model developed by Blanco [67] is modified: The processes for producing III-V top absorbers are removed from the model, leaving a model equivalent to PERC solar cells at the state of the art of 2018. The model includes the entire balance of system of the PV system including the module, inverter, and mounting system [67].
To test the environmental impact of different scenarios the following main parameters, among others, are varied:

- Average conversion efficiency based on the output of the efficiency scenario.
- Loss of cells in the production process based on the rigorous scenario.
- Change of silver use.

The LCA model is implemented in the open source LCA framework Brightway2 [68]. The Activity Browser software [69] is used as a front-end to conduct the scenario analysis.

[1] Dr.-Ing. Sebastian Nold, Baljeet Goraya, Vasileios Georgiou-Sarlikiotis (Fraunhofer ISE). Presentation Techno-Economical Analysis within Internal NextTec Project Meeting. March 31st, 2022.

9 | Sustainability Assessment

To make use of the defined assessment framework from the previous chapter, scenarios have to be created. Scenarios allow testing the impact of the "smart factory" vision to which new defect detection algorithms are supposed to contribute. The third section introduces additional scenarios, explains how all scenarios are connected to the LCA model, and how the impact of cleanroom factory environments is modeled within the LCA.

9.1 Rigorous Scenario

9.1.1 Overview

The aim of the rigorous approach is to demonstrate the need for increased process monitoring at increased production throughput. Furthermore, the scenario is used to showcase that an increase in throughput can lead to higher absolute losses even if material usage is drastically reduced.

The scenario definition is based on a Fraunhofer ISE-internal cost of ownership analysis for increasing the production throughput of screen printing processes while saving material, especially silver. In the majority of today's solar cells, screen printing is used for the metalization process of the solar cell. Metalization is required to form the electrical contacts of the cells. Screen printing most commonly uses silver paste to print front- and rear-side contacts of the solar cell. Currently, silver is responsible for around 10 % of the costs for a typical solar cell [70]. The current standard process is flatbed screen printing (FSP). Rotary screen printing (RSP) is currently under development. The cost of ownership analysis compares the application of FSP versus RSP for HJT solar cells at different production throughputs. As shown in Figure 9.1, the key indicator is the monetary cost of ownership. In all scenarios, the process consumables, meanly silver, make up for by far the largest share of the cost of ownership (COO) per cell. The COO analysis compares three different screen printing scenarios per production output (solar cell).

- The **baseline FSP scenario** is flatbed screen printing (FSP) at 5,500 wafers per hour. This is the current industrial state of the art.
- The **conservative RSP scenario** assumes the application of RSP instead of FSP while scaling up production to 12,300 wafers per hour.
- The **progressive RSP scenario** assumes the application of RSP instead of FSP while scaling up production to 24,600 wafers per hour.

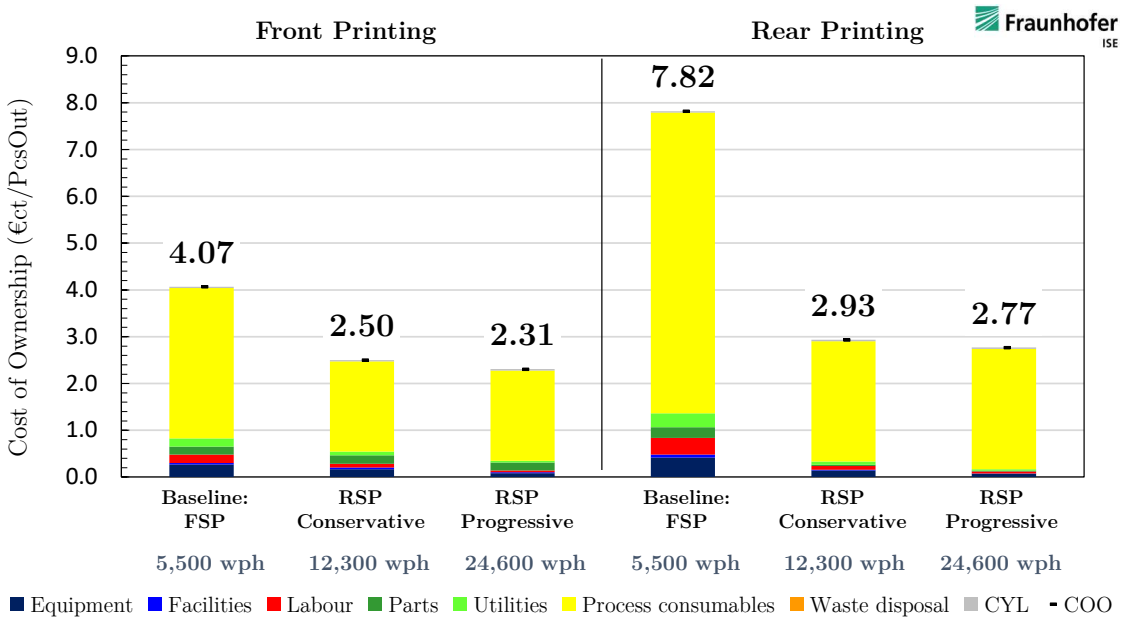


Figure 9.1: Cost of ownership (COO) analysis for scaling up screen printing of HJT cells. Flatbed screen printing (FSP) at 5,500 wafers per hour forms the baseline. The scenarios at increased throughput of 12,300 wafers per hour and 24,600 wafers per hour assume the application of rotary screen printing (RSP).[1]

An important advantage of the RSP over FSP is the reduced silver consumption. Researchers at Fraunhofer ISE have observed a reduction in silver consumption of -40 % for the front grid and -60 % for the rear grid of a HJT solar cell [16].

9.1.2 Methodology

It is assumed that the defect detection algorithm developed in the technical part of this thesis project is applied to detect systematic defects. The assumption is that every systematic defect is so severe that it leads to the disposal of the produced cells until the defect is resolved.

Table 9.1: Assumed parameters for the rigorous scenario.

Parameter	Value	
Probability loosing production of one hour (%)	1.00%	
Production hours per year (h)	8760	
Process stop time in case systematic defect (h)	0.0833	(5 minutes)
Expected number of defect occurances per year (-)	87.6	
Price per kg silver (EUR/kg)	662.81 [2]	

Table 9.1 shows the assumed parameters for the rigorous scenario. A production of 24 hours per day at 365 days a year is assumed at an up-time of all processes of 100 %. The probability that a

[1] Adopted from Dr.-Ing. Sebastian Nold (Fraunhofer ISE). Presentation Internal NextTec Project Meeting. March 31st, 2022.
[2] Global silver price per May 31, 2022. Retrieved from silverprice.org.

systematic error in screen printing occurs that renders all cells useless is assumed at 1 % per hour. Finding realistic numbers on this probability is difficult. Dr.-Ing. Sebastian Nold, head of team techno-economic and ecological analyses stated in a personal communication that 92-95 % uptime is a common estimate for machines across the PV production chain.[3] In the remaining 5-8 % of the time, the machine is not available for production, e.g., due to maintenance. However, it is difficult to find a temporal estimate of how often a systematic error affecting production occurs. As this assessment has the purpose of a sensitivity analysis, the precise hourly production error rate is not of utmost importance. Production data of manufacturers could possibly reveal this information, but detailed process insights are mostly kept confidential by the cell manufacturers.

Algorithm detection accuracy is defined as the probability of the defect detection algorithm finding a systematic production defect within 5 minutes. A systematic defect is defined as a defect that would lead to the disposal of cells produced within 1 hour.

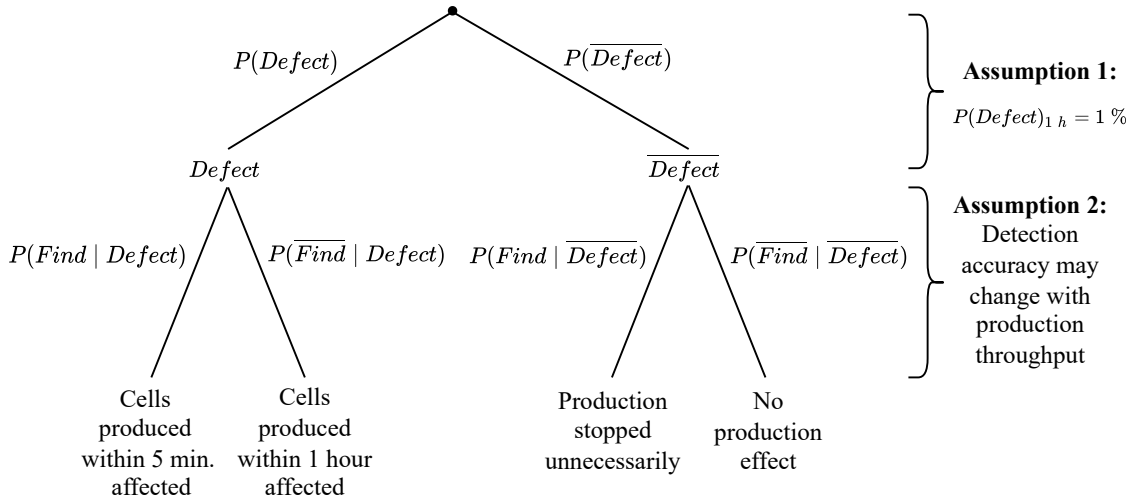


Figure 9.2: Illustration of probability experiment

The above assumptions can be translated into a probabilistic experiment as shown in Figure 9.2. A systematic defect occurs with the probability $P(Defect)_{1\,h} = 1\ \%$. It is assumed that if the defect is detected by the algorithm, the production can be stopped after 5 minutes, i.e., only the production of 5 minutes needs to be discarded. In case a defect occurs and the defect is not detected by the algorithm, it is assumed that the production of the consecutive hour has to be discarded.
For multiple detection accuracies, the expected number of defective and discarded wafers in case of a single systematic defect event is calculated as given in Equation 9.1.

[3] Dr.-Ing. Sebastian Nold (Fraunhofer ISE). Personal communication. May 17th, 2022.

$$E(X) = P(Find \cap Defect) \cdot W_{5m} + P(\overline{Find} \cap Defect) \cdot W_{1h} \tag{9.1}$$

With:

$E(X)$: Expected number of defect wafers

W_{5m} : Number of wafers produced within 5 minutes at given throughput rate

W_{1h} : Number of wafers produced within 1 hour at given throughput rate

To calculate the expected number of yearly discarded wafers, $E(X)$ is then multiplied by the number of expected defect occurrences per year (87.6 occurrences per year at $P(Defect)_{1\,h} = 1\ \%$ as presented in Table 9.1).

9.1.3 Results

Table 9.2: Results of the rigorous scenario. Even if silver usage per cell is drastically reduced, the absolute losses increase at increased throughput if the defect monitoring is inaccurate.

Variable	Baseline		Conservative		Progressive	
	w/o insp.	w/ insp.	w/o insp.	w/ insp.	w/o insp.	w/ insp.
Production rate (%)	5,500	5,500	12,300	12,300	24,600	24,600
Defect detection accuracy (%)	0.00	90.00	0.00	90.00	0.00	90.00
Ag/wafer front (mg)	60.00	60.00	36.00	36.00	36.00	36.00
Ag/wafer rear (mg)	120.00	120.00	48.00	48.00	48.00	48.00
Lost wafers if def. detected (waf)	458	458	1,025	1,025	2,050	2,050
Lost wafers if def. not detected (waf)	5,500	5,500	12,300	12,300	24,600	24,600
Expected loss in def. case (waf)	5,500	963	12,300	2,153	24,600	4,305
Expected loss per year (waf)	481,800	84,315	1,077,480	188,559	2,154,960	377,118
Ag in lost wafers (kg)	86.72	15.18	90.51	15.84	181.02	31.68
Ag losses (EUR)	57,482	10,059	59,990	10,498	119,980	20,996
Yearly production loss (%)	1.00	0.18	1.00	0.18	1.00	0.18

From Table 9.2 and Figure 9.3 it can be seen that even though RSP uses less silver, scaling up production without accurate monitoring can still lead to similar (at 12,300 wafers per hour) or even more (at 24,600 wafers per hour) absolute silver losses. The higher the throughput, the more significant accurate monitoring becomes if absolute losses should be minimized. For now, only silver losses have been considered because specific numbers are available from ISE. However, the expected yearly number of lost wafers can easily be translated into a change in average life-cycle impact across the entire production. The table on the next page shows the calculated scenarios. 0.00 % detection accuracy equals the case of no inspection present.

