

# Computer vision-based method for quantifying iron-related defects in silicon solar cell

Oleg Olikh<sup>1,‡</sup>, Oleksii Zavhorodnii<sup>1</sup>, Yulia Perets<sup>1</sup>

<sup>1</sup>Physics Faculty, Taras Shevchenko National University of Kyiv, Kyiv 01601, Ukraine

E-mail: olegolikh@knu.ua

**Abstract.** abstract abstract abstract abstract abstract abstract

*Keywords:* defect, Si

Submitted to: *Semicond. Sci. Technol.*

<sup>‡</sup> Author to whom any correspondence should be addressed.

## 1. Introduction

Due to the urgent need to address environmental challenges and the growing global demand for renewable energy, the deployment of photovoltaic (PV) systems has been rapidly increasing worldwide. In particular, solar PV generation exceeded 1,600 TWh in 2023 [?, ?], rose by about 30% in 2024 [?], and forecasts indicate that total installed capacity will surpass 6 TW by 2030 [?]. Meanwhile, crystalline silicon photovoltaics, which have benefited from decades of scientific advancement and continuous cost reductions, continued to dominate the market in 2024, accounting for approximately 98% of the global share [?, ?].

As in other semiconductor devices, defects play a decisive role in determining the operating parameters of solar cells. Therefore, diagnosing defects, particularly determining their concentrations, is critically important for maintaining the stable performance of PV systems. In recent years, researchers have increasingly complemented established defect-characterisation methods with machine learning (ML) approaches that improve the accuracy, speed, and cost efficiency of these analyses. The use of ML methods for analysing macroscopic defects (such as cracks, finger failures, hotspots, and scratches) and point defects, however, differs significantly. Researchers typically detect macroscopic defects in PV systems using two main approaches [?, ?]. The first approach, electrical testing techniques, involves analysing characteristic electrical curves of parameters such as current, voltage, and power. The second approach, Imaging-Based Techniques, involves analysing electroluminescence (EL) [?] or photoluminescence (PL) [?] images of solar cells. Numerous review studies demonstrate extensive use of ML in both approaches [?, ?, ?, ?, ?, ?, ?].

Regarding point defects, although they represent one of the primary limitations of PV devices, researchers have developed considerably fewer ML techniques specifically for their analysis. Existing ML-based approaches include extracting recombination-active centre parameters from lifetime measurements [?, ?, ?], detecting radiation-induced defects using Raman spectroscopy [?, ?], estimating the concentration of contaminant impurities from the ideality factor of current-voltage characteristics [?], and analysing variations in photovoltaic conversion parameters [?]. Nevertheless, researchers still offer a relatively limited number of such methods.

One of the main challenges in applying ML methods effectively is that training the models requires a large amount of labelled data [?]. In practice, researchers often cannot obtain such volumes of experimental data; therefore, they commonly use approaches such as simulations, in which they simulate hundreds of thousands of dependencies [?, ?, ?, ?, ?],

Physics-Informed Neural Networks (PINNs), which incorporate physical laws into the loss function to generate synthetic data [?, ?], or Transfer Learning, in which researchers adapt a model trained on one task to a related task [?, ?]. Simulations can be highly demanding in terms of time and computational resources. PINNs are primarily suited for phenomena described by partial differential equations, and researchers do not have pre-trained models available for all types of physical problems. At the same time, one of the most widely addressed tasks in machine learning is computer vision, for which researchers have publicly released many pre-trained models. For example, researchers trained EfficientNetB7 on approximately 1.2 million images from standard datasets.

This work primarily aimed to apply standard pre-trained computer vision models to analyse electrophysical measurement results related to point defects. In particular, we focused on quantifying iron in boron-doped crystalline silicon solar cells by examining the short-circuit current ( $I_{sc}$ ) relaxation following intense illumination. Iron represents one of the most prevalent, ubiquitous, and efficiency-limiting metallic impurities in such structures [?, ?], which, as noted earlier, form the foundation of the photovoltaic market. In Si:B, iron tends to form iron-boron pairs under equilibrium conditions, which intense illumination can dissociate [?, ?]. Observed changes in  $I_{sc}$  directly reflect the recovery process of the iron-boron pairs [?].

