

# 3D CNN and grad-CAM based visualization for predicting generation of dislocation clusters in multicrystalline silicon

Cite as: APL Mach. Learn. 1, 036106 (2023); doi: 10.1063/5.0156044

Submitted: 25 April 2023 • Accepted: 12 July 2023 •

Published Online: 24 July 2023



View Online



Export Citation



CrossMark

Kyoka Hara,<sup>1</sup> Takuto Kojima,<sup>2</sup> Kentaro Kutsukake,<sup>3</sup> Hiroaki Kudo,<sup>2</sup> and Noritaka Usami<sup>1,a</sup>

## AFFILIATIONS

<sup>1</sup> Graduate School of Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8603, Japan

<sup>2</sup> Graduate School of Informatics, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan

<sup>3</sup> Center for Advanced Intelligence Project, RIKEN, 1-4-1, Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan

<sup>a</sup>Author to whom correspondence should be addressed: [usa@material.nagoya-u.ac.jp](mailto:usa@material.nagoya-u.ac.jp)

## ABSTRACT

We propose a machine learning-based technique to address the crystallographic characteristics responsible for the generation of crystal defects. A convolutional neural network was trained with pairs of optical images that display the characteristics of the crystal and photoluminescence images that show the distributions of crystal defects. The model was trained to predict the existence of crystal defects at the center pixel of the given image from its optical features. Prediction accuracy and separability were enhanced by feeding three-dimensional data and data augmentation. The prediction was successful with a high area under the curve of over 0.9 in a receiver operating characteristic curve. Likelihood maps showing the distributions of the predicted defects are in good resemblance with the correct distributions. Using the trained model, we visualized the most important regions to the predicted class by gradient-based class activation mapping. The extracted regions were found to contain mostly particular grains where the grain boundaries changed greatly due to crystal growth and clusters of small grains. This technique is beneficial in providing a rapid and statistical analysis of various crystal characteristics because the features of optical images are often complex and difficult to interpret. The interpretations can help us understand the physics of crystal growth and the effects of crystallographic characteristics on the generation of detrimental defects. We believe that this technique will contribute to the development of a better fabrication process for high-performance multicrystalline materials.

© 2023 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0156044>

06 July 2023 15:20:39

## I. INTRODUCTION

In multicrystalline materials, the formation of crystal defects or imperfections in the atomic arrangement has a profound effect on the material behavior. A variety of crystal defects can be found in multicrystalline materials: point defects, such as vacancies and impurities; line defects, including edge and screw dislocations; and surface defects, which divide the solid crystal into regions with different structures and crystallographic orientations. The presence of these various crystal defects is found to change the mechanical, physical, and optical properties of crystalline materials.<sup>1–8</sup> The mechanical strength strongly depends on the density of grain boundaries, the thermal conductivity in metals as well as the absorption of light around certain wavelengths is reduced by point defects, and plastic

deformation is controlled by dislocations.<sup>4–8</sup> While the increase in mechanical strength has allowed for the production of various applications, the deterioration in electrical and thermal properties has raised concerns. Hence, the study of understanding and controlling the defects of crystalline materials is attracting attention for fabricating materials with suitable properties. While many studies have been conducted to control defect generation,<sup>9,10</sup> they are often operated on a series of physical experiments and human-based analysis on a limited area. If we could analyze the crystallographic characteristics holistically with the aid of machine learning, it would be possible to eliminate human-based bias or preconceptions and conduct a comprehensive analysis.

Multicrystalline silicon is employed as a model material in this study. Multicrystalline silicon is a common material used to

fabricate solar cells due to its abundance and affordability. However, the electrical properties of the cells, particularly their conversion efficiencies, tend to be limited by dislocation clusters with high concentrations of dislocations up to  $10^7 \text{ cm}^{-2}$  acting as recombination centers for photogenerated carriers.<sup>11–14</sup> Thus, it is essential to understand the generation mechanism of dislocation clusters and to control them. Typically, dislocation clusters are formed near grain boundaries during silicon ingot casting due to the relaxation of strain energy.<sup>15</sup> While the generation of dislocations has been studied by means of observation of silicon ingots and their stress distributions during growth processes, the relationships between the dislocations and other crystalline characteristics, including grain boundary arrangements, grain size distributions, and crystallographic orientations, have not been clarified.<sup>16,17</sup>

In this study, we construct a machine learning model that reproduces the behaviors of the crystal and analyze the internal structures to gain an understanding of the physical phenomenon behind the behavior. We address the various characteristics in order to gain a deep and comprehensive understanding of the generation of dislocation clusters. A machine learning model consisting of convolutional and dense layers is trained with optical images containing information about the crystalline characteristics and photoluminescence (PL) images displaying the distributions of dislocation clusters. The construction of the model and its outcomes are reported, and we provide insights from the analysis of the model.

## II. METHODS

Multicrystalline silicon wafers were used in this study. In order to associate the crystalline structures with the generation of dislocation clusters, a machine learning model conducted binary classification that predicts the presence of dislocation clusters from the given images of the wafers. The pre-trained model was then analyzed to identify the relevant structures in the generation of dislocation clusters.