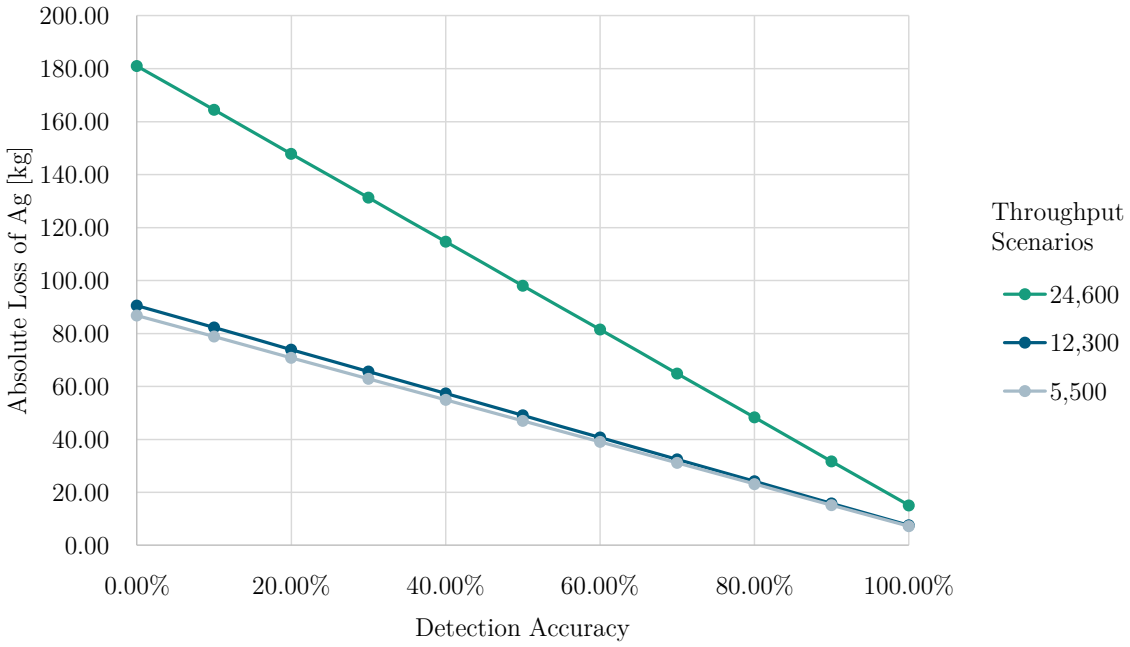


Figure 9.3: Expected silver loss per year for the three different throughput scenarios (5,500 wph with FSP, 12,300 wph with RSP, and 24,600 wph with RSP) at different detection accuracies.

9.1.4 Discussion

A personal interview with an expert from the industry revealed that cells with only little defects visible in the EL image are indeed highly undesired in the industry and may either be sold at a lower price or may even be discarded. Every manufacturer defines their own criteria for how to make decisions based on EL images [3], therefore it is not possible to provide a conclusion that is valid across the industry. The assumption that a defect that renders all cells of an hour useless and is not noticed without visual inspection is still drastic. If the current-voltage (I-V) characteristics are still monitored simultaneously, a very severe defect would be noticed this way. However, because I-V measurements become more challenging at increased throughput as well, this scenario still allows for establishing a somewhat realistic initial link between algorithm accuracy and sustainability performance.

9.2 Efficiency Scenario

9.2.1 Overview

Similar to the rigorous scenario, the aim is to show the need for increased process monitoring at increased production throughput. Therefore, the efficiency scenario makes use of similar assumptions and the same probabilistic experiment as shown in Figure 9.2. However, the baseline differs: Instead of the quite radical assumption that all cells which are affected by a systematic production defect are discarded, it is assumed that the cell efficiency η is lower in the case of a systematic defect. This can be considered more realistic because it is unlikely that cells with minor differences in conversion efficiency are discarded even though every manufacturer makes different trade-offs [3], [46].

9.2.2 Methodology

In this scenario, the best cell in a dataset of cells from a production line can be defined as the baseline. In case no defect occurs, it is assumed cells with the conversion efficiency of the best cell of the batch, η_{good}, are produced. In case a systematic defect occurs, it is assumed that cells with a lower η are produced until the defect is resolved because the defect is not noticed. The same assumption from the rigorous scenario regarding faulty production applies: 5 minutes of faulty production occurs in case the defect is detected and 1 hour in case the defect is not detected. It is assumed that the faulty wafers have the low efficiency η_{faulty} while good wafers are assumed to have the high efficiency η_{good}. The average efficiency is then calculated by Equation 9.2.

$$E_\eta(X) = \frac{W_{faulty} \cdot \eta_{faulty} + W_{good} \cdot \eta_{good}}{W_{total}} \tag{9.2}$$

With:

$E_\eta(X)$: Expected average efficiency of production

W_{faulty} : Number of "faulty" wafers with low efficiency

W_{good} : Number of "good" wafers with high efficiency

Table 9.3: Assumed parameters for the efficiency scenario.

Parameter	**Value**	**Notes**
Probability of low eff. cell event per hour (%)	1.00	
Production hours per year (h)	8760	
Process stop time in case systematic defect (h)	0.0833	(5 minutes)
Expected number of systematic defect occurrences (-)	87.6	
Efficiency of cell with defects (%)	23	
Efficiency of cell without defects (%)	25	

Table 9.3 states the assumptions for the efficiency scenario. As mentioned above, the best cell in a dataset of cells from a production line can be defined as the baseline. However, the best

experimental TOPCon cell in the dataset created within this project only shows an η of 21.981 %. This is rather low for a TOPCon cell. Therefore, an η of 22.5 % is assumed which is used in existing LCA studies for current industrial PERC cells [47]. For a TOPCon cell in mass production, this is still a conservative assumption. All other assumptions remain the same as for the rigorous scenario.

9.2.3 Results

As can be seen in Table 9.4, the loss in average efficiency does not change with scaling up production. The ratio of low-efficiency cells and total production stays constant. However, the expected absolute number of wafers with η_{faulty} changes significantly. In Figure 9.4 the impact of detection accuracy on both the absolute number of cells with low efficiency at 24,600 wph and on the change in average conversion efficiency is shown. While the change in conversion efficiency between 0 % and 100 % detection accuracy seems low (0.0138 % difference in average efficiency), the number of affected low-efficiency cells is still high, i.e., 2,154,960 affected cells without detection, 179,580 with 100 % detection accuracy. The low difference in average efficiency may indicate that the assumptions for the efficiency scenario are too conservative.

Table 9.4: Results of the efficiency scenario. In the simplified model that has been used the loss in average efficiency does not depend on the production throughput but the absolute number of affected cells increases with increasing throughput.

Variable	Baseline		Conservative		Progressive	
	w/o insp.	w/ insp.	w/o insp.	w/ insp.	w/o insp.	w/ insp.
Production rate (waf)	5,500	5,500	12,300	12,300	24,600	24,600
Defect detection accuracy (%)	0.00	90.00	0.00	90.00	0.00	90.00
Low eff. wafers if detected (waf)	458	458	1,025	1,025	2,050	2,050
Wafers low eff. not detected (waf)	5,500	5,500	12,300	12,300	24,600	24,600
Expected low eff. wafers (waf)	5,500	963	12,300	2,153	24,600	4,305
Expected low eff. wafers/year (waf)	481,800	84,315	1,077,480	188,559	2,154,960	377,118
Expected average efficiency (%)	24.9800	24.9965	24.9800	24.9965	24.9800	24.9965
Expected loss in avg. eff. (%)	0.0200	0.0035	0.0200	0.0035	0.0200	0.0035

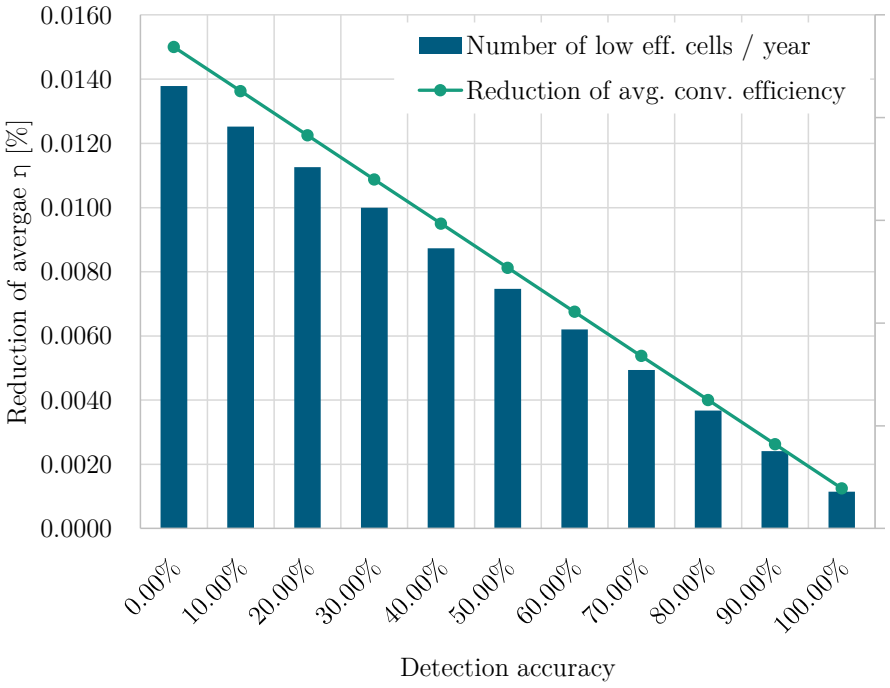


Figure 9.4: Expected loss in average conversion efficiency at a production throughput of 24,600 wph) at different detection accuracies. The bar chart shows the expected absolute number of wafers with η_{faulty}, i.e., W_{faulty}.

9.3 Additional Scenarios and LCA Connection

The impact of the intermediate results from this experiment chapter is tested in the LCA model from Blanco [67] introduced in Chapter 8. The chosen functional unit is 1 kWh electricity produced on a 7 kW_p slanted roof PV installation in Europe.

9.3.1 Definition of Additional Scenarios

To consider more than only the two defined scenarios from the algorithm level, further scenarios are defined in Table 9.5 to test for other potentially significant impact levers in a simplified manner. In the reference scenario, the model assumes a production in China, i.e., it also considers the Chinese energy mix as modeled in the ecoinvent 3.6 database. In the quality inspection process, a yield of 99.7 % is assumed, i.e., only 0.03 % of the cells that enter the quality inspection process are discarded. The reference module conversion efficiency is assumed with 25 %. This is an ambitious estimate as currently these efficiencies are currently achieved at cell level with TOPCon cells while modules show efficiencies around 22.5 % [4]. Because mainly changes with respect to the reference are investigated, the absolute number chosen only plays a minor role.

Table 9.5: Definition of scenarios to be tested in LCA model including underlying assumptions.

