Dear Editor and Reviewers,

We sincerely thank you for taking the time to review our manuscript "Computer vision-based method for quantifying iron-related defects in silicon solar cells" (Ref. No.: SST–111269). Your insightful comments and constructive suggestions have greatly helped us improve the quality of our work. We particularly appreciate your careful reading and thoughtful feedback, which have led to significant improvements in both the technical content and presentation clarity of our manuscript. We have carefully addressed all the comments and made corresponding revisions to the manuscript. The location of revisions is pointed by red color and highlighted in yellow in "CompleteDocumentForReview.pdf". Below we provide our detailed point-by-point responses to each comment. We hope the revised manuscript better meets your expectations and standards for publication in Solar Energy.

Response to Reviewer #1

Major Comment 1. *Novelty and Contribution: The idea of applying pre-trained CV models to wavelet-transformed kinetic data is interesting and potentially generalizable. However, similar signal-to-image ML transformations exist in other domains. The manuscript should emphasize what new physical or methodological insight this work provides beyond prior Fourier/wavelet-based ML approaches..*

Reply:

The additional information was added to the revised manuscript (from page 7, left column, paragraphs 2 to page 8, left column, paragraphs 2; page 8, right column, paragraph 2).

Major Comment 2. *Dataset Size and Overfitting Risk: The study relies on extremely small datasets (25 simulated and 28 experimental samples). Although data augmentation is performed, flipping or rotating spectrograms likely introduces redundant samples rather than independent data points. The near-perfect R^2 values (0.996–0.999) strongly suggest overfitting. A robust cross-validation (e.g., k-fold or leave-one-out) with uncertainty quantification is needed.*

Reply:

The minimal level shift, which was used as evidence of defect transformation, was 40 meV.

Major Comment 3. *Simulation–Experiment Gap: A major discrepancy is observed between simulated and experimental predictions, requiring a post-hoc quadratic correction. This implies that the CNN–regressor models primarily learn the synthetic data distribution rather than physical correlations. The authors should explore physics-based domain adaptation or partial fine-tuning using experimental data instead of empirical correction.*

Reply:

The description of observation technique was modified (page 3, left column, paragraph 3)

Major Comment 4. *Physical Model Validation: The SCAPS-1D simulations use fixed FeB parameters (binding energy, migration energy, and pre-exponential factors). Yet, these parameters vary widely in literature (0.55–0.69 eV for migration energy). Without sensitivity analysis or error quantification, the generated synthetic dataset may not reflect realistic kinetics. Validation against first-principles or experimental benchmarks would strengthen the study.*

Reply:

Allow us to say a few words in favor of GaAs and 6H-SiC.

Major Comment 5. *Regression and Feature Interpretation: Although multiple regressors (SVR, XGB, DNN, RF, GB) are compared, no insight is given into the learned features or their physical correlation with iron concentration. Incorporating explainable AI techniques (e.g., SHAP, PCA loading analysis) would add interpretability to what the CNN features represent.*

Reply:

Allow us to say a few words in favor of GaAs and 6H-SiC.

Major Comment 6. *Post-Hoc Correction: The quadratic correction (Eq. 10) is an empirical adjustment that artificially improves metrics but lacks theoretical justification. The authors should either (i) replace it with a physics-informed calibration (e.g., temperature or diffusion based scaling) or (ii) clearly acknowledge its heuristic nature and limitations.*

Reply:

Allow us to say a few words in favor of GaAs and 6H-SiC.

Major Comment 7. *Statistical Reporting: The reported MSE, MAPE, and R^2 values are given without variance or confidence intervals. Given the small datasets, reporting mean ± standard deviation across multiple random splits would be essential to establish statistical robustness.*

Reply:

Allow us to say a few words in favor of GaAs and 6H-SiC.

Minor Comment 1. *The introduction is overly broad; it should focus more on ML for microscopic defects rather than general PV or macro-defect analysis.*

Reply:

The body of work specifically focused on applying ML to microscopic defect analysis remains relatively limited. However, we conducted a detailed review of the publications cited in the initial manuscript draft and carried out an extensive supplementary literature search. As a result, it has been established that existing applications of ML in microscopic defect characterization can be broadly categorized into several distinct approaches. One such approach focuses on enhancing conventional defect-analysis techniques through the integration of Artificial Intelligence methods for processing and interpreting the resulting experimental signals.