### A. Collecting optical and photoluminescence images

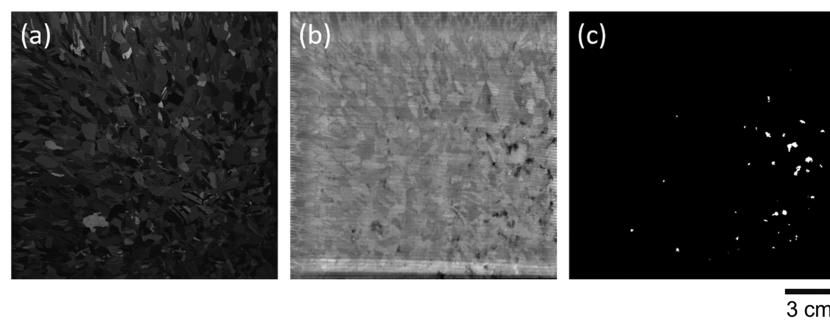
The model was trained using two types of images: optical and PL images. The images were taken from silicon wafers cut out from a single ingot by a diamond wire saw and textured using alkaline solutions. Texturing removes contamination and damage caused by

the sawing process and creates pyramid-like structures consisting of {111} planes. The wafer size was  $15.6 \times 15.6 \text{ cm}^2$ , and the thickness of each wafer was  $\sim 180 \mu\text{m}$ , cut out with an interval of  $290 \mu\text{m}$ . Optical images were taken using a house-made apparatus.<sup>18,19</sup> The apparatus mainly consists of an imaging photometer (ProMetric IP-PMY29), a rotating collimated light for illumination, and a stage. Irradiation on the wafers generates reflection patterns specific to the crystallographic orientation, which we integrate into full-sized optical images of  $500 \times 500 \text{ pixels}^2$ . An example of the optical images is shown in Fig. 1(a), in which the grain boundaries, grain sizes and shapes, and crystallographic orientations (represented by the color differences) are distinguishable.

For the same wafers, PL images were taken. The images were resized into images of  $500 \times 500 \text{ pixels}^2$ , with each pixel of the optical and PL images corresponding to each other. The PL images were then prepared for training by binarization. The images were binarized with a fixed threshold and reconstructed into an image shown in Fig. 1(c), with the white areas representing dislocations and the black areas representing non-dislocations. The spatial resolution is limited by the pixel size, which is about  $300 \mu\text{m}$ . Empirically, there is a marked decrease in PL intensity at dislocation densities of  $10^5 \text{ cm}^{-2}$ , where 10 or more dislocations exist per pixel. Since it is difficult to estimate the dislocation density from the PL intensity with high accuracy, only the presence or absence of significant dislocation generation in each pixel is treated by thresholding.

### B. Model design

Our proposed model is a convolutional neural network (CNN)<sup>20</sup> consisting of three convolutional layers and one dense layer. We have chosen CNN since it provides translational invariance and the ability to extract a variety of features directly from the image without prior parameter selection. The convolutional layers extract the key features from the input optical images. Then, based on the binarized PL images, the dense layers classify the given images into outputs between 1 and 0, which represents whether the center pixels of the given images contain dislocations or not. The hyperparameters of the model were adjusted to increase the area under the curve (AUC) of the test. Since the balance of the classes in the training dataset is not ideal, the accuracy may not reflect the actual model performance due to overlearning. By using the AUC as an indicator,



**FIG. 1.** Example of collected images: (a) optical image, (b) acquired PL image, and (c) binarized PL image.

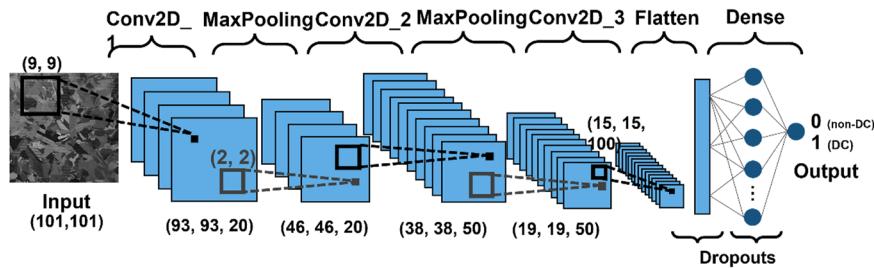


FIG. 2. Architecture of the 2D CNN.

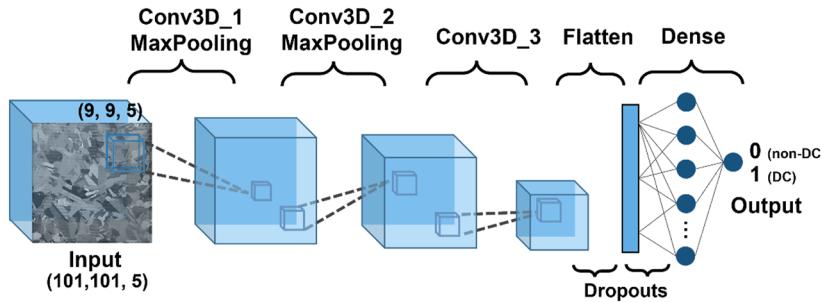


FIG. 3. Architecture of the 3D CNN.

we confirmed that the prediction was not random and determined the hyperparameters.