Researchers commonly use well-established benchmark architectures from computer vision, such as YOLO, MobileNetV2, EfficientNet, ResNet, Xception, and GoogleNet, to detect macroscopic defects based on EL measurements [?, ?, ?, ?, ?, ?, ?, ?]. In these cases, the measurement output consists of images, which makes the approach relatively straightforward. In our study, we needed to convert temporal dependencies into images. To address this task, we applied standard techniques such as the Fourier transform or wavelet transform. In photovoltaics, researchers typically apply wavelet transforms to images of solar cells to enhance macroscopic defect detection [?, ?]. However, wavelet transforms can also convert one-dimensional non-stationary signals into two-dimensional spectrograms, allowing efficient extraction of subtle features [?].

By applying computer vision models to wavelet spectrograms represented as images, we generated high-dimensional feature vectors and used them as inputs for traditional regression models. We obtained the  $I_{sc}(t)$  dependencies for training the regression models through both simulation and experimental measurements. In both cases, the hybrid Transfer Learning approach produced predictions with

sufficiently high accuracy (within a few percent), even when training models on small datasets containing fewer than 30 samples. Importantly, our approach is highly versatile and can be applied to a wide range of tasks, not only defect characterisation but also to other applications beyond photovoltaics.

## 2. Experimental details

### 2.1. General Scheme

Figure ?? illustrates the workflow of the ML pipeline used to extract iron contamination from  $I_{sc}(t)$  dependencies. The process consists of three main blocks: Data Acquisition, CNN (Convolutional Neural Network) Feature Processing, and Predictive Regression. In the first stage, we either simulated or experimentally measured the time-dependent short-circuit current of a solar cell after inducing FeB-pair dissociation. ?? and ?? describe these procedures in detail. For experimental curves, we smoothed the data using a Savitzky–Golay filter [?]. Next, we applied a continuous wavelet transform (CWT) [?] to convert one-dimensional temporal dependencies into two-dimensional spectrograms, represented as images. Each pixel in these images corresponds to the amplitude of a wavelet coefficient at a specific time and frequency. We used the Morlet wavelet for this transformation and implemented it with the Python library PyWavelets. Figure ?? and Figure ?? show representative examples of the resulting images. To expand the dataset, we augmented the generated images by flipping them along the x- and y-axes and rotating them by 90°, 180°, and 270°. Such augmentation procedures are well known to improve the prediction accuracy of ML models, especially when working with small datasets [?].

During the CNN Feature Processing stage, we processed all images, both original and augmented, using standard computer vision models to extract feature sets for each image. ?? describes the models and the corresponding feature extraction methods. We did not fine-tune the CNN models; instead, we used them in their pre-trained form as downloaded. In general, the feature dimensionality produced by CNN models greatly exceeds the number of available samples, meaning that many extracted features are not independent. To address this imbalance and facilitate comparison of results, we applied Principal Component Analysis (PCA) in selected cases. PCA reduces feature dimensionality while preserving most of the total variance in the data.

The obtained feature sets served as inputs to regression models based on one of the standard algorithms described in ??, which aimed to predict the iron concentration ( $N_{Fe}$ ) in the solar cell. The models

were trained on a training dataset and validated on a test dataset using the evaluation metrics outlined in ??. During training, we treated the feature sets corresponding to the original wavelet spectrogram and the augmented images as individual samples. During testing, we obtained the final prediction by taking the median value from the outputs corresponding to the augmented images. In one case, we trained the models on simulated data and tested them on both simulated and experimental datasets. In another, we used part of the experimental data for training and the remaining portion for testing.

### 2.2. Simulation details

This study performed I–V curve simulations for a silicon  $n^+p$ - $p^+$  structure under monochromatic illumination using SCAPS 3.3.11 to obtain the  $I_{sc}(t)$  dependencies. The SCAPS-1D software [?] serves as a widely used tool for modelling solar cells with defect-state effects [?, ?, ?, ?, ?, ?]. The study extracted  $I_{sc}$  values from the simulated I–V curves following a standard procedure [?].