Scenario name	Changed parameters	Value	Underlying assumptions + motivation
Baseline	No changes	-	Assumes conversion efficiency of 25.0 % and wafer yield in sorting and testing of 0.997 (measurement system already present)
Rigorous 0 % acc.	Wafer yield in sorting and testing	0.99	Assumes no algorithm present. Process-related defects are only found within an hour. Defect wafers are discarded.
Rigorous 90 % acc.	Wafer yield in sorting and testing	0.9982	Assumes algorithm finding process-related defects within 5 minutes of production with 90 % accuracy. Defect wafers are discarded.
Rigorous bad process	Wafer yield in sorting and testing	0.95	To test what would happen in a very unstable process without quality inspection.
Efficiency 0 % acc.	PV conversion efficiency	0.2498	Assumes algorithm finding process-related defects within 5 minutes of production with 90 % accuracy. Defect wafers are discarded.
Efficiency 90 % acc.	PV conversion efficiency	0.249965	Assumes algorithm finding process-related defects within 5 minutes of production with 90 % accuracy. Defect wafers are discarded.
Efficiency process feedback bonus	PV conversion efficiency	25.5	Assumes significantly increased cell efficiency due to better process monitoring and fine-tuning.
Throughput increase 2	Electricity for clean room + electricity and pressured air in cell testing and sorting	2	Required clean room area and process consumables of cell testing and sorting reduce linearly with increasing throughput.
Throughput increase 4	Electricity for clean room + electricity and pressured air in cell testing and sorting	4	Required clean room area and process consumables of cell testing and sorting reduce linearly with increasing throughput.
Silver usage 50 %	Silver consumption in printing processes	0.5	Silver consumption in cell production is reduced to 50 %. Tests RSP scenario.
Silver usage 10 %	Silver consumption in printing processes	0.1	Silver consumption in cell production is reduced to 10 %. Tests RSP scenario.
Production DE	Energy mix location	DE (instead CN)	To roughly assess the impact of producing in Germany instead of China. Ignored the origin of other process consumables.
Very low efficiency	PV conversion efficiency	0.18	To test impact of low conversion efficiency.

9.3.2 Modeling Cleanroom Usage

While most changes in the LCA model are sufficiently explained in Table 9.5, the cleanroom energy demand per cell requires additional documentation. Solar cell manufacturing is a contamination-sensitive activity. Therefore, the manufacturing environment has to fulfill specific requirements regarding air quality. To achieve this, cleanrooms are commonly used which control the particulate contamination of the air. Cleanrooms are usually classified by the ISO 14644-1 standard [71] from ISO 9 (unclassified space such as a normal office) to ISO 1 (used in highly sensitive environments such as chip manufacturing processes for space applications). For solar cell manufacturing, vendors offer mixed ISO 7 and ISO 5 cleanrooms [72]. While the overall production environment is classified as ISO 7, some processes require ISO 5 zones with purer air.

Based on personal communication with Dr.-Ing. Sebastian Nold, 13,000 m^2 per $\mathrm{GW_p}$ production capacity per year is a good estimate for the total required building area in cell production. The sole manufacturing area makes up for around half of this estimate, i.e., 6,500 m^2 per $\mathrm{GW_p}$.[4] It is assumed that the cell manufacturing area of 6,500 m^2 per $\mathrm{GW_p}$ is certified as ISO 7 cleanroom. Additionally, it is assumed that doubling production throughput leads to half the used factory space per manufactured cell as shown in Equation 9.3.

$$A_{new} = \frac{throughput_{old}}{throughput_{new}} \cdot A_{old} \tag{9.3}$$

Here, A_{new} and A_{new} represent the new and old factory space, respectively. Cleanroom Technology [73] states an average power demand of 0.22 kW per m^2 for an ISO 7 cleanroom. This information is used to calculate the cleanroom-related energy consumption per cell as shown in Equations 9.4 and 9.5.

$$E_{cell} = P_{clean\ room\ 1\ m^2} \cdot A_{1\ GW_p} \cdot operating\ hours \cdot \frac{P_{cell}}{1\ GW_p} \tag{9.4}$$

With:

E_{cell}	: Energy for cleanroom per produced solar cell
$P_{clean\ room\ 1\ m^2}$	: Power required to operate 1 m^2 of cleanroom
$A_{1\ GW_p}$	: Ar cells equivalent to 1 $\mathrm{GW_p}$
P_{cell}	: Peak power of one solar cell

Using the assumptions from above, this amounts to:

$$E_{cell} = 0.22\frac{kW}{m^2} \cdot 6,500\ m^2 \cdot 8,760\ h \cdot \frac{7.5\ W_p}{1\ GW_p} = 0.093\ kWh \tag{9.5}$$

E_{cell} is added to the unit process *6.1 cell* in the LCA model as a technosphere inflow. Currently, the smaller ISO 5 cleanroom zones which are much higher in power demand (2.40 kW per m^2) [73] are not modeled. In order to get an even more detailed picture of possible impact reductions by factory space reductions, it is recommended to model this in greater detail.

[4] Dr.-Ing. Sebastian Nold (Fraunhofer ISE). Personal communication. July 25th, 2022.

9.3.3 Factory Building Materials

Next to the cleanroom-related energy consumption, the change in building materials used per cell is modeled as well. The ecoinvent 3.6 database [74] contains a unit process *photovoltaic cell factory construction* which is based on a solar cell factory built in 1999 in Germany for a yearly production of 10 million cells. This factory is already out of service, but the use of construction material to build this factory can still be used to estimate the difference in impact due to changing facility needs.

Using the factory area of this factory from 1999 as well as the current estimate of 13,000 m^2 per $\mathrm{GW_p}$ production capacity per year, the number of factories per cell is calculated assuming a factory lifetime of 25 years. The detailed calculations are shown in the *6.1 cells* sheet of the supplementary table of Appendix B - LCA Unit Processes. Based on the calculations, $5.50 \cdot 10^{-11}$ units of the 1999 factory are required per cell at current production efficiency which is added to the unit process *6.1 cell* in the LCA model as a technosphere inflow.

10 Integration of Sustainability Results

In this chapter, the findings from the initial assessment in Chapter 8 as well as the results from the scenario experiments in Chapter 9 are combined to a comprehensive analysis. The results on process level and algorithm level are projected into a large-scale level scenario as indicated in Figure 8.2 in Chapter 8.

10.1 Main Results of Assessment Scenarios

The impact compared in six impact categories using the ILCD 2.0 2018 midpoint indicators [75]. Next to *climate change total*, the *resources*, *ecosystem quality*, and *human health* categories are of special interest because the impact of silver usage reduction is investigated as well. Figure 10.1 shows a heatmap of the LCA results. Each row represents one tested scenario. The columns show the different impact categories. Each cell shows the relative change in impact for the respective scenario and impact category with respect to the baseline scenario. The cells are colored based on the impact change. Blue stands for a reduction in impact and red stands for an increase in impact.

The *rigorous bad process scenario* shows what would happen if 5 % of all produced cells are disposed after the cell testing and sorting process: The climate change impact would increase by 1.76 % with respect to the baseline scenario. However, this loss rate can be considered unrealistically high as the baseline scenario assumes a loss rate of 0.3 %.

More realistic assumptions in the scenarios that are supposed to test the impact of algorithm accuracy lead to much smaller changes in impact. In the rigorous scenario, the difference in climate change impact between 0 % detection accuracy and 90 % detection accuracy is 0.28 %. The efficiency scenarios show an even lower difference in the impact of 0.07 % in all impact categories between 0 % detection accuracy and 90 % detection accuracy. This shows the difficulties of demonstrating the often stated benefits of the “smart factory” vision. Either the simple probabilistic model developed in Chapter 9 is insufficient to capture the benefits or the benefits of a “smart factory” are lower than expected. While an assumed 0.5 % increase in PV conversion efficiency thanks to better process feedback could lead to reductions of roughly 2 % in all impact categories (efficiency bonus scenario), it remains to be proved how a “smart factory” could actually achieve these improvements.

Reducing silver usage per cell by 50 % as assumed in the application of rotational screen printing (RSP) in the rigorous scenario leads to a reduction of around 9 % in the *resources: minerals, and metals* impact category. If silver usage is even reduced to 10 % (currently unrealistic), the impact in this category would even decrease by 16 %. Reducing silver usage by 50 % also leads to

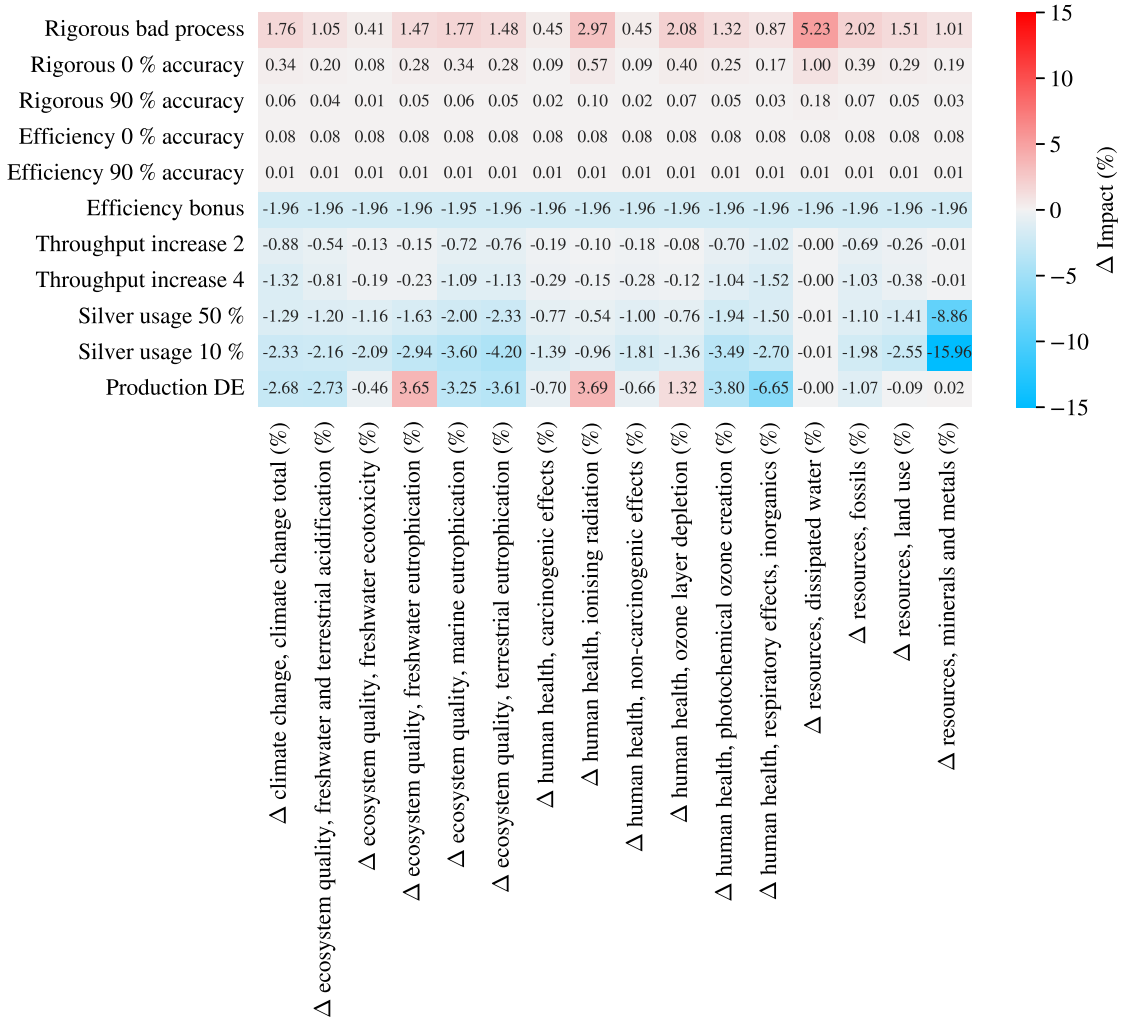
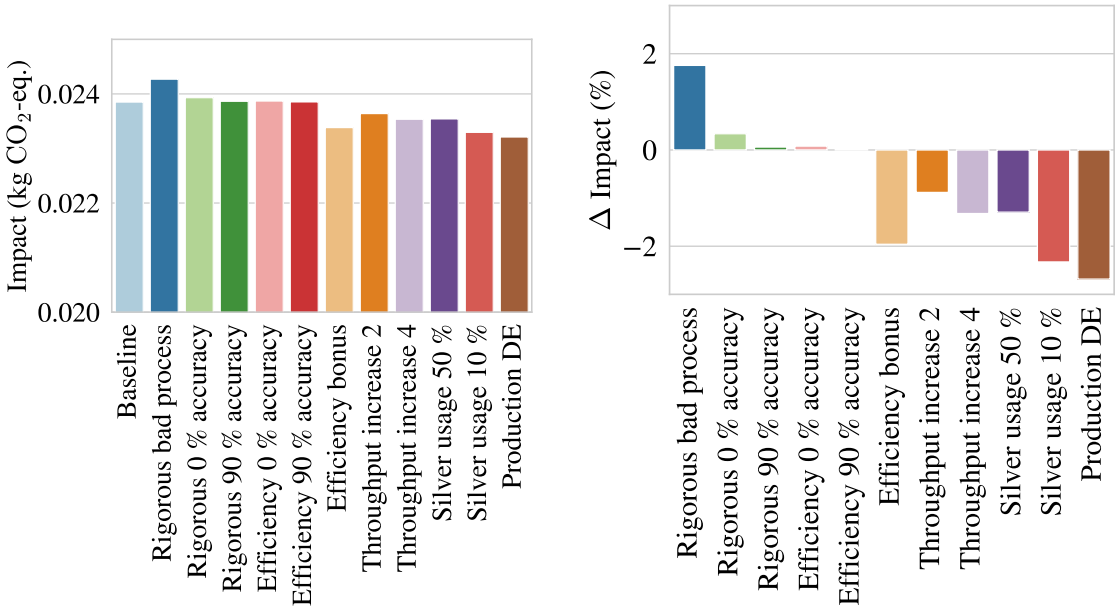