In our study, two CNNs were constructed with different dimensions: 2D and 3D.<sup>21,22</sup> In the 2D model, inputs were simple optical images and they were trained with labels from the corresponding PL images. It should be noted that balancing the distribution of classes is an essential process in the dataset. As can be seen in Fig. 1(c), the original given images contain more non-dislocation areas than dislocation areas. Namely, in the original images, only 0.6% of the pixels contain dislocations. This means that a model predicting “non-dislocation” for every single pixel still acquires an accuracy of 99.4%. To avoid this, we randomly selected images to create a training dataset containing ~25% of dislocations. The architecture is shown in Fig. 2. The dataset contained 37 953 images, which were split into training and validation datasets by a ratio of 80:20, each consisting of 30 362 and 7592 images. Test data were taken from alternate wafers. The architecture of the 3D model is shown in Fig. 3, where the model takes five optical images stacked in the growth direction and predicts the dislocations of the top wafer. Here, the model learns not only from the crystal structures, such as grain boundaries and orientations, but also from their changes in association with crystal growth. The dataset contained 21 426 images, split into training and validation datasets, each consisting of 17 140 and 4286 images.

### C. Grad-CAM, providing visual explanations of the model

The analysis of the model is performed by Gradient-based Class Activation Mapping (Grad-CAM).<sup>23</sup> Generally, the interpretations that a machine learning model makes are not visible to the eye. This

is referred to as the “black-box problem,” where the lack of interpretability and visibility delimits the model’s trustworthiness.<sup>24–26</sup> Recent developments in interpretability have suggested using the weights of the pre-trained network to visualize the model’s reasoning in producing the output. While this technique has provided more trust and reliability in neural networks, we believe that visualizing the model’s reasoning in complicated problems may provide extended benefits, such as teaching us something that humans are not aware of yet. Here, we use gradients of the pre-trained model to calculate the weights of the feature maps on the class score and generate coarse heatmaps highlighting the most discriminant areas of the given images.

The Grad-CAM is operated in three simple steps: First, the gradients of  $y^c$  (the score for a predicted class  $c$ ) are calculated with respect to the feature map  $A^k$ , which is  $\frac{\partial y^c}{\partial A^k}$ . Note that  $k$  is the number of feature maps produced from the respective convolutional layer, with width dimensions  $i$  and height dimensions  $j$ . Then, the gradients  $y^c$  are global-average-pooled with the height and width dimensions of the feature maps to obtain the neuron importance weights  $\alpha_k^c$ ,

$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k}. \quad (1)$$

Finally, a coarse heatmap indicating the most important regions is obtained by calculating a weighted sum of the feature maps followed by a ReLU,

$$L_{Grad-CAM}^c = ReLU \left( \sum_k \alpha_k^c A^k \right). \quad (2)$$

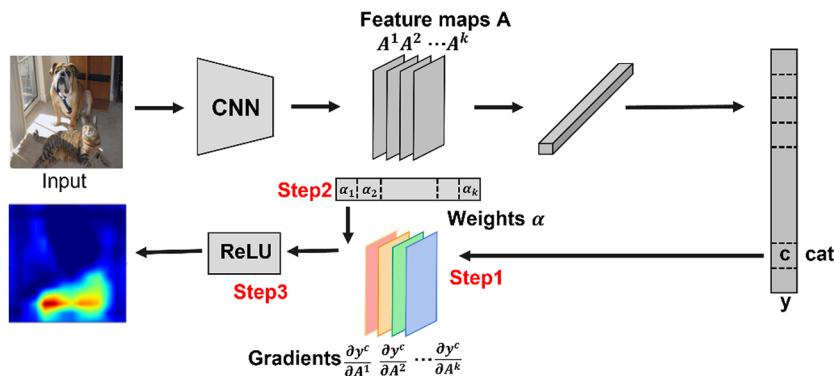


FIG. 4. The three steps required to obtain Grad-CAM.

The obtained coarse highlights of the discriminant regions are overlaid on the original images to visualize the most important regions. A graphical explanation of the Grad-CAM is shown in Fig. 4.