During the simulations, we assumed that the base thickness of the structure was 380  $\mu\text{m}$ , that the dopant was boron with a concentration of  $N_B = 1.36 \times 10^{15} \text{ cm}^{-3}$ , and that the temperature was 340 K. We provided illumination using light with a wavelength of 940 nm and an intensity of 5  $\text{W/m}^2$ , matching the experimental conditions (see ??). We set the total concentration of iron impurity atoms,  $N_{Fe}$ , as one of the simulation parameters. We assumed that the atoms were uniformly distributed throughout the base and  $p^+$  layer of the solar cell and could exist either in the interstitial position (concentration  $N_{Fe_i}$ ) or as iron–boron pairs (concentration  $N_{FeB}$ ). The time dependence of  $N_{Fe_i}$  after pair dissociation follows the well-known expression [?, ?]:

$$N_{Fe}(t) = (N_{Fe_i,0} - N_{Fe,eq}) \times \exp(-t/\tau_{ass}) + N_{Fe,eq}, \quad (1)$$

$$\tau_{ass} = 5.7 \times 10^5 \frac{\text{s}}{\text{Kcm}^{-3}} \times \frac{T}{N_A} \exp\left(\frac{E_m}{kT}\right), \quad (2)$$

where  $\tau_{ass}$  is the characteristic of the complex association,  $N_{Fe_i,0}$  is the concentration of interstitial iron atoms formed due to FeB pair dissociation,  $N_{Fe_i,0} = N_{Fe_i}(t=0) = N_{Fe}$ , and  $N_{Fe,eq}$  is the portion of interstitial iron atoms that remain unpaired in the equilibrium state  $N_{Fe,eq} = N_{Fe_i}(t \rightarrow \infty)$ . According to [?, ?]

$$[Fe_{i,eq}] = \frac{Fe_{bulk}}{[1 + N_A 10^{-23} \exp(\frac{E_b}{kT})][1 + \exp(\frac{E_F - 0.39 \text{ eV}}{kT})]}, \quad (3)$$

where  $E_b$  is the binding energy of the FeB pairs and  $F$  is the Fermi level (donor level  $E_{Fe_i} = E_V + 0.394 \text{ eV}$  is associated with  $Fe_i$ ).

In turn, the study estimated the iron–boron pair concentration  $N_{\text{FeB}}$  from

$$N_{\text{FeB}}(t) + N_{\text{Fe}}(t) = N_{\text{Fe},0}, \quad (4)$$

Overall, the concentrations of iron-related defects depended not only on time but also on their spatial position within the structure, reflecting the non-uniformity of the  $F-E_{\text{Fei}}$  difference. Olikh et al. [?, ?] provide a detailed description of the modelling approach, including the values of silicon and defect parameters. To create the training dataset, we used 25  $N_{\text{Fe}}$  values, uniformly distributed on a logarithmic scale from  $10^{10} \text{ cm}^{-3}$  to  $10^{14} \text{ cm}^{-3}$ . Figure ?? shows examples of the resulting  $I_{\text{sc}}(t)$  dependencies and the corresponding wavelet spectrograms. We generated the simulated test dataset using 10  $N_{\text{Fe}}$  values not included in the training set.

### 2.3. Experiment details

We validated the proposed method using measurements of  $n^+p$ - $p^+$  silicon solar cells fabricated on 380  $\mu\text{m}$  thick  $p$ -type boron-doped Cz-wafers ( $N_{\text{B}} = 1.36 \times 10^{15} \text{ cm}^{-3}$ ). The  $n^+$  (0.7  $\mu\text{m}$ , 20–30  $\Omega/\square$ ) and  $p^+$  (0.6  $\mu\text{m}$ , 10–20  $\Omega/\square$ ) layers were formed by phosphorus and boron diffusion, respectively; details of the fabrication are given elsewhere [?]. Current–voltage curves and  $I_{\text{sc}}$  kinetics were measured with a Keithley 2450 source meter. A 940 nm SN-HPIR940nm-1W LED (intensity 5  $\text{W}/\text{m}^2$ ) served as a low-intensity monochromatic source, with its output stabilized by a W1209 thermostat and a feedback-controlled power supply. The cell temperature was regulated by a thermoelectric cooler with an STS-21 sensor and maintained constant by a PID algorithm in the control software. FeB-pair dissociation was induced by intense (about 7000  $\text{W}/\text{m}^2$ ) halogen-lamp illumination. The illumination intervals were selected according to a previous study [?]. We measured the kinetics of the short-circuit current in the dark at 340 K over a period of 3000 s. According to Eq. (??), this interval is sufficient for the pairs to recover fully to their equilibrium concentration.