Figure 10.1: Relative results for four chosen impact categories with respect to the baseline scenario.

a reduction in impact in the toxicity categories. Compared to the baseline, the freshwater ecotoxicity impact of 1 kWh solar energy is reduced by 1.17 %. The decrease of carcinogenic effects is lower (reduced by 0.78 %).

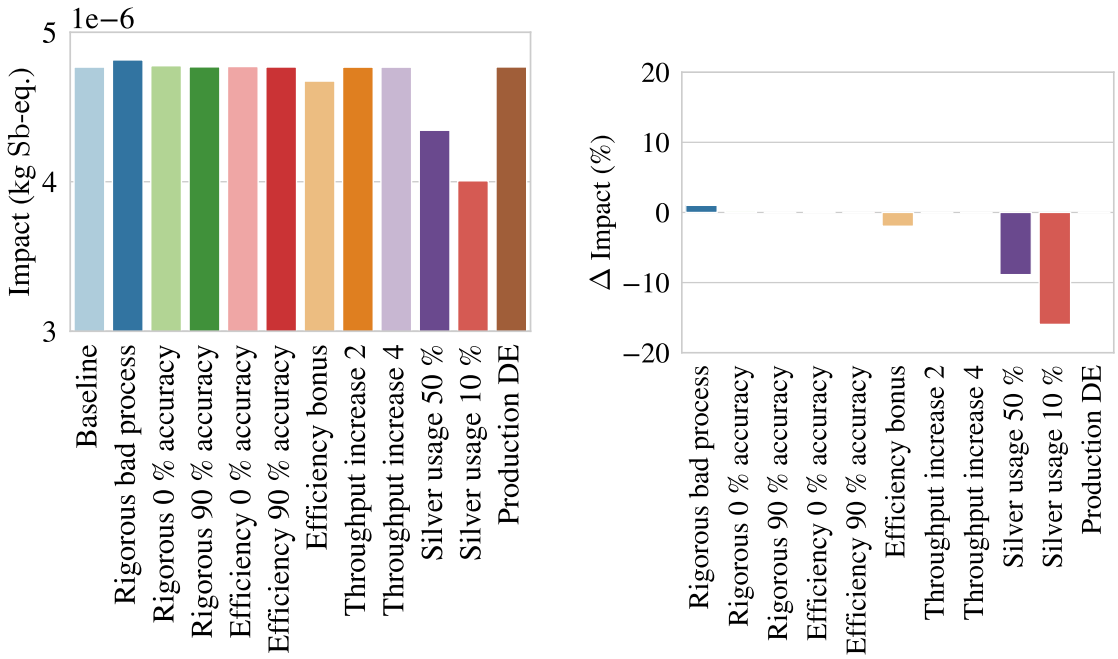
Figure 10.2 visualizes the absolute impact (in kg CO_2 -eq.) in (a) and difference relative impact with respect to the baseline scenario in the *climate change total* impact category in (b). Figure 10.3 visualizes absolute (a) and relative impact (b) in the *resources: minerals and metals* impact category which is especially impacted by silver usage reductions.



(a) Absolute change

(b) Relative change in impact w.r.t. baseline

Figure 10.2: LCA results for the impact category *climate change total.*



(a) Absolute change

(b) Relative change in impact w.r.t. baseline

Figure 10.3: LCA results for the impact category *resources: minerals and metals.*

10.2 Process Level

The direct change of factory space related to EL measurements is rather low. The EL system responsible for quality inspection requires around 20 m^2 factory space. At the current throughput

of 5,000 wafers per hour, this amounts to factory space requirements of 62 m^2 per $\mathrm{GW_p}$ per year.[1] At the target throughput of 16,000 wafers per hour, one machine could handle cells equivalent to slightly above 1 $\mathrm{GW_p}$ per year, i.e., the factory space requirement reduces to 19 m^2 per $\mathrm{GW_p}$ per year.

However, if a whole cell production line is viewed, the picture changes. The estimate of 13,000 m^2 factory space per $\mathrm{GW_p}$ production capacity per year is starting point for a rough estimate of the associated cost savings. COO of 1,500 Euro per m^2 fab area is assumed at a depreciation period of 10 years.[2] This leads to facility-related COO of 1.95 million Euros per $\mathrm{GW_p}$ production capacity which potentially could be lowered to 0.975 million Euros when halving the required factory space per cell. Additionally, utilities, i.e., resource consumption for electricity, cooling, etc., potentially decrease at a similar scale depending on the process.

At a per-cell level, the impact change due to saved building materials from the factory construction is neglectable as only $5.50 \cdot 10^{-11}$ units of the 1999 factory modeled in ecoinvent 3.6 are required per cell. The major change in impact in the *throughput increase* scenarios arises from reduced cleanroom usage due to factory space reductions. Doubling the production throughput could lead to reductions of 0.88 % in climate change impact, mainly due to reduced ISO 7 cleanroom usage per cell which in turn reduces electricity consumption. Quadrupling the throughput would lead to a climate change impact reduction of 1.32 %.

As a social factor to consider on process level, higher throughput leads to a reduction of staff required per cell reducing labor cost per $\mathrm{GW_p}$.

10.3 Large-Scale Level

The EU Market Outlook for Solar Power 2021-2025 [76] assumes that 700 $\mathrm{GW_p}$ in solar capacity have to be added until 2030 to stay on track in achieving the 1.5 °C Paris agreement. The found results from the process and algorithm level are projected into a large-scale scenario considering the need for PV growth in Europe within the next 5-10 years.

Overview

Production in Europe could become more attractive at increased production throughput. As shown in the process-level assessment, the factory area demand per cell reduces when the production throughput is scaled up and the labor costs decrease. Especially in relatively densely populated Europe with high regulatory barriers to opening new production facilities and high wages this makes local production more attractive. On top of that the EU Market Outlook from Solar Power Europe sees “a new window of opportunity for a domestic solar manufacturing sector” [76, p. 25] thanks to the current market growth phase.

[1] Dr. Johannes Greulich (Fraunhofer ISE). NextTec Equipment Data Input Sheet for Quality Inspection Processes. July 2nd, 2020.

[2] Dr.-Ing. Sebastian Nold, Lorenz Friedrich, Alma Spribille, Ingo Brucker (Fraunhofer ISE), SCost Model for NextTec Project. March 10th, 2020.

From a purely environmental perspective, it is not straightforward to judge if production in Europe leads to big benefits. Müller et al. [47] identified that glass-backsheet and glass-glass PV modules including PERC cells produced in the EU are linked to around 40 % less greenhouse gas (GHG) emissions than those produced in China, with the carbon intensity of electricity in production being one of the main levers. The results from the LCA in this work show a reduction in the impact of 2.68 % in climate change impact when the German electricity mix is used. The main difference is that Müller et al. [47] regionalized a higher number of processes.

It is worth noting that using the German electricity mix in production in the used LCA model increases freshwater eutrophication impact by 3.65 %. By investigating the process contributions to freshwater eutrophication in the German and Chinese electricity mix, the cause can be found: In China, the treatment of spoil from hard coal mining in a surface landfill is the biggest contributor to freshwater eutrophication. In Germany, the treatment of spoil from lignite mining in a surface landfill contributes most significantly to freshwater eutrophication. This indicates that the relatively high share of lignite coal in the German energy mix is responsible for the high eutrophication impact.

Projecting EU-Wide Silver Consumption and Understanding Its Impact

As a simplified example to emphasize the impact of the detection accuracy on the algorithm level, the silver losses assumed in the rigorous scenario in Chapter 9 are used to project the possible absolute silver losses at 0 % detection accuracy and 90 % detection accuracy. Silver is used because silver use is seen as one of the most critical factors in scaling up PV production to Terawatt scale [45]. Similar to the silver losses, rough estimates could be calculated for other emissions and resource use associated with cell production. Equation 10.1 is used to estimate the silver losses for considering the required production for the EU market until 2030.

$$Ag_{loss} = \frac{P_{total}}{P_{cell}} \cdot Ag_{cell} \cdot \%_{loss} \tag{10.1}$$

With:

Ag_{loss} : Total lost silver

P_{total} : Total peak power of produced cells

P_{cell} : Peak power of one solar cell

Ag_{cell} : Silver content of one solar cell

$\%_{loss}$: Solar cell yield loss rate

With no defect detection (assuming $\%_{loss} = 1\ \%$) the total silver losses until 2030 would amount to 78.4 metric tonnes at a power of 7.5 W per cell and a total production of 700 $\mathrm{GW_p}$. With 90 % detection accuracy, this reduces to 13.4 metric tonnes. In 2021, the global silver production was around 30,000 tonnes [77]. Assuming constant same silver production until 2030, the silver loss at $\%_{loss} = 1\ \%$ would amount to 0.0327 % of global production until 2030. This is reduced to

one-fifth in the scenario with 90 % detection accuracy.

Silver is a limited resource on planet earth and is considered a “precious metal” in the 2020 EU critical raw materials report [78]. Market analyses state that the production has decreased by 12 % since 2016 [79]. Therefore, basing the future renewable energy supply on this critical resource poses supply chain risks that should be avoided.

Next to possible economic consequences of relying on silver in the PV supply chain, especially silver mining is associated with various negative environmental and social impacts. The LCA results for the scenarios considering reduced silver usage give an indication of which impact categories are most affected by silver. The impact in the eutrophication-related categories reduces stronger than the climate change impact. Therefore, an independent investigation is carried out to understand which parts of the silver supply chain contribute most to these impact categories. Additionally, the impact category *freshwater and terrestrial acidification* is considered.