### III. RESULTS

#### A. Results of training on the 2D model

Figure 5 shows the predicted likelihoods for two test wafers along with the ground truth for 2D prediction. The reliability of the model can be seen from the receiver operating characteristic (ROC) curve shown in Fig. 6, which plots true positive rate vs false positive rate.<sup>27,28</sup> The AUC is a typical efficiency statistic for the ROC curves, and the larger the area under the red curve is, the more separability the model shows. Note that the model was trained on

30 362 images randomly selected from nine wafers (Nos. 502, 506, 509, 513, 516, 519, 523, 526, and 541 in the supplied datasets in the repository<sup>29</sup>), each shaped (101, 101). For test, two full-sized wafers were used. Each (101, 101) area was sequentially taken out, and the existence of dislocations at the center pixels was predicted. The outputs were integrated into predicted likelihood maps with the size of (400, 400). Furthermore, we improved the predictions by data augmentation. Data augmentation is a method to increase data when it is assumed that there are insufficient data to train the network. In many cases, simple transformations, such as rotating or flipping the original image, filtering, or desaturating, are effective.<sup>30,31</sup> The selection of data augmentation techniques is embraced on the condition that they are “safe” as in preserving the labels post-transformation.<sup>30</sup> Although these minorly altered images are inherently the same data, the neural network treats them as novel data, hence helping

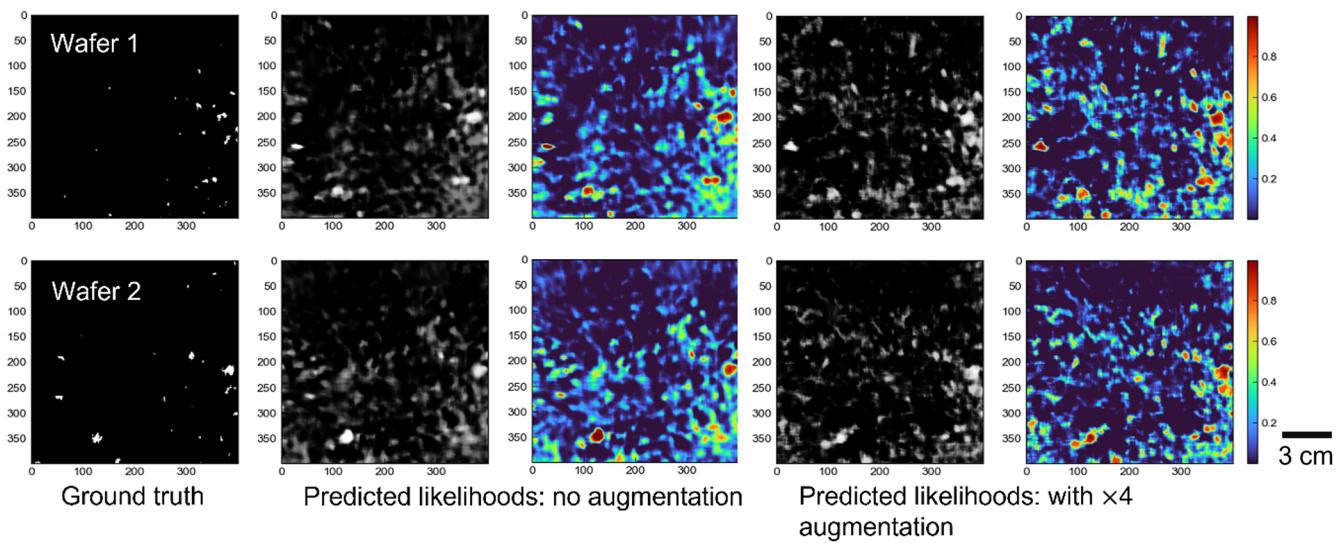
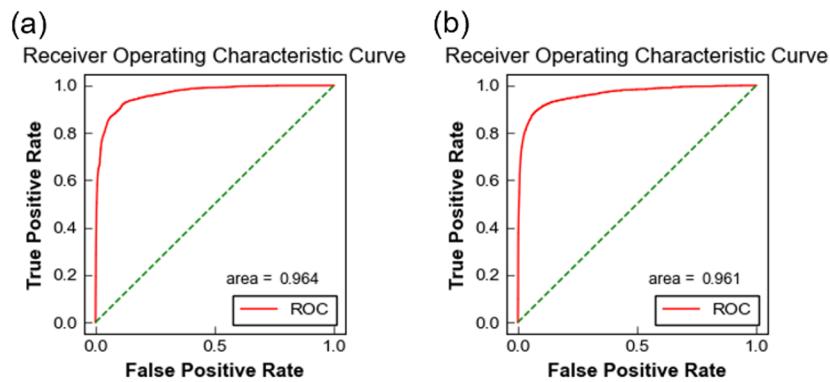
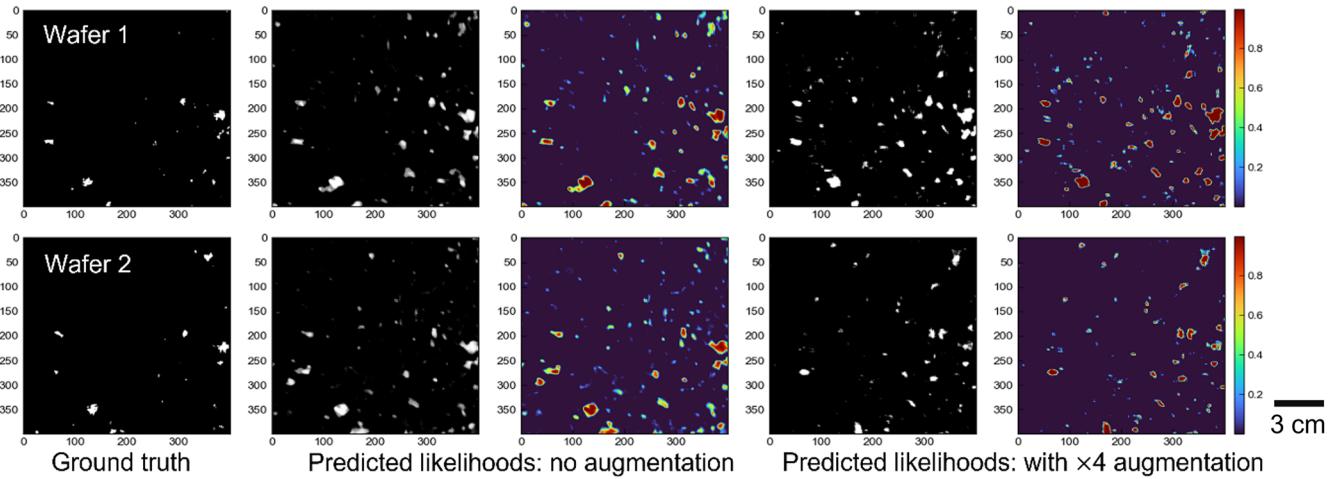