Figure ?? presents an example of a measured  $I_{\text{sc}}(t)$  dependence. The signal contains some noise because, despite using a thermostat, the LED temperature fluctuated by approximately 0.4 K. We smoothed the data using a Savitzky–Golay filter, adaptively selecting the window length and filter order according to Krishnan and Seelamantula [?]. For the wavelet transform, we used only the current values at the same time points as in the simulations. Figure ?? shows the corresponding smoothed curve, while the other panels display the spectrograms derived from the raw experimental curve and the processed dependence.

The iron concentration  $N_{\text{Fe}}$  was determined using a methodology described in [?, ?], which is based on fitting the kinetics of the short circuit current following FeB pairs dissociation. We analysed 28 samples with iron concentrations ranging from  $10^{11} \text{ cm}^{-3}$  to  $2 \times 10^{13} \text{ cm}^{-3}$ . To test models trained on simulated data, we used all experimental measurements. For models trained on experimental results, we randomly selected 28 samples for training and assigned the remaining 8 samples to the test set.

### 2.4. Computer vision models

To extract graphical features from the wavelet spectrograms, we applied several computer vision models available in Keras, namely EfficientNetB7, ResNet152V2, MobileNetV2, Xception, and NASNet-Large. All these models belong to the class of convolutional neural networks (CNNs) designed for object classification, and researchers have successfully applied them to process EL images of solar cells [?, ?, ?, ?, ?]. For each model, we employed two feature extraction approaches: we either passed the class probabilities, indicating the likelihood of the image belonging to a specific class, to the next stage of the pipeline, or we used the features directly extracted by the CNN.

We also applied the CSPDarknet53 model, which serves as the CNN backbone for YOLOv4. This family of models features a more complex CNN architecture, optimised for detecting multiple objects in an image rather than single-object classification. Researchers widely use these models in Imaging-Based Techniques [?, ?, ?]. CSPDarknet53 returns three feature layers, and we processed either only the highest-level layer or the last two layers, depending on the analysis.

Increasing the feature dimension does not constantly improve the total information variance. To reduce the impact of redundant data, we applied PCA, which constructs new, uncorrelated features (principal components). PCA is a widely used and effective tool in machine learning, particularly for improving results in Electrical Testing Techniques [?, ?]. We applied PCA to the training datasets using an explained variance threshold of 99.9%, retaining only the principal components that together account for at least 99.9% of the total variance of the original features. This procedure significantly reduced the feature dimensionality. We applied PCA only to the computer vision models that performed best on the test datasets without PCA, in order to assess the usefulness of this preprocessing step. Because YOLOv4 extracts a vast number of features, we also applied an additional dimensionality reduction technique. Specifically, we performed global average pooling on each convolutional feature map, replacing the spatial map with its mean value and thus producing one scalar per channel. We summarise the

computer vision model variants in table 1, along with the notations used in the subsequent analysis.

### 2.5. Regression algorithms

We used five ML algorithms to develop regression models for predicting iron concentration: Random Forest (RF), Gradient Boosting (GB), eXtreme Gradient Boosting (XGB), Support Vector Regression (SVR) and Deep Neural Network (DNN). The models were implemented using Python toolkits: Keras for DNN, Scikit-learn for RF, GB, and SVR, and XGBoost for XGB. We applied each regression model to the features extracted from the variants listed in table 1. The only exception was the uncompressed YOLOv4 features: our available computational resources (2.9 GHz AMD Ryzen 7 4800H CPU, 8 GB RAM, GeForce GTX 1650 4GB) allowed us to use only SVR for these features. We set the target variable to  $\log N_{\text{Fe}}$ , which is a standard transformation for achieving high prediction accuracy when quantities vary over a wide range [?, ?]. We normalised both input and target variables to have a mean of zero and a standard deviation of one over the training set.