From Figure 10.4 it can be seen that the treatment of sulfidic tailings from silver mine operations is the main contributor to freshwater eutrophication. The ILCD 2.0 2018 impact categories only consider phosphorus, phosphoric acid, and phosphates as contributors to freshwater eutrophication. 1 kg of silver on the market leads to 1.47 P-eq. in freshwater eutrophication due to the release of phosphates from the tailings impoundment to long-term groundwater. Based on the global silver market modeled in ecoinvent 3.6 [74] the major eutrophication impact of silver from the global market occurs in Mexico (57 %), Australia (15 %), and the United States (11 %). Amongst others, freshwater eutrophication has negative impacts on aquatic life, and adverse effects on humans and domestic animals [80]. If long-term groundwater storage is affected by the discharges from tailing impoundment the freshwater supply of future human generations may be at risk in the areas surrounding the mines.

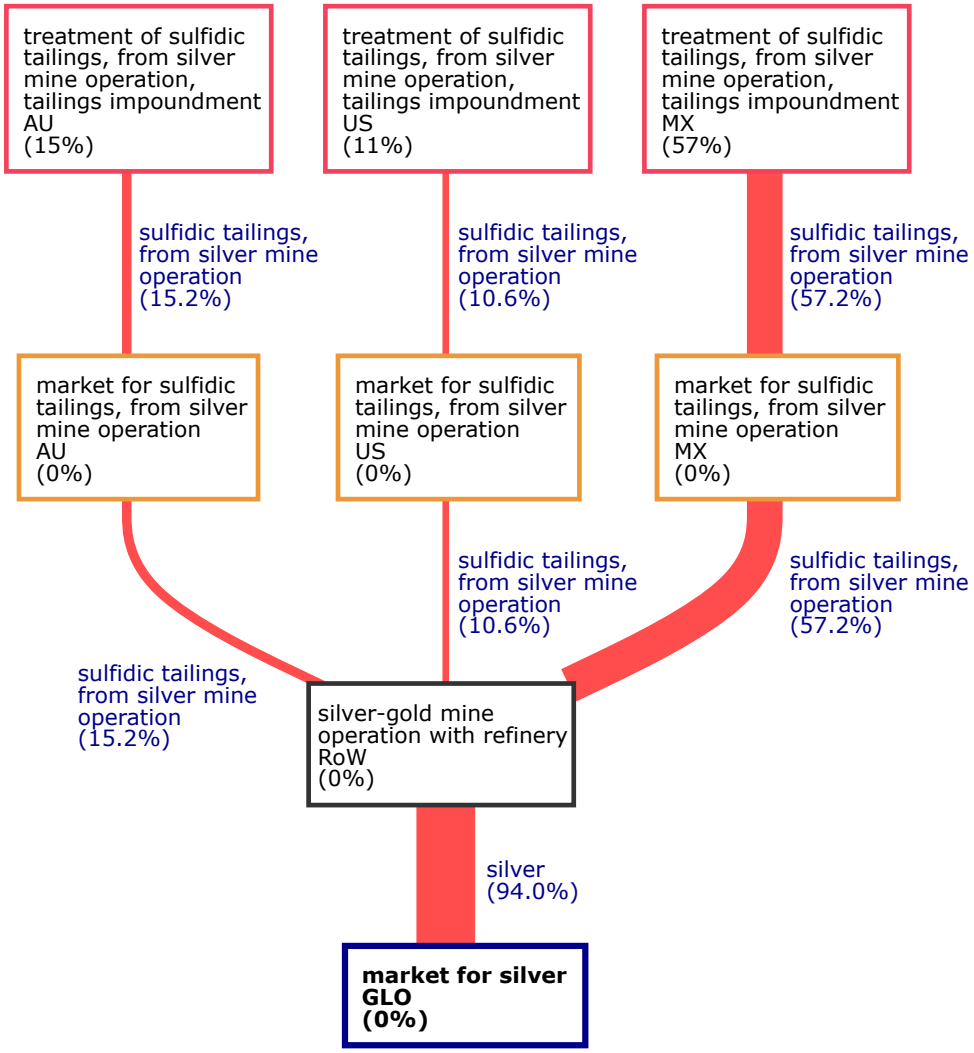


Figure 10.4: Sankey diagram of freshwater eutrophication impact for the functional unit of 1 kg of silver from the global market.

Figure 10.5 shows that blasting in mining operations contributes heavily to freshwater and terrestrial acidification (2.18 mol H^+-eq.). Acidification is associated with a loss in plant species richness due to a degradation in plant performance in acidic environments [81]. The main contributors to acidification are nitrogen and sulfur deposition [82]. In the ecoinvent *blasting* process, the detonation of 1 kg Tovex explosive releases 59 g of ammonia, a compound of nitrogen and hydrogen, as well as 331 g of nitrogen oxides into the air. The Sankey diagram for the *terrestial eutrophication* impact (not shown in this report) looks similar. This means blasting not only contributes to acidification but also to terrestrial eutrophication due to the ammonia and nitrogen oxide release.

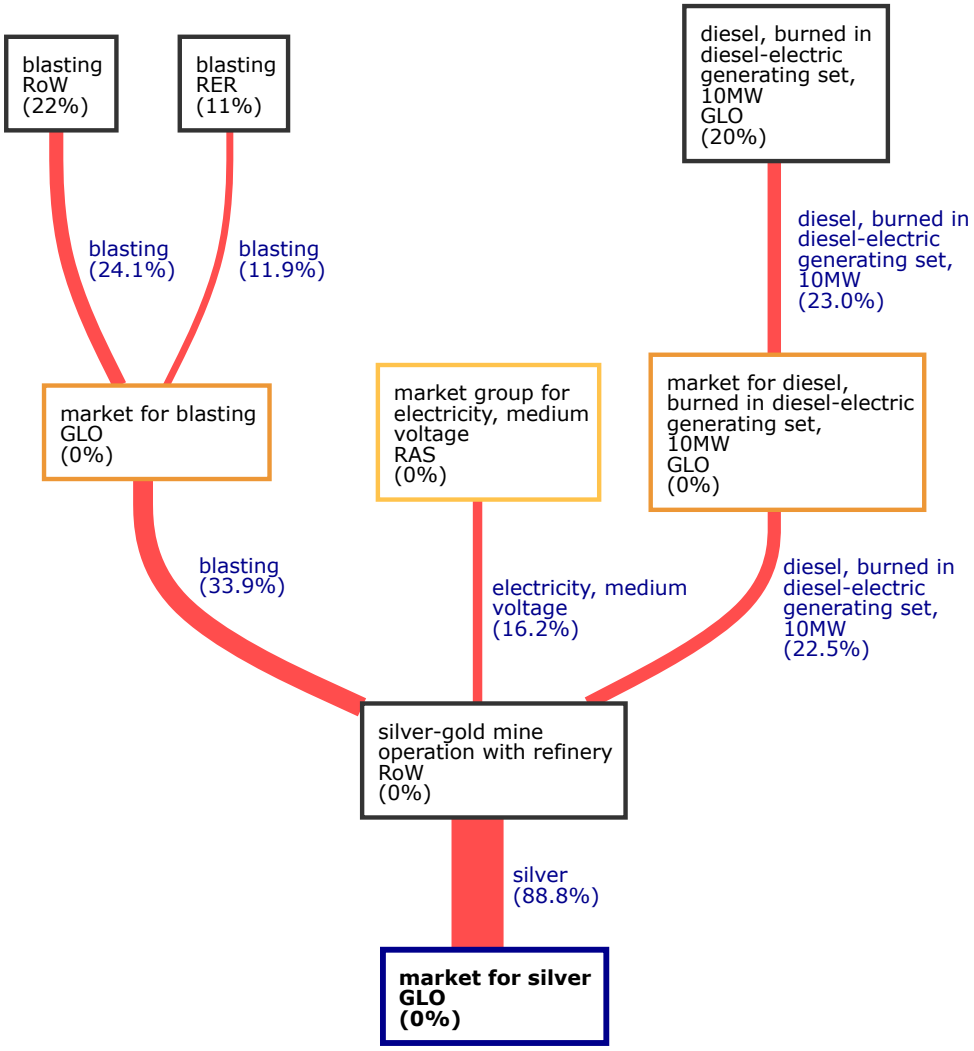


Figure 10.5: Sankey diagram of freshwater and terrestial acidification impact for the functional unit of 1 kg of silver from the global market.

Silver mining often takes place in remote areas. Land use change due to mine construction makes the land unavailable to both humans and previously present ecosystems. An important human health risk related to silver mining is the discharge of mercury [83]. Especially individual entrepreneurs (artisanal miners) carrying out small-scale mining use mercury amalgamation to recover silver from ore [84]. Artisanal miners are exposed to mercury and related health risks. However, mercury amalgamation processes can have even more far-reaching impacts. One example of negative health and social impacts is mercury discharge into water in remote areas of the Americas. Many communities rely on fish in their diet and may not be aware of both the presence of mercury in the ecosystem and the associated health risks [85]. This excludes humans and other species from safe food and water supply.

Overall, reducing silver usage and especially the associated mining is key in reducing the environmental impact of solar cell manufacturing. It is important to emphasize that this overview is by far not extensive, e.g., blasting releases many more harmful compounds that have an impact on human health such as aluminum. Sverdrup et al. [86] highlight the importance of consistent

recycling to increase the sustainability of the silver market supply and to avoid irreversible environmental damage. Increasing PV recycling rates as well as reducing the silver amount per solar cell can contribute to these goals.

Factory Space Consumption

The production of 700 $\mathrm{GW_p}$ in PV capacity for the EU market will require a significant increase in existing production capacity. It is unrealistic that the required 700 $\mathrm{GW_p}$ are produced in Europe until 2030 as the current cell production capacity in the EU is only around 0.8 $\mathrm{GW_p}$ per year as of today [76]. Assuming 100 $\mathrm{GW_p}$ yearly cell production capacity required to meet the EU's demands until 2030, this would amount to 1.3 $\mathrm{km^2}$ in cell production facilities $\mathrm{m^2}$ and facility-related COO of 1.95 billion Euros based on the assumptions stated above. However, as mentioned in the process-level investigations, the reduced facilities required per cell have almost no measurable environmental impact per produced cell or per kWh of solar energy in a Gigawatt scale production.

11 | Sustainability Discussion

The framework defined in this research leaves a lot of freedom to the applying individual. This is intended because this way, it can be applied in various ways: On the one hand, it is simplified enough to challenge current technical researchers to think about the impact of their work and possible sustainability implications without too much time investment required. On the other hand, it gives the opportunity to base profound design choices on it if the integration level is applied and a link to existing or new LCA studies is established.

The scaled approach helps to find contradictions: A sophisticated algorithm could help make a process faster or more efficient while being responsible for new, previously non-existing losses in energy or material. If only evaluated separately, those consequences could be undermined in the development process or be noticed too late. This also holds the other way around. For example, Verlinden [45] sees "no challenges" related to the technology in scaling up production but the technical part of this project clearly points out challenges that are still to be solved. They can potentially be overcome, but it is a good example of how high-level research ignores profound technological challenges.

Deploying spatially resolved quality inspection alone is not associated with any significant increase in environmental impact. In the used LCA model only the process consumables in quality inspection, i.e., electricity and pressured air are considered but also the non-considered additional equipment requirements are low. The hardware required to run a CV algorithm is usually already present in current production lines. Additionally, the burden of any potential additional hardware is shared with more cells at high throughput. Already one consumer-grade graphics card (GPU) for an entire production line is sufficient to execute the algorithm developed in this project. This is a significant difference from large-scale AI algorithm deployment, e.g., on social media platforms where the impact of inefficient algorithms scales up via millions of end-consumer devices or leads to massive computational effort in data centers.