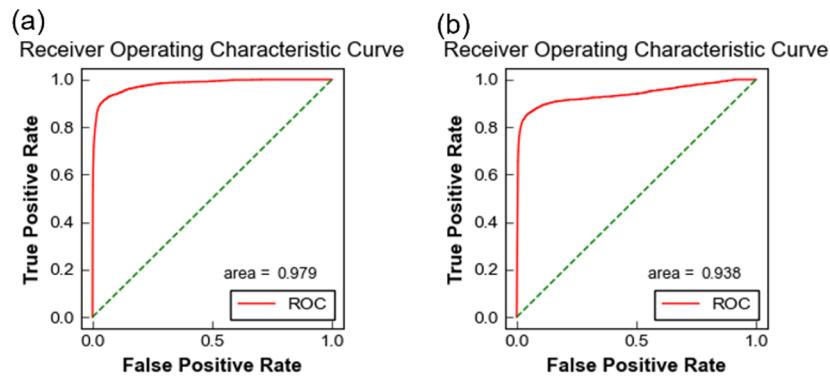
FIG. 5. Ground truth and predicted likelihoods in grayscale and colormap for two wafers for the 2D model. Predicted likelihoods are shown for two models: one trained with non-augmented inputs and the other trained with  $\times 4$  augmented inputs.



**FIG. 6.** ROC curve for test data predicted on the 2D model pre-trained with (a) non-augmented and (b) augmented inputs.



**FIG. 7.** Ground truth and predicted likelihoods in grayscale and colormap for two wafers for the 3D model. Predicted likelihoods are shown for two models: one trained with non-augmented inputs and the other trained with  $\times 4$  augmented inputs.



**FIG. 8.** ROC curve for test data predicted on the 3D model pre-trained with (a) non-augmented and (b) augmented inputs.

**TABLE I.** Training dataset size, accuracy, and test AUC for 2D and 3D models.

Model	Augmentation	Training dataset size	Binary accuracy	Test AUC
2D	None	30 362	0.975	0.964
2D	×4	121 645	0.960	0.961
3D	None	21 426	0.979	0.979
3D	×4	85 704	0.988	0.938

to improve the model's generalizability and trustworthiness. Particularly in our case, the original luminance profiles were rotated by 90°, increasing data to four times the original size. The model was trained and tested on the augmented data and compared with the non-augmented training.

## B. Results of training on the 3D model

Figure 7 shows the predicted likelihoods for two predicted wafers for the 3D model with non-augmented and augmented inputs. The ROC curve is displayed in Fig. 8. Note that training was performed on 21 426 image sets consisting of five wafers stacked in the growth direction (Nos. 461, 481, 502, 521, and 543 for 3D in the supplied datasets in the repository<sup>29</sup>), each shaped (101, 101, 5).

Test was performed on two full-sized wafer sets. The existence of dislocations at the center pixel of the top wafer was predicted. These predictions were integrated into likelihood maps of (400, 400). The size of the training dataset, the accuracy of predictions, and the test AUC are displayed in Table I, showing the comparison of the 2D augmented and non-augmented models and 3D augmented and non-augmented models.