For example, Buratti *et al.* [1] employed regression algorithms, including Random Forest (RF), Gradient Boosting (GB), and Deep Neural Networks (DNN), to analyze dependencies derived from temperature- and injection-dependent lifetime spectroscopy (TIDLS). They trained the models on more than one million simulated curves, which enabled accurate estimation of silicon defect energy levels and carrier capture cross-sections. In addition, unlike the conventional fitting of signals with the Shockley–Read–Hall equation, their approach can also predict the energy level position at half of the bandgap. An extension of this approach was presented in [2], where the methodology incorporated a Convolutional Neural Network (CNN) to analyse images derived from a family of lifetime curves measured at different temperatures, in addition to applying a RF model to the standard TIDLS signal. In that study, the CNN was used both to perform the classification of the half-bandgap position of the energy level and to extract features, which were subsequently used by the RF. As in the earlier work, the models were trained on a dataset consisting of several hundred thousand synthetic samples. An alternative TIDLS signal processing strategy was also investigated by Wang *et al.* [3], who used CNNs to analyse one-dimensional signals and thereby extract parameters associated with two-energy-level defects in silicon. Machine-learning methods are also employed for the analysis of Raman spectra [4]. In this study, the spectra of electron-irradiated GaAs were examined using linear discriminant analysis models. These models were trained on 6,000 experimentally acquired spectra, enabling the identification of radiation-induced defects.

An alternative approach is based on determining defect parameters by analysing the characteristics of devices, primarily solar cells, that are directly affected by such defects. For silicon solar cells, for example, a methodology has been proposed to estimate the concentration of contaminant impurities from the magnitude of the ideality factor obtained from current–voltage (I–V) characteristics [5] or from variations in photovoltaic conversion parameters [6]. In both studies, classical regression algorithms (DNN, RF, Support Vector Regression (SVR), and GB) were employed. The numerical values of parameters extracted from the I–V characteristics served as input features, and the models were trained on tens of thousands of current–voltage curves simulated under different defect parameters. A closely related approach was presented by Haidari *et al.* [7], who used thirteen parameters extracted from the I–V curves of CIGS solar cells as inputs to a DNN to predict the spatial distribution and concentration of six bulk and surface defects. The inverse problem, namely the determination of photovoltaic conversion parameters based on predefined defect concentrations, was examined by *Kim et al.* [8] for perovskite solar cells. In that study, RF, XGBoost, Linear Regression, and Multilayer Perceptron algorithms were evaluated, and the RF model delivered the highest performance. In all four studies mentioned above, the SCAPS-1D simulation tool was consistently used to generate the synthetic I–V characteristics that formed the training datasets.

Beyond the detection and characterization of defects in actual devices, a distinct research direction focuses on accelerating and improving density functional theory (DFT) and molecular dynamics (MD) calculations of defect parameters. For example, several studies have demonstrated the use of Graph Neural Networks, trained on DFT-calculated data, to estimate vacancy formation energies [9, 10] and to evaluate the electronic structure of charged defects in GaAs [11]. Graph Convolutional Networks have also been applied to MD-generated datasets for predicting vacancy diffusion paths in high-entropy alloys [12] and intrinsic defects in perovskites [13]. In addition, DFT datasets have been integrated with ML methods to identify formation enthalpies

and ionization energies of impurity defects [14] and to determine the equilibrium configurations of defects in emerging materials [15].

The relevant information has been incorporated into the Introduction section (pages 1–2, paragraphs three through six).

Minor Comment 2. *Some typographical errors exist (e.g., "where where" in Eq. 1). Please proofread carefully.*

Reply: We carefully reviewed the manuscript and corrected multiple typographical errors. Namely:

- changed phrase "Standard approaches to solving such problems involve the use of Fourier or wavelet transforms, and last were applied in this study" to "Standard approaches to solving such problems involve the use of Fourier or wavelet transforms, and the latter were applied in this study" on page 2;
- changed "where where" to "where" after Eq. (1);
- changed "A is the constant" to "A is the pre-exponential constant" after Eq. (3);
- changed phase "Panels (b) and (c) show the wavelet spectrograms corresponding to the curves with filled squares and open circles, respectively" to "Panels (b) and (c) show the wavelet spectrograms corresponding to the curves with open squares and filled circles, respectively" in the caption of Fig 3;
- a sentence was added to clarify the applicability of the formula for the median absolute percentage error "Eq. (8) implies that $\texttt{MAPE}_i$ must be arranged in order of magnitude; " after Eq. (8)

etc.