The uncertainty of the LCA model and the assumptions can be regarded as high. Therefore, no absolute conclusion about the potential emissions savings should be drawn from this analysis. However, the results create value by adding different perspectives. The scenarios also come with limitations. The rigorous scenario in Chapter 9 is based on a prospective screen printing process. The calculations from Fraunhofer ISE are based on HJT cells. The LCA model is modeling PERC cells with the state of the art of 2018 but the current silver content of TOPCon cells. The algorithms in the technical part of the projects are optimized on TOPCon cells but can easily be adapted to other cell technologies. While this limits the accuracy of the results it does not change the order of magnitude of the results.

How high the environmental benefits of moving production facilities to Europe would be is dif-

ficult to answer. Most surprisingly changing from the Chinese to the German electricity mix in the cell production chain only leads to a reduction of 2.68 % of climate change impact per 1 kWh PV electricity production. Müller et al. [47] sees a far more significant reduction in emissions when changing the electricity mix in production. A main difference between the approach of both LCAs is that Müller et al. [47] uses 1 $\mathrm{kW_p}$ module manufacturing as functional unit as compared to 1 kWh electricity production. Additionally, [47] regionalize more processes and consider transport emissions as opposed to the model by [67] that is used in this project.

The impact of throughput increase is currently not extensively modeled in the LCA used in this project. Only in the quality inspection and sorting process, a throughput increase has been modeled, i.e., more cells share the burden of the electricity consumption of the EL measurement device. The advantages of throughput increase such as factory space reduction are shown outside of the LCA with some simplified calculations. While the calculations of factory space reductions are rough estimates, they reveal space and facility COO savings potential.

The LCA model could be improved by adding a more accurate measure of the energy consumption per factory space to consider the advantages arising from saving factory space when increasing production throughput in greater detail. Especially the need for higher clean room classes than ISO 7 in the production facilities should be considered. Furthermore, it is recommended to increase the level of detail by considering the impact of factory equipment production. For the quality inspection process, this may include the required information and communications technology such as graphics cards for CV algorithms and camera equipment.

Several studies indicate that increasing recycling rates and considering the end-of-life (EoL) phase will be of key importance and will become major challenges in the PV sector [45], [87]–[89]. In current research projects, this is still not directly interlinked: Technical research usually develops the technology faster than the technologies required for EoL treatment. The techno-economical assessments carried out for the project within Fraunhofer ISE are as well still mostly ignoring the EoL aspect. Especially research should show increased openness to presenting new business models and exploring recycling technology even if the monetary benefits are not immediately evident at this point in time.

It is argued that by contributing to a significant throughput increase a responsibility arises to consider EoL scenarios for the produced cells as well. The rigorous scenario in Chapter 9 shows that even if the material usage per cell is reduced significantly, there can still be higher absolute material losses in the production line in case the throughput is increased. When using *per cell metrics* or the functional unit of producing 1 kWh electricity this is easily overlooked. Especially for scarce resources, also absolute material losses should be considered. In the considered scenarios a higher algorithm accuracy is linked to lower absolute material losses.

As a future vision for ML research at Fraunhofer ISE, it should be explored how the knowledge gained from this high-end technological research can also be used for innovations with a more

significant direct impact on establishing more circular supply chains in the PV sector. Research on increasing production throughput should go hand in hand with research on increasing recycling throughput. For example, especially the lamination of the cells inhibits the disassembly of modules to recover materials from the cells [87]. There are start-ups and research teams working on a more modular module design that enhances maintainability and recyclability, e.g., *Biosphere Solar* [90] founded by students of TU Delft and Leiden University. These concepts demonstrate that sustainability improvements in the PV sector do not always require the most sophisticated technology.

Without the right technology to dissemble these modular solar panels at their end of use on a large scale, a more modular panel design does not help much at a large scale. Robots may be used to disassemble modules and separate the cells from them. Robust CV could play a key role to make this suitable for a variety of different modules. In fact, knowledge gained in high-speed semantic segmentation could potentially be applied to recycling robots.

With this discussion, the author seeks to inspire ML researchers at Fraunhofer ISE to keep their strong technical research focus while thinking about future research pathways that could have an even higher and more direct sustainability impact than current research. Within the field of PV research, Fraunhofer ISE is already an established institution that has proven its impact in many cases. Currently, the machine learning team of ISE is mostly focusing on quality inspection and characterization but is already involved in other projects such as the optimization of agrivoltaics. Founding an institute-wide competence team moving further forward with the research of applying ML techniques within the scope of achieving the SDGs could reveal huge potential.

Under the condition that EoL aspects are considered more in future technical research that concerns the upscaling of production, efforts directed towards increasing production throughput are favorable from a sustainability perspective. It can be assumed that c-Si PV is the fastest pathway towards the Terrawatt scale. The ITRPV (2021 results) [4] expects the mass production of Si-based tandem cells and modules to start around 2026. Even then the production of existing cell concepts will remain important. The results of this project point out that production in Europe might become more attractive thanks to reduced CAPEX, less staff required, and less required factory space per $\mathrm{GW_p}$ production capacity.

12 | Conclusion

12.1 Technical

The main technical research question of this project is *How well can electroluminescence images of TOPCon solar cells taken in motion at high production throughput be analyzed by deploying CNNs?*

This work focused on identifying key challenges in moving EL images. To do so, this project proposes the first known C-DCR-based Smart Labeling for TOPCon cells and the first known investigation of the impact of moving solar cells on detection accuracy. The results show it is possible to detect defects in moving EL images where human observers cannot detect much. Additionally, the experiments indicate that the refined concept of Smart Labeling can be applied to TOPCon solar cells. There is no known reference work that has focused on this combined challenge. Therefore, there are no references available to quantitatively compare the quality of the results. Compared to other publications on non-moving images mentioned in the related work (Chapter 3), the achieved quantitive detection accuracy is lower. It needs to be considered that results from related work are evaluated on non-reproducible human labels. On similar labels, it can be assumed that the chosen architecture in this project would lead to similar results as other publications. Because the results cannot easily be compared in numbers, this project contains a comprehensive qualitative analysis section.

The experiments show that a reduction in detection accuracy has to be expected on moving EL measurements. On still images, an F_1 score of 0.715 for R_S defects and 0.857 for j_0 defects in the best scenario is achieved. In the best scenario on real moving images, this reduces to 0.567 for R_S defects and 0.738 for j_0 defects at $T_L = 4\ ms$. This degradation in performance is not solely due to the motion blur. A low signal-to-noise ratio in the EL images due to low brightness resulting from short illumination times and high camera gain proved to be another main challenge.

A major challenge is aligning moving images with the corresponding labels to train the network. It has been attempted to avoid this by training networks with artificial motion blur. However, the network trained on artificial motion blur does not perform sufficiently on the real data. Estimating only the characteristics of linear motion blur occurring in moving measurements is not sufficient. More emphasis should be given to other contributors to the noise in the moving EL images.

Clear advice is given that images with motion blur should not be cropped too tightly as the border regions contain important image information that the CNN can make use of.

Another main challenge identified is the quality of C-DCR reference measurements. It is recommended to fine-tune the measurement technique to TOPCon to provide more homogeneous

output images for Smart Labeling. In the long-term, it could be desirable to predict physical characteristics or different degrees of defect severity by extending the binary classification. Additionally, the prediction could be expanded to more classes than only R_S and j_0, e.g., cracks and other relevant defects.

There are only a few C-DCR devices available across the world that could be used to acquire reference measurements for Smart Labeling. Companies or research institutes owning the equipment could potentially offer Smart Labeling as a service: Cell manufacturers send their cells to the "labeler" and receive the cells and labels back.

The future development of the moving EL measurements is still uncertain. For the EL measurement technique itself and for the image processing there is only little to no additional hardware required making it cost attractive. Next to improving the detection accuracy of the neural networks used for image analysis, there are still challenges to be solved regarding the process integration such as fast and reliable on-the-fly contacting.

Overall, the technical part of the project has delivered successful initial results and helped to gain important insights for further improvement of the concept of moving measurements.

12.2 Sustainability

The sustainability assessment in this project seeks to answer the research question *What is the prospective sustainability impact of novel quality inspection algorithms in high throughput PV production?*

The sustainability impact of the sole innovation in quality inspection alone is rather low but becomes significant if the innovation is viewed in a greater context. The results show that an increased average cell conversion efficiency of 0.5 % thanks to enhanced process feedback could lead to reductions of around 2 % per 1 kWh electricity from a 7 $\mathrm{kW_p}$ slanted roof PV installation in Europe in all impact categories. In one assumed scenario, early defect detection thanks to spatially resolved defect detection help reduce the yield loss after screen printing processes from 1 % to 0.18 %. This would lead to savings of 0.03 % of the yearly global silver production assuming a cell production of 100 $\mathrm{GW_p}$ per year. This could especially reduce acidification, eutrophication, mercury release, and other harmful consequences of silver mining.

Another important sustainability lever is the reduction of cleanroom usage when scaling up production throughput. Doubling throughput could lead to a decrease of 0.88 % in climate change emissions per kWh solar energy thanks to better cleanroom area utilization.

No immediate reasons were found that would advise against deploying AI-based high throughput quality inspection. The environmental impact of the algorithm due to energy consumption is neglectable on a per-cell level. This is because the algorithm is only deployed locally in the production line. Therefore, its impact does not scale up with increasing throughput. Still, especially

researchers seeking to limit their impact are advised to reconsider unnecessary (thoughtless) training of algorithms in the research and development phase.

In the short and mid-term, scaling up c-Si PV production is required to meet global market demand. Other more disruptive technologies are not expected to gain sufficient momentum within the next couple of years. Therefore, also seemingly marginal improvements to c-Si PV manufacturing processes are still highly welcome.

However, projects aiming at the further scale-up of production should automatically integrate an end-of-life (EoL) perspective to mitigate possible negative impacts made possible by technological advancements. Currently, this is often carried out in different research projects. The technical research in this project is still funded and carried out under the assumption of linear business models and does not include any circularity considerations. To cope with some of the biggest challenges arising from PV mass production it is recommended to Fraunhofer ISE to explore other possible applications of machine learning to contribute to the sustainable development of the entire PV supply chain even more significantly. The brief analysis of silver impact provides one example of why recycling is key in mitigating the negative consequences of mining. A vision developed in this thesis is to support research in circular module design and automatized disassembly assisted by CV building up on the knowledge gained from this and other research projects.

Bibliography

[1] F. Clement and S. Nold, *NextTec Project Proposal: Identifikation, Evaluation und Selektion von Produktionstechnologien der nächsten Generation mit dem Potential für eine signifikante Steigerung der Durchsatzraten in der PV-Produktion*, 2019.

[2] B. Kafle, B. S. Goraya, S. Mack, F. Feldmann, S. Nold, and J. Rentsch, "TOPCon – technology options for cost efficient industrial manufacturing," *Solar Energy Materials and Solar Cells*, vol. 227, p. 111 100, Aug. 2021. DOI: 10.1016/j.solmat.2021.111100.

[3] M. Alt, S. Fischer, S. Schenk, S. Zimmermann, K. Ramspeck, and M. Meixner, "Electroluminescence imaging and automatic cell classification in mass production of silicon solar cells," *2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC)*, pp. 3298–3304, Nov. 2018. DOI: 10.1109/PVSC.2018.8547983.

[4] VDMA e. V., "International technology roadmap for photovoltaic (ITRPV)," *Solar Energy Materials and Solar Cells*, 2022.