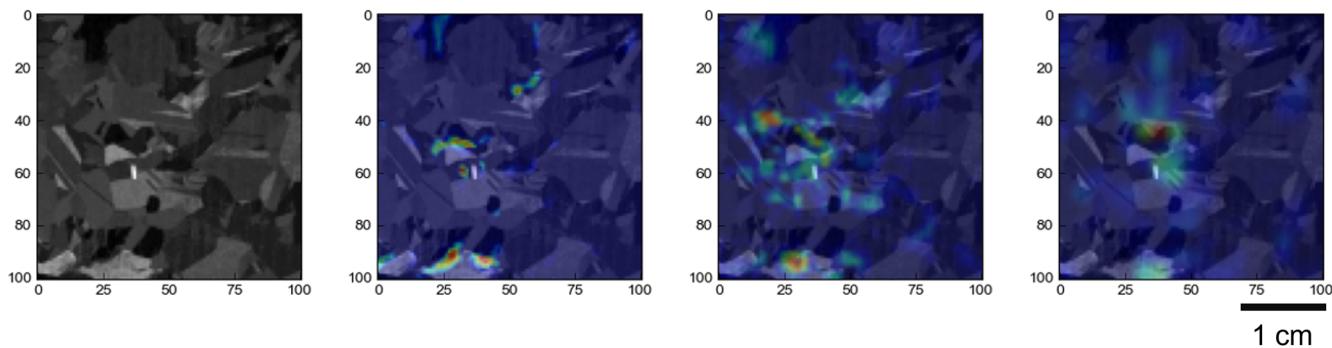
## C. Results of analysis by Grad-CAM

The Grad-CAM was operated on the 2D and 3D models trained with the augmented data. In Fig. 9, we show an example of the highlights for the 2D model. The highlights are created for each of the three convolutional layers. Figure 10 shows an overlay of the true and predicted dislocation clusters from the 3D model on a corresponding optical image. For the same model, Fig. 11 shows the examples of highlights obtained for nine regions.

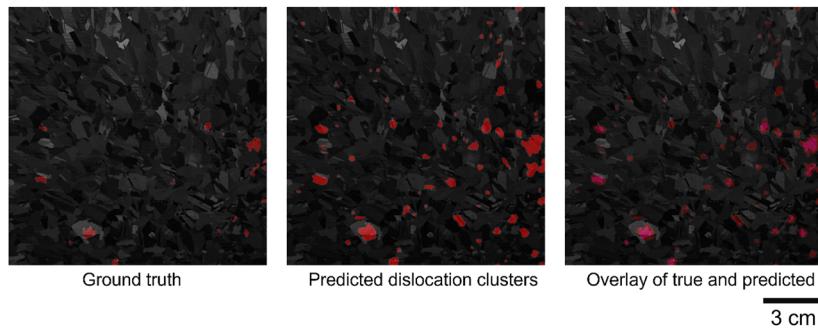
## IV. DISCUSSION

### A. Performance of prediction

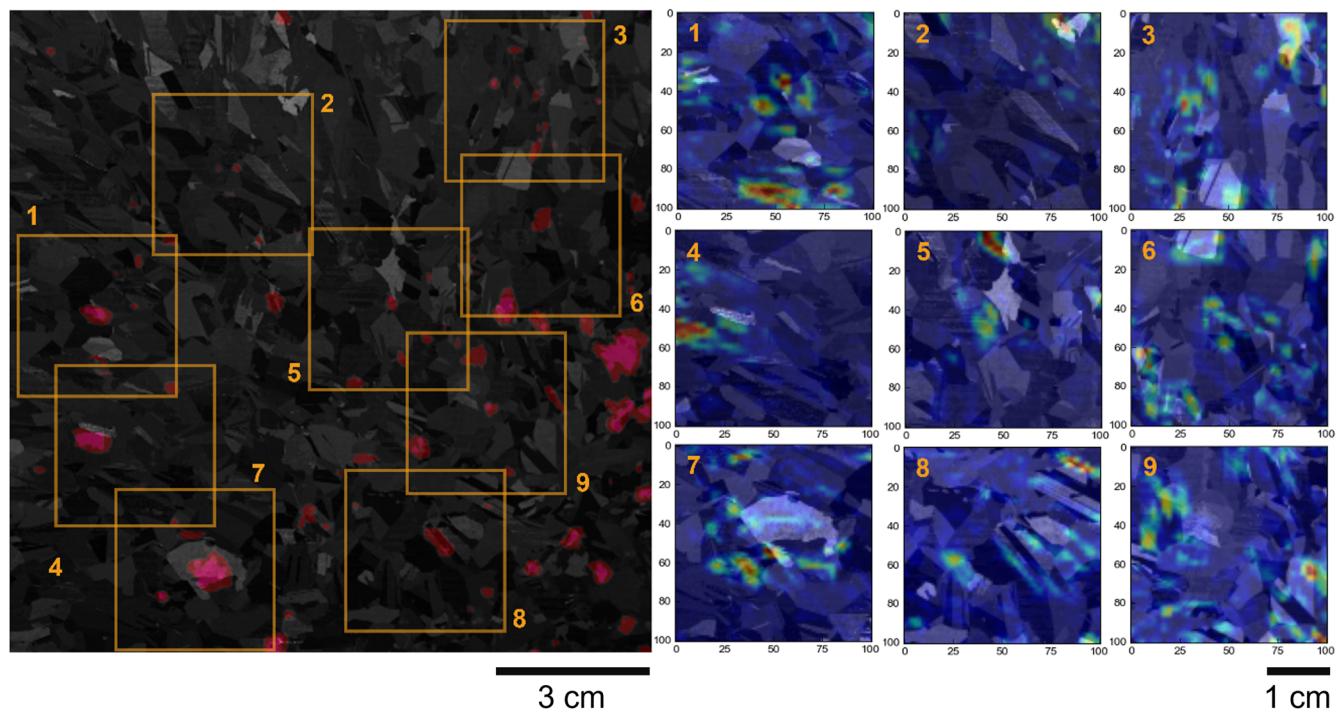
Our models have a high AUC of over 0.93, successfully separating the dislocation areas from the rest. The likelihood maps show that the main dislocation areas are predicted correctly, while the model has predicted a larger area as dislocations than