We hope that these revisions have resolved the issue as thoroughly as possible.

Minor Comment 3. *Figures should include axis units, consistent color scales, and indicate whether values are in linear or logarithmic scale.*

Reply:

We thank the Reviewer for identifying the shortcomings in the figures. Axes, units of measurement, and color scales have been added to Figures 2b, 2c, 3b, and 3c. The units of measurement (%) corresponding to the MAPE and MedAPE values have also been included next to the color scale labels, as shown in Figures 5, 6, 7, 9, and 10. The mutual arrangement of the ticks and tick labels in all figures enables unambiguous identification of the scale type.

Minor Comment 4. *The Supplementary Figures (S1–S10) are repeatedly referenced but insufficiently summarized in the main text. A concise overview table would be helpful.*

Reply:

Allow us to say a few words in favor of GaAs and 6H-SiC.

Minor Comment 5. *The data availability statement ("upon reasonable request") should be replaced with a public repository link for transparency.*

Reply:
Allow us to say a few words in favor of GaAs and 6H-SiC.

Minor Comment 6. *A comparison with simpler ML baselines (e.g., direct regression on ISC(t) data without wavelet transformation) would contextualize the improvement due to CV-based transfer learning.*
Reply:
Allow us to say a few words in favor of GaAs and 6H-SiC.

Minor Comment 7. *References [6], [35], [38] should be verified for year and page accuracy. Some reference formatting inconsistencies (journal abbreviations, italics) should be corrected.*
Reply:
Allow us to say a few words in favor of GaAs and 6H-SiC.

Response to Reviewer #2

Comment 1. *It is additionally essential to examine temperatures between 270 to 350 kelvin.*
Reply:
To our shame, the reviewer is correct about some fog in Results and discussion. We hopefully rephrased the section to add sunlight.

Comment 2. *You should compare and review your manuscript with other new articles such as "Novel Design of Multi-Layer Cubic Nanoparticles for Achieving Efficient Thin-Film Perovskite Solar Cells"*
Reply: We have expanded the Introduction section, as detailed in our responses to Minor Comment 1 and Major Comment 1 from Reviewer 1. In particular, the article mentioned in Reviewer 2's comment is now cited as Reference [38] in the revised manuscript.

Comment 3. *Put the solar cell parameters in a table with references.*
Reply: We added Table 1 to the revised manuscript (page 5). This table summarizes the parameters of the solar cell, silicon, and defect states at the temperature used in the main calculations.

Comment 4. *Actually, all solar cells have Rs and Rsh values. By investigating parasitic losses on cell performance, the article could be made more interesting.*
Reply: We fully agree with the Reviewer that, for real solar cells, it is essential to consider the influence of series and shunt resistances on the efficiency of photovoltaic conversion. In particular, the presence of these resistances reduces the short-circuit

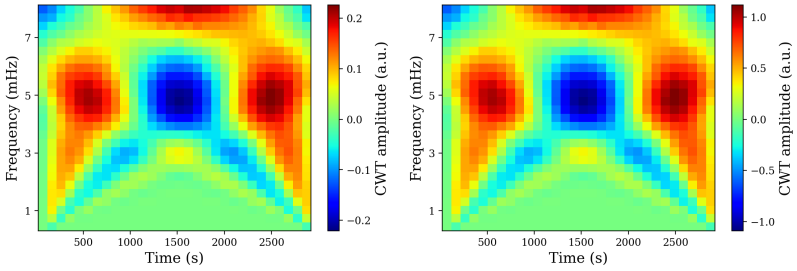

Figure 1. Wavelet spectrograms corresponding to the simulated short-circuit curves for a solar cell with $N_{\mathrm{Fe}} = 10^{13}$ cm^{-3}, obtained before (left) and after (right) normalization.

current. In an ideal case, the short-circuit current I_{sc} is equal to the photogenerated current I_{ph}:

$$I_{\mathrm{sc}} = I_{\mathrm{ph}} \, .$$

In the presence of resistances, within the single-diode model approximation

$$I_{\mathrm{sc}} \approx I_{\mathrm{ph}} \frac{R_{\mathrm{sh}}}{R_{\mathrm{sh}} + R_{\mathrm{s}}} \, .$$

However, when FeB pairs are restored, the values of R_{sh} and R_{s} remain unchanged. Therefore, the presence of parasitic resistances will effectively scale the dependence $I_{\mathrm{sc}}(t)$ by a constant factor only. In this case, the resulting wavelet spectrogram remains unaffected. As an example, Figure 1 presents the spectrograms obtained for the kinetic dependencies of the short-circuit current before and after normalization to the initial current value. It can be seen that normalization affects the amplitude of the continuous wavelet transform, but this change is proportional across all frequency and time values. As a result, the graphical appearance of the spectrogram remains unchanged, and computer vision models extract the same features regardless of normalization (and, consequently, regardless of the presence of parasitic resistance).