[5] P. Kunze, J. Greulich, S. Rein, et al., "Efficient Deployment of Deep Neural Networks for Quality Inspection of Solar Cells Using Smart Labeling," in *37th European Photovoltaic Solar Energy Conference and Exhibition*, WIP, Oct. 2020, pp. 427–432.

[6] A. Smets, K. Jäger, O. Isabella, R. van Swaaij, and M. Zeman, *Solar energy: The physics and engineering of photovoltaic conversion, technologies and systems*. Cambridge, England, UK: UIT Cambridge Ltd, 2016.

[7] G. M. Wilson, M. Al-Jassim, W. K. Metzger, et al., "The 2020 photovoltaic technologies roadmap," *Journal of Physics D: Applied Physics*, vol. 53, no. 49, p. 493 001, Sep. 2020. DOI: 10.1088/1361-6463/AB9C6A.

[8] B. Kafle, B. S. Goraya, S. Mack, F. Feldmann, S. Nold, and J. Rentsch, "TOPCon – technology options for cost efficient industrial manufacturing," *Solar Energy Materials and Solar Cells*, vol. 227, p. 111 100, Aug. 2021. DOI: 10.1016/j.solmat.2021.111100.

[9] A. Smets, K. Jäger, O. Isabella, R. Swaaij, van, and M. Zeman, *Solar energy : the physics and engineering of photovoltaic conversion technologies and systems*, English. UIT, 2016.

[10] A. Cuevas, "The recombination parameter j0," *Energy Procedia*, vol. 55, pp. 53–62, 2014, Proceedings of the 4th International Conference on Crystalline Silicon Photovoltaics (SiliconPV 2014). DOI: https://doi.org/10.1016/j.egypro.2014.08.073.

[11] T. Fuyuki and A. Kitiyanan, "Photographic diagnosis of crystalline silicon solar cells utilizing electroluminescence," *Applied Physics A 2008 96:1*, vol. 96, no. 1, pp. 189–196, Dec. 2008. DOI: 10.1007/S00339-008-4986-0.

[12] T. Trupke, R. A. Bardos, M. C. Schubert, and W. Warta, "Photoluminescence imaging of silicon wafers," *Applied Physics Letters*, vol. 89, no. 4, p. 044 107, Jul. 2006. DOI: 10.1063/1.2234747.

[13] M. Glatthaar, J. Haunschild, M. Kasemann, J. Giesecke, W. Warta, and S. Rein, "Spatially resolved determination of dark saturation current and series resistance of silicon solar cells," *physica status solidi (RRL) – Rapid Research Letters*, vol. 4, no. 1-2, pp. 13–15, Feb. 2010. DOI: 10.1002/PSSR.200903290.

[14] L. Li, J. Tu, Y. Yang, J. Wu, K. Hu, and S. Yu, "Effect of finger interruption mode on the performance of crystalline silicon solar cells," *Solar Energy*, vol. 238, pp. 381–391, 2022. DOI: https://doi.org/10.1016/j.solener.2022.03.065.

[15] H. Stolzenburg, A. Fell, F. Schindler, et al., "Edge recombination analysis of silicon solar cells using photoluminescence measurements," in *15th International Conference on Concentrator Photovoltaic Systems (CPV-15)*, AIP Publishing, 2019. DOI: 10.1063/1.5123822.

[16] S. Nold, *Presentation NextTec project meeting, 31.03.2022, Fraunhofer ISE internal*, 2022.

[17] J. Schmidhuber, "Deep Learning in Neural Networks: An Overview," *Neural Networks*, vol. 61, pp. 85–117, Apr. 2014. DOI: 10.1016/j.neunet.2014.09.003.

[18] J. Gu, Z. Wang, J. Kuen, et al., "Recent Advances in Convolutional Neural Networks,"

[19] A. Gron, *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, 1st. O'Reilly Media, Inc., 2017.

[20] A. Paszke, S. Gross, F. Massa, et al., "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, Eds., Curran Associates, Inc., 2019, pp. 8024–8035.

[21] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9351, pp. 234–241, May 2015.

[22] R. Lifset and T. E. Graedel, "Industrial ecology: goals and definitions," in *A Handbook of Industrial Ecology*, Northampton MA, USA: Edward Elgar Publishing, May 2002, pp. 3–15. DOI: 10.4337/9781843765479.00009.

[23] A. Kapur and T. E. Graedel, "Industrial Ecology," *Encyclopedia of Energy*, pp. 373–382, Jan. 2004. DOI: 10.1016/B0-12-176480-X/00533-7.

[24] T. E. Graedel and B. R. Allenby, *Industrial ecology*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall, 2003.

[25] R. Heijungs and S. Suh, *The computational structure of life cycle assessment*, ser. Eco-Efficiency in Industry and Science. Dordrecht, Netherlands: Springer, Dec. 2010.

[26] B. Su, H. Chen, and Z. Zhou, "Baf-detector: An efficient cnn-based detector for photovoltaic cell defect detection," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 3, pp. 3161–3171, 2021.

[27] B. Su, H. Chen, P. Chen, G. Bian, K. Liu, and W. Liu, "Deep learning-based solar-cell manufacturing defect detection with complementary attention network," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 6, pp. 4084–4095, 2020.

[28] Y. Jiang, C. Zhao, W. Ding, L. Hong, and Q. Shen, "Attention m-net for automatic pixel-level micro-crack detection of photovoltaic module cells in electroluminescence images," in *2020 IEEE 9th Data Driven Control and Learning Systems Conference (DDCLS)*, IEEE, 2020, pp. 1415–1421.

[29] J. Balzategui, L. Eciolaza, N. Arana-Arexolaleiba, et al., "Semi-automatic quality inspection of solar cell based on convolutional neural networks," in *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, IEEE, 2019, pp. 529–535.

[30] O. D. Singh, A. Malik, V. Yadav, S. Gupta, and S. Dora, "Deep segmenter system for recognition of micro cracks in solar cell," *Multimedia Tools and Applications*, vol. 80, no. 5, pp. 6509–6533, 2021.

[31] M. Mayr, M. Hoffmann, A. Maier, and V. Christlein, "Weakly supervised segmentation of cracks on solar cells using normalized l p norm," in *2019 IEEE International Conference on Image Processing (ICIP)*, IEEE, 2019, pp. 1885–1889.

[32] H. Chen, S. Wang, and J. Xing, "Detection of cracks in electroluminescence images by fusing deep learning and structural decoupling," in *2019 Chinese Automation Congress (CAC)*, IEEE, 2019, pp. 2565–2569.

[33] A. Bartler, L. Mauch, B. Yang, M. Reuter, and L. Stoicescu, "Automated detection of solar cell defects with deep learning," in *2018 26th European signal processing conference (EUSIPCO)*, IEEE, 2018, pp. 2035–2039.

[34] Z. Ying, M. Li, W. Tong, and C. Haiyong, "Automatic detection of photovoltaic module cells using multi-channel convolutional neural network," in *2018 Chinese Automation Congress (CAC)*, IEEE, 2019, pp. 3571–3576.

[35] A. Chindarkkar, S. Priyadarshi, N. S. Shiradkar, A. Kottantharayil, and R. Velmurugan, "Deep learning based detection of cracks in electroluminescence images of fielded pv modules," in *2020 47th IEEE Photovoltaic Specialists Conference (PVSC)*, IEEE, 2020, pp. 1612–1616.

[36] X. Zhang, T. Hou, Y. Hao, H. Shangguan, A. Wang, and S. Peng, "Surface defect detection of solar cells based on multiscale region proposal fusion network," *IEEE Access*, vol. 9, pp. 62 093–62 101, 2021.

[37] E. Sovetkin, E. J. Achterberg, T. Weber, and B. E. Pieters, "Encoder–decoder semantic segmentation models for electroluminescence images of thin-film photovoltaic modules," *IEEE Journal of Photovoltaics*, vol. 11, no. 2, pp. 444–452, 2020.

[38] R. Pierdicca, E. Malinverni, F. Piccinini, M. Paolanti, A. Felicetti, and P. Zingaretti, "Deep convolutional neural network for automatic detection of damaged photovoltaic cells.," *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, vol. 42, no. 2, 2018.

[39] A. S. Kovvali, M. Demant, and S. Rein, "Spatially resolved material quality prediction via constrained deep learning," in *2019 IEEE 46th Photovoltaic Specialists Conference (PVSC)*, IEEE, 2019, pp. 3059–3062.

[40] D.-M. Tsai, S.-C. Wu, and W.-C. Li, "Defect detection of solar cells in electroluminescence images using fourier image reconstruction," *Solar Energy Materials and Solar Cells*, vol. 99, pp. 250–262, Apr. 2012. DOI: 10.1016/j.solmat.2011.12.007.

[41] D. C. Tseng, Y. S. Liu, and C. M. Chou, "Automatic Finger Interruption Detection in Electroluminescence Images of Multicrystalline Solar Cells," *Mathematical Problems in Engineering*, vol. 2015, 2015. DOI: 10.1155/2015/879675.

[42] G. Dost, H. Höffler, and J. M. Greulich, "Advanced Series Resistance Imaging for Silicon Solar Cells via Electroluminescence," *physica status solidi (a)*, vol. 218, no. 6, p. 2 000 546, Mar. 2021. DOI: 10.1002/PSSA.202000546.

[43] R. Vinuesa, H. Azizpour, I. Leite, et al., "The role of artificial intelligence in achieving the Sustainable Development Goals," *Nature Communications*, vol. 11, no. 1, Dec. 2020. DOI: 10.1038/s41467-019-14108-y.

[44] F. W. Geels, "Technological transitions as evolutionary reconfiguration processes: A multi-level perspective and a case-study," *Research Policy*, vol. 31, no. 8, pp. 1257–1274, 2002, NELSON + WINTER + 20. DOI: https://doi.org/10.1016/S0048-7333(02)00062-8.

[45] P. J. Verlinden, "Future challenges for photovoltaic manufacturing at the terawatt level," *Journal of Renewable and Sustainable Energy*, vol. 12, no. 5, p. 053 505, 2020. DOI: 10.1063/5.0020380. eprint: https://doi.org/10.1063/5.0020380.

[46] M. Aßmus, S. Nold, S. Rein, M. Hofmann, J. Rentsch, and R. Preu, "Performance requirements of crack detection systems in silicon solar cell production," *Energy Procedia*, vol. 27, pp. 147–153, Dec. 2012. DOI: 10.1016/j.egypro.2012.07.043.

[47] A. Müller, L. Friedrich, C. Reichel, S. Herceg, M. Mittag, and D. H. Neuhaus, "A comparative life cycle assessment of silicon PV modules: Impact of module design, manufacturing location and inventory," *Solar Energy Materials and Solar Cells*, vol. 230, p. 111 277, Sep. 2021. DOI: 10.1016/J.SOLMAT.2021.111277.

[48] C. F. Blanco, S. Cucurachi, J. B. Guinée, et al., "Assessing the sustainability of emerging technologies: A probabilistic LCA method applied to advanced photovoltaics," *Journal of Cleaner Production*, vol. 259, p. 120 968, Jun. 2020. DOI: 10.1016/j.jclepro.2020.120968.

[49] C. H. Wu, K.-k. Tseng, C. K. Ng, and W. H. Ip, "An effective motion-blurred image restoration approach for automated optical inspection," *HKIE Transactions*, vol. 22, no. 4, pp. 252–262, 2015. DOI: 10.1080/1023697X.2015.1102655. eprint: https://doi.org/10.1080/1023697X.2015.1102655.

[50] J. M. Greulich, M. Demant, P. Kunze, et al., "Comparison of Inline Crack Detection Systems for Multicrystalline Silicon Solar Cells," *IEEE Journal of Photovoltaics*, vol. 10, no. 5, pp. 1389–1395, Sep. 2020. DOI: 10.1109/JPHOTOV.2020.2996750.