**FIG. 9.** Highlights of discriminant regions for each layer in the 2D CNN. Shown in grayscale is the given optical image. The colormaps show the overlays of highlights for conv2D\_1, conv2D\_2, and conv2D\_3 layers, respectively.



**FIG. 10.** Overlay of ground truth and predicted dislocation clusters, both shown on a given optical image. In the overlay of ground truth and predicted dislocation clusters, the predicted dislocations are displayed in red and the true dislocations are shown in magenta.



**FIG. 11.** Highlights of discriminant regions for several regions in the 3D CNN.

the actual. Data augmentation does not change the prediction scores significantly, but the pixels with intermediate likelihood values declined after augmentation. This is shown by the decline of the green or yellow areas in the colored likelihood maps. This indicates that by data augmentation, the model has generated more confidence in its predictions. This is because the data variety and abundance provided by rotating the input images have helped the model generalize more. Moreover, the 3D model has a higher accuracy and good separability compared to the 2D model. The likelihood maps are outstandingly clearer and in good contrast, indicating that the model has certainly separated areas with and without dislocation clusters.

From the likelihood maps and the ROC curve, it can be said that our model has been sufficiently trained to undergo analysis.

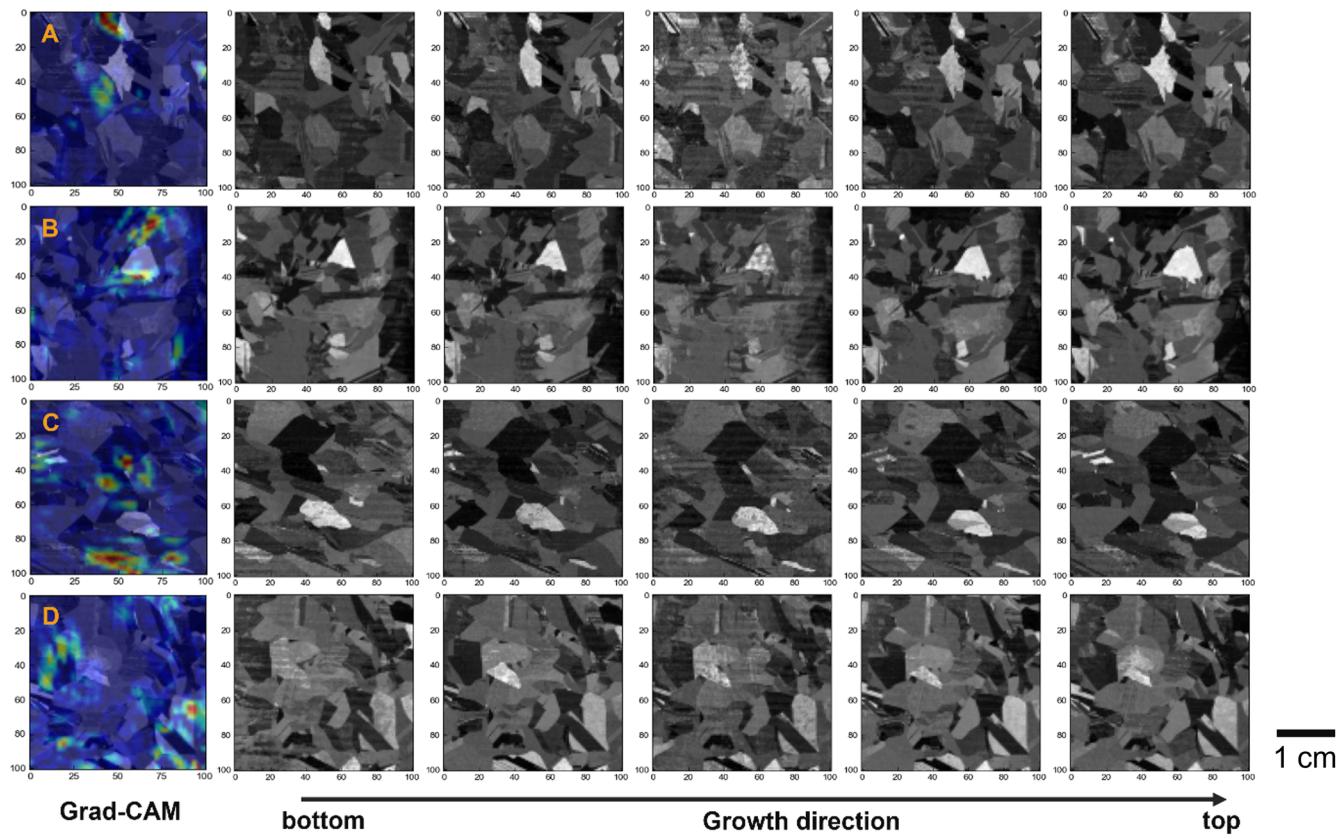
### B. Visualization by Grad-CAM

Here, we discuss the visualizations from the 3D model with augmented inputs, which is expected to have the most reliability. In Fig. 11, the regions where the change in the values of the input pixels affects the class of prediction are distinctly observed. For regions 1, 4, and 5, the highlights are significant for a selected area. In regions 6 and 9, the highlights are broader and the model seems to be looking at the multiple complex structures. The areas where dislocations existed are commonly highlighted but do not seem to have the most impact. It is remarkable that none of the highlights were of the center pixels alone. Despite the fact that the model was trained on the dislocation presence of the center pixel of the given image,

the characteristics of the areas far from the center are often more discriminative. It has also been found that when fed with slightly shifted or 90°-rotated images, the same grains were highlighted. It seems that the data augmentation resulted in providing the rotation invariance to the model in addition to the intrinsic translation invariance of CNN. As a result, the model acts as if it were chasing several specific grains or areas, which indicates the presence of characteristics with a large responsibility for the generation of dislocations.

Shown in Fig. 12 are the visualizations compared with the changes in crystal structures for four regions. In region A, the grain boundary at the highlighted area has arched due to the growth of an adjoining crystal. In region B, the model focuses on the area that is pushed outward by an enlarging grain. In region C, the highlights are near two shrinking grains at the center and also above complex structures at the bottom of the image. In region D, highlighted are several clusters of small grains and extending or narrowing grain boundaries. These changing crystallographic structures are of great importance in the generation of dislocation clusters, and the model has successfully extracted the key structures.