Thus, the method proposed in this work for estimating the concentration of impurity iron is inherently resistant to the presence of series and shunt resistances. The corresponding information has been added to the manuscript in the final paragraph preceding the Conclusion section. We are grateful to the Reviewer for highlighting this additional advantage of the proposed approach.

A direct investigation of the influence of parasitic losses on cell performance was not the objective of this study.

Response to EDITOR REPORT

REPORT. *We have found that your manuscript contains text which appears to have been replicated from the following published articles:*

www.sciencedirect.com/science/article/abs/pii/S0038092X25005171?via%3Dihub

Please reduce the level of overlap in your revised manuscript by rewriting the appropriate sections.

Reply:

First, we apologize for the observed similarities. The cited article is our own and also addresses the determination of iron concentration in silicon solar cells. However,

the approaches used in the cited work and in the present study are fundamentally different. In the former, regression models are employed that utilize changes in photoelectric parameters during the decay of FeB pairs, whereas in the present manuscript, the primary approach involves converting the kinetic dependencies of the short-circuit current into images and extracting features using computer vision models. Nevertheless, both studies use similar solar cell models, apply standard regression algorithms and metrics, and perform testing on experimental samples from the same batch. This explains certain similarities in wording.

We have revised the relevant sections and reduced the extent of overlap.

[16]

References

[1] Buratti Y, Le Gia Q T, Dick J, Zhu Y and Hameiri Z 2020 *npj Computational Materials* **6** 142
[2] Buratti Y, Dick J, Le Gia Q and Hameiri Z 2022 *ACS Appl. Mater. Interfaces* **14** 48647–48657
[3] Wang S, Wright B, Zhu Y, Buratti Y and Hameiri Z 2024 *Sol. Energ. Mat. Sol.* **277** 113123
[4] Chia J Y, Thamrongsiripak N, Thongphanit S and Nuntawong N 2024 *J. Appl. Phys.* **135** 025701
[5] Olikh O, Lozitsky O and Zavhorodnii O 2022 *Prog. Photovoltaics Res. Appl.* **30** 648–660
[6] Olikh O and Zavhorodnii O 2025 *Sol. Energy* **300** 113754
[7] Haidari G 2025 *Appl. Phys. A* **131** 850
[8] Kim S, Jeong Y, Han D W and Mo C B 2023 *J. Electron. Mater.* **52** 5861–5871
[9] Choudhary K and Sumpter B G 2023 *AIP Adv.* **13** 095109
[10] Kumagai Y 2025 *Materia Japan* **64** 518–523
[11] Ma Y, Yu H, Zhong Y, Chen S, Gong X and Xiang H 2025 *Appl. Phys. Lett.* **126** 044103
[12] Reimer C, Saidi P, Casert C, Beeler C, Tetsassi Feugmo C G, Whitelam S, Mansouri E, Martinez A, Beland L and Tamblyn I 2025 *J. Appl. Phys.* **138** 074306
[13] Tyagi V, Pols M, Brocks G and Tao S 2025 *The Journal of Physical Chemistry Letters* **16** 5153–5159 pMID: 40371867 (*Preprint* https://doi.org/10.1021/acs.jpclett.5c01139) URL https://doi.org/10.1021/acs.jpclett.5c01139
[14] Mannodi-Kanakkithodi A, Xiang X, Jacoby L, Biegaj R, Dunham S T, Gamelin D R and Chan M K 2022 *Patterns* **3** 100450 ISSN 2666-3899 URL https://www.sciencedirect.com/science/article/pii/S266638992200023X
[15] Mosquera-Lois I, Kavanagh S R, Ganose A M and Walsh A 2024 *npj Comput. Mater.* **10** 121
[16] Wijaranakula W 1993 *J. Etectrochem. Soc.* **140** 275–281