Defect content characterization in solar cells with the assistance of machine learning

Oleg Olikh, Oleksii Zavhorodnii

*Taras Shevchenko National University of Kyiv, 64/13, Volodymyrska Street, 01601, Kyiv, Ukraine*

olegolikh@knu.ua

Integrating artificial intelligence into sustainable clean energy research, particularly using machine learning (ML) for defect characterization, is increasingly compelling. A prevalent strategy is to identify extended defects from electroluminescence images. Recently, however, there has been a notable shift in focus towards point defects [1,2]. Our research aims to create an efficient, low-cost machine learning-based methodology for evaluating recombination defects in solar cells (SCs) using current-voltage (IV) measurement data, eliminating the need for extra equipment.

To demonstrate our methodology, we focused on identifying iron-related defect concentrations in silicon SC. Fig. 1(a) presents the workflow. Using SCAPS-1D software, the performance of back surface field SCs under both standard AM1.5 and monochromatic (940 nm) illumination were modeled. Simulated IV curves captured the behaviour of iron-acceptor pairs and scenarios with only interstitial iron. Then the relative changes in short-circuit current ε*Isc*, open-circuit voltage ε*Voc*, efficiency εη, and fill factor ε*FF* were extracted. ML techniques – deep neural networks (DNN), random forest (RF), and gradient boosting (GB) – were employed to estimate iron concentrations. The accuracy of predictions from various models was compared using data obtained under different lighting conditions and with varying numbers (ranging from 4 to 7) of descriptors. In the simplest case, the descriptors included the SC's base depth and doping level, temperature, and ε*Isc*. For cases involving 5, 6, and 7 descriptors, the εη, ε*Voc*, and εFF, were added respectively. The results are shown in Tables 1 and 2 and Fig. 1(b).

[1] O. Olikh, O. Lozitsky and O. Zavhorodnii, *Prog Photovolt Res Appl.* **2022**, *30*, 648.

[2] Y. Buratti, J. Dick, Q.L. Gia and Z. Hameiri, *ACS Appl. Mater. Interfaces* **2022**, *14*, 48647.

Залежності частки зразків, для яких похибка не перевищує певну величину від величини похибки передбачення

|  |  |
| --- | --- |
|  |  |
|  |  |

**Fig. 1.** (a) Workflow. (b) The relationship between the fraction of samples for which the error does not exceed a certain threshold and the prediction error magnitude.

**Table 1.** Results of 5-fold cross-validation for train dataset

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Illumination | Mean squared error (10-3) | | | |
| Number of descriptors | | | |
| 4 | 5 | 6 | 7 |
| DNN | AM1.5 | 42±5 | 9±3 | 4±2 | 2±1 |
| 940 nm | 10±5 | 6.1±0.4 | 6±2 | 1.5±0.7 |
| RF | AM1.5 | 33±2 | 11±3 | 5±2 | 4±1 |
| 940 nm | 6±1 | 4.6±0.2 | 3.0±0.5 | 3.0±0.8 |
| GB | AM1.5 | 34±2 | 9±2 | 5±2 | 4±1 |
| 940 nm | 4.2±0.6 | 3.5±0.2 | 2.3±0.6 | 2.1±0.5 |

**Table 2.** Prediction accuracy for test dataset

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Number of descriptors | Mean squared error (10-3) | | Mean relative error (%) | | R2 | |
| Illumination | | | | | |
| AM1.5 | 940 nm | AM1.5 | 940 nm | AM1.5 | 940 nm |
| DNN | 4 | 58 | 6 | 53 | 10 | 0.905 | 0.977 |
| 5 | 4 | 33 | 7 | 36 | 0.988 | 0.881 |
| 6 | 0.9 | 0.6 | 5 | 5 | 0.992 | 0.993 |
| 7 | 5 | 0.8 | 11 | 5 | 0.990 | 0.988 |
| RF | 4 | 41 | 3 | 142 | 11 | 0.930 | 0.968 |
| 5 | 10 | 3 | 15 | 10 | 0.959 | 0.967 |
| 6 | 4 | 3 | 10 | 9 | 0.972 | 0.956 |
| 7 | 5 | 3 | 11 | 10 | 0.958 | 0.963 |
| GB | 4 | 33 | 3 | 43 | 8 | 0.947 | 0.965 |
| 5 | 9 | 2 | 13 | 8 | 0.955 | 0.980 |
| 6 | 5 | 2 | 10 | 7 | 0.969 | 0.967 |
| 7 | 5 | 2 | 10 | 8 | 0.960 | 0.961 |