[51] E. Sovetkin, E. Achterberg, T. Weber, and B. Pieters, "Encoder–decoder semantic segmentation models for electroluminescence images of thin-film photovoltaic modules," *IEEE Journal of Photovoltaics*, vol. PP, pp. 1–9, Dec. 2020. DOI: 10.1109/JPHOTOV.2020.3041240.

[52] T. R. Singh, S. Roy, O. I. Singh, T. Sinam, and K. M. Singh, *A new local adaptive thresholding technique in binarization*, 2012. arXiv: 1201.5227 [cs.CV].

[53] OpenCV. "Image thresholding." (2022), [Online]. Available: https://docs.opencv.org/3.4/d7/d4d/tutorial_py_thresholding.html (visited on 05/16/2022).

[54] H. Iqbal, *Harisiqbal88/plotneuralnet v1.0.0*, 2018. DOI: 10.5281/ZENODO.2526396.

[55] S. Ioffe and C. Szegedy, *Batch normalization: Accelerating deep network training by reducing internal covariate shift*, 2015. DOI: 10.48550/ARXIV.1502.03167.

[56] PyTorch Contributors. "CrossEntropyLoss - PyTorch 1.12 documentation." (2022), [Online]. Available: https://pytorch.org/docs/stable/generated/torch.nn.CrossEntropyLoss.html (visited on 05/17/2022).

[57] PyTorch Contributors. "Adam Optimizer - PyTorch 1.12 documentation." (2022), [Online]. Available: https://pytorch.org/docs/stable/generated/torch.optim.Adam.html#torch.optim.Adam (visited on 05/17/2022).

[58] Y. Sasaki, "The truth of the F-measure," 2007.

[59] W. Zhou, X. Hao, K. Wang, et al., "Improved estimation of motion blur parameters for restoration from a single image," *PLOS ONE*, vol. 15, no. 9, R. Damasevicius, Ed., e0238259, Sep. 2020. DOI: 10.1371/journal.pone.0238259.

[60] D.-H. Kim and I. Gratchev, "Application of optical flow technique and photogrammetry for rockfall dynamics: A case study on a field test," *Remote Sensing*, vol. 13, no. 20, 2021. DOI: 10.3390/rs13204124.

[61] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, pp. 600–612, May 2004. DOI: 10.1109/TIP.2003.819861.

[62] Y. Buratti, A. Sowmya, R. Evans, T. Trupke, and Z. Hameiri, "Half and full solar cell efficiency binning by deep learning on electroluminescence images," *Progress in Photovoltaics: Research and Applications*, vol. 30, no. 3, pp. 276–287, 2022. DOI: https://doi.org/10.1002/pip.3484. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pip.3484.

[63] P. Kunze, S. Rein, M. Hemsendorf, K. Ramspeck, and M. Demant, "Learning an empirical digital twin from measurement images for a comprehensive quality inspection of solar cells," *Solar RRL*, vol. 6, no. 5, p. 2 100 483, 2022.

[64] Wealth Daily, *Special Report: Solar Technology*, 2017.

[65] F. Dimroth and S. Wrigley. "SiTaSol." (2022), [Online]. Available: https://sitasol.com/ (visited on 08/23/2022).

[66] EU Publications Office, "SiTaSol (application relevant validation of c-si based tandem solar cell processes with 30 % efficiency target)," *CORDIS EU research results*, 2021. DOI: 10.3030/727497.

[67] C. F. Blanco, "Guiding safe and sustainable technological innovation under uncertainty: A case study of iii-v/silicon photovoltaics," Ph.D. dissertation, Leiden University, Leiden, The Netherlands, 2022.

[68] C. Mutel, "Brightway: An open source framework for life cycle assessment," *The Journal of Open Source Software*, vol. 2, no. 12, p. 236, Apr. 2017. DOI: 10.21105/joss.00236.

[69] B. Steubing, D. de Koning, A. Haas, and C. L. Mutel, "The activity browser — an open source LCA software building on top of the brightway framework," *Software Impacts*, vol. 3, p. 100 012, Feb. 2020. DOI: 10.1016/j.simpa.2019.100012.

[70] T. Hatt, J. Bartsch, S. Schellinger, et al., "Stable copper plated metallization on shj solar cells amp; investigation of selective al/alox laser patterning," en, *38th European Photovoltaic Solar Energy Conference and Exhibition; 326-329*, 2021. DOI: 10.4229/EUPVSEC20212021-2DV.3.27.

[71] International Organization for Standardization, *ISO Standard 14644-1. Cleanrooms and associated controlled environments. Part 1: Classification of air cleanliness.* Case Postale 56, CH-1211 Genève 20, Switzerland, 1999.

[72] ProCleanroom. "Cleanroom - solar cell production - cleanroom specifications." (2022), [Online]. Available: https://procleanroom.com/portfolio/cleanroom-solar-cell-production/ (visited on 08/23/2022).

[73] Cleanroom Technology. "Saving energy in cleanrooms." (2014), [Online]. Available: https://www.cleanroomtechnology.com/news/article_page/Saving_energy_in_cleanrooms/100623 (visited on 08/23/2022).

[74] G. Wernet, C. Bauer, B. Steubing, J. Reinhard, E. Moreno-Ruiz, and B. Weidema, "The ecoinvent database version 3 (part i): Overview and methodology," *The International Journal of Life Cycle Assessment*, vol. 21, no. 9, pp. 1218–1230, Apr. 2016. DOI: 10.1007/s11367-016-1087-8.

[75] S. Fazio, V. Castellani, S. Sala, et al., "Supporting information to the characterisation factors of recommended ef life cycle impact assessment methods: New methods and differences with ilcd," no. KJ-NA-28888-EN-N, 2018. DOI: 10.2760/671368.

[76] SolarPower Europe, "EU market outlook for solar power 2021-2025," *Solar Energy Materials and Solar Cells*, 2021.

[77] The Silver Institute. "World silver survey 2022." (2022), [Online]. Available: https://www.silverinstitute.org/wp-content/uploads/2022/04/World-Silver-Survey-2022.pdf (visited on 08/23/2022).

[78] European Commission. "Study on the EU's list of critical raw materials – final report." (2020), [Online]. Available: https://ec.europa.eu/docsroom/documents/42883/attachments/1/translations/en/renditions/native (visited on 08/23/2022).

[79] J. Clark. "New silver supply is drying up faster than death valley." (2022), [Online]. Available: https://www.cpmgroup.com/update-new-silver-supply-is-drying-up-faster-than-death-valley/ (visited on 08/24/2022).

[80] W. K. K. Dodds and E. B. Welch, "Establishing nutrient criteria in streams," *Journal of the North American Benthological Society*, vol. 19, no. 1, pp. 186–196, Mar. 2000. DOI: 10.2307/1468291.

[81] L. B. Azevedo, R. van Zelm, A. J. Hendriks, R. Bobbink, and M. A. Huijbregts, "Global assessment of the effects of terrestrial acidification on plant species richness," *Environmental Pollution*, vol. 174, pp. 10–15, 2013. DOI: https://doi.org/10.1016/j.envpol.2012.11.001.

[82] F. Dentener, J. Drevet, J. F. Lamarque, et al., "Nitrogen and sulfur deposition on regional and global scales: A multimodel evaluation," *Global Biogeochemical Cycles*, vol. 20, no. 4, 2006. DOI: https://doi.org/10.1029/2005GB002672. eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2005GB002672.

[83] S. Strode, L. Jaeglé, and N. E. Selin, "Impact of mercury emissions from historic gold and silver mining: Global modeling," *Atmospheric Environment*, vol. 43, no. 12, pp. 2012–2017, 2009. DOI: https://doi.org/10.1016/j.atmosenv.2009.01.006.

[84] W. Salomons, *Mercury from Gold and Silver Mining: A Chemical Time Bomb?* Jan. 1998. DOI: 10.1007/978-3-642-58793-1.

[85] M. Lucotte, R. Canuel, and S. B. de Grosbois, "Mercury cycling and human health concerns in remote ecosystems in the americas," 2009.

[86] H. Sverdrup, D. Koca, and K. V. Ragnarsdottir, "Investigating the sustainability of the global silver supply, reserves, stocks in society and market price using different approaches," *Resources, Conservation and Recycling*, vol. 83, pp. 121–140, 2014. DOI: https://doi.org/10.1016/j.resconrec.2013.12.008.

[87] S. Weckend, A. Wade, and G. Heath, "End of life management: Solar photovoltaic panels," Tech. Rep., Aug. 2016. DOI: 10.2172/1561525.

[88] L. El-Khawad, D. Bartkowiak, and K. Kümmerer, "Improving the end-of-life management of solar panels in germany," *Renewable and Sustainable Energy Reviews*, vol. 168, p. 112 678, Oct. 2022. DOI: 10.1016/j.rser.2022.112678.

[89] Gervais, Estelle, Herceg, Sina, Nold, Sebastian, and Weiß, Karl-Anders, "Sustainability strategies for pv: Framework, status and needs," *EPJ Photovolt.*, vol. 12, p. 5, 2021. DOI: 10.1051/epjpv/2021005.

[90] Biosphere Solar. "Biosphere Solar - A circular revolution in solar energy." (2022), [Online]. Available: https://www.biosphere.solar/wp-content/uploads/2022/07/Biosphere-Solar-V.0.3.-Download-Kit-1.zip (visited on 07/26/2022).

Appendix A - Code and Models

The archive file *Appendix A - Code and Models - DetectionInMotion.zip* contains the program code created in this thesis project. The code of the ISE-internal "Phillephalle" package used for the organization of solar cell samples is not included. Furthermore, the ISE-internal alignment package used to rectify and crop measurement images is not included.

The project folder in the archive consists of two main parts looking from the root folder:
The *NN_experiment_files* folder contains trained models of all scenarios and related plots. In the submitted version, this folder does not contain all created files due to size restrictions (all files of the project require 308 GB of storage). However, for each experiment, one model file of the best-performing scenario for each illumination is submitted.

The *detectioninmotion* folder consists of the following subfolders:

- *alignment* (scripts for the rectangular alignment of measurement images)
- *dashboard* (scripts of an application used in the final defense, not part of this report)
- *database_creation* (scripts to create a solar cell database based on measurements at Fraunhofer ISE)
- *smart_labeling* (scripts to calculate smart labels for R_S and j_0 defects
- *neural_networks* (main scripts for the neural network experiments, including training, test, and validation)
- *blurring* (contains a script to investigate artificial motion blur and to create clean plots of samples with motion blur applied)
- *tables_plotting* (scripts for plotting publishing-grade plots of various types)

A more detailed overview of the contained scripts and their purpose is given in the *README.md* file in the archive.

Appendix B - LCA Unit Processes and Absolute Results

The LCA used in this work is based on an existing model from Blanco [67]. To reproduce the results, please contact the model author mentioned in the bibliography. The background processes from ecoinvent 3.6 remain unchanged. The foreground processes changed in the model are:

- 2.4 Ag nanoink printing TOPCon (varied silver amount)
- 5.13 Cell testing and sorting (varied throughput and yield loss)
- 5.8 PVD rear metal (varied silver amount)
- 6.5 Electricity production, slanted-roof installation (varied PV conversion efficiency)
- 6.1 Cell (added cleanroom energy consumption and factory construction, varied throughput)

The archive file *Appendix B - LCA Details - DetectionInMotion.zip* contains the LCA-related files. The detailed process changes and associated calculations are documented in the file *Appendix B - LCA Unit Processes.xlsx* which is part of the submission. The analysis in the report mostly focuses on the relative impact changes with respect to the baseline scenario. The absolute results in all ILCD 2.0 2018 midpoint impact categories are available in the file *Appendix B - LCA Absolute Results.xlsx*.