For each case, we tuned the regression models to optimize performance. We used the Optuna toolkit to tune model hyperparameter, employing the TPE sampler and Hyperband pruner for efficient hyperparameter selection. Tables S1–S5 (Supplementary Material) provide the tuned hyperparameters and their search ranges. We applied 5-fold cross-validation during model tuning, using 20% of the training data as a validation set to evaluate models trained on the remaining 80%. The chosen hyperparameter combinations are presented in Tables S6–S10.

In total, we considered 87 combinations of computer vision model and regression model, training and testing each on both simulated and experimental data. To label the results, we combined the label from the last column of table 1 with the abbreviated name of the regression algorithm.

### 2.6. Model evaluation

We conducted a rigorous assessment of model performance using multiple evaluation metrics to construct a robust regression model. For iron quantification, we used the mean squared error (MSE), mean absolute percentage error (MAPE), median absolute percentage error (MedAPE), and the coefficient of determination ( $R^2$ ), as defined in Eqs. (5)–(8)

$$MSE = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2, \quad (5)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|N_{\text{Fe,PRED},i} - N_{\text{Fe,TRUE},i}|}{N_{\text{Fe,TRUE},i}} \times 100\% \quad (6)$$

$$MedAPE = \text{median} \sum_{i=1}^N \frac{|N_{\text{Fe,PRED},i} - N_{\text{Fe,TRUE},i}|}{N_{\text{Fe,TRUE},i}} \times 100\% \quad (7)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (N_{\text{Fe,TRUE},i} - N_{\text{Fe,PRED},i})^2}{\sum_{i=1}^N (N_{\text{Fe,TRUE},i} - \bar{N}_{\text{Fe,TRUE}})^2}, \quad (8)$$

where  $\hat{y}_i$  represents the predicted value for the  $i$ -th data point,  $y_i$  is the known value for the  $i$ -th data point, and  $N$  is the number of samples in the dataset ( $N = 25$  for the simulated training set, 20 for the experimental training set, 10 for the simulated test set, and 28 and 8 for the experimental datasets used to test models trained on the simulated and experimental sets, respectively);  $N_{\text{Fe,PRED},i}$  is the predicted iron concentration,  $N_{\text{Fe,TRUE},i}$  is the known value (either the parameter used in the simulation or obtained from experimental iron determination);  $N_{\text{Fe,TRUE}}$  is the mean of the true values in the dataset.

We selected MSE as a primary metric for assessing model accuracy and trained the models specifically to minimise it. However, because computing  $y_i$  involves normalisation and logarithmic transformation of  $N_{\text{Fe}}$ , MSE does not fully reflect the accuracy of iron-contamination estimation. To complement it, we used MAPE, which measures the mean relative error. We also considered MedAPE, which indicates the error threshold below which 50% of the predictions fall. MedAPE is more robust to individual outliers, which can strongly influence the mean error in small datasets. Finally, we employed  $R^2$  to quantify the fraction of variance in the target variable captured by the model, indicating how well the predicted values reproduce the observed data; a value of 1 corresponds to perfect agreement.

## 3. Results and discussion

### 4. Conclusion

### References

**Table 1.** Summary of used pretrained CNN models and feature extraction variants

Base model	Model type	Feature processing	Output dimension	Model Label
EfficientNetB7	Classifier	None	1000	ENB7:CL
	Feature extractor	None	2560	ENB7:FE
		PCA	39	ENB7:FE:P
MobileNetV2	Classifier	None	1000	MNV2:CL
	Feature extractor	None	1280	MNV2:FE
		PCA	124	MNV2:FE:P
NASNetLarge	Classifier	None	1000	NAS:CL
		PCA	30	NAS:CL:P
	Feature extractor	None	4032	NAS:FE
ResNet152V2	Classifier	None	1000	R152:CL
	Feature extractor	None	2048	R152:FE
Xception	Classifier	None	1000	XCP:CL
	Feature extractor	None	2048	XCP:FE
YOLOv4 (CSPDarknet53)	Feature extractor (raw, top layer)	None	86528	YL:FE1
		PCA	137	YL:FE1:P
CSPDarknet53 (YOLO backbone)	Feature extractor (raw, top & penultimate layers)	None	433640	YL:FE2
		PCA	142	YL:FE2:P
	Feature extractor (pooled, top layer)	None	512	YL:FP1
	Feature extractor (pooled, top & penultimate layers)	None	1024	YL:FP2