Dislocation generation in polycrystalline materials depends not only on the complexity of the microstructure and the diversity of grain boundaries but also on the material manufacturing process. Even for materials science experts, it is difficult to determine the presence or absence of dislocations from optical images. The results of this study will provide experts with insights into the fundamental nature of this complex problem and make a significant contribution to the deepening of scientific knowledge.



**FIG. 12.** Grad-CAM visualizations and the optical images fed into the model. The optical images are aligned in the direction of crystal growth.

## V. CONCLUSIONS

We developed a method to address and analyze dislocation clusters, a type of crystal defect that is detrimental to solar cells. This was carried out through feature extraction of optical images by convolutional neural networks and an analysis based on gradient-based class activation mapping. This technique outperforms baseline techniques in the analysis of the crystal structure in the time required and in the elimination of human bias. Remarkably, areas far from the center of the given images were often of more interest to discriminate than the center pixels. From a close look at the highlights, it was seen that the model reasons on significant changes in grain boundaries and grain shapes or complex small structures. This indicates that a machine learning model is extremely effective in extracting the tendencies statistically from the images. This is beneficial in analyzing a large amount of crystallographic data because image features are difficult to interpret. We believe that this technique will become a guideline for addressing large data in order to clarify mechanisms regarding crystal growth. Future work seeks a comprehensive understanding of the generation of crystal defects through increased data from various crystals.

## ACKNOWLEDGMENTS

This work was supported by JST CREST (Grant No. JPMJCR17J1).

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Author Contributions

**Kyoka Hara:** Conceptualization (equal); Data curation (lead); Investigation (lead); Methodology (lead); Visualization (lead); Writing – original draft (lead). **Takuto Kojima:** Conceptualization (lead); Investigation (supporting); Supervision (equal); Validation (lead). **Kentaro Kutsukake:** Conceptualization (equal); Investigation (supporting); Methodology (equal); Supervision (supporting). **Hiroaki Kudo:** Conceptualization (supporting); Investigation (supporting); Methodology (supporting); Supervision (supporting).

**Noritaka Usami:** Funding acquisition (lead); Supervision (lead); Writing – review & editing (lead).

## DATA AVAILABILITY

The data that support the findings of this study are openly available in GitHub at [https://github.com/UsamiCREST/DC\\_prediction](https://github.com/UsamiCREST/DC_prediction).<sup>29</sup>

## REFERENCES

- <sup>1</sup>Z. Fang, B. Bueken, D. E. De Vos, and R. A. Fischer, *Angew. Chem., Int. Ed.* **54**, 7234 (2015).
- <sup>2</sup>J. Weertman, D. Farkas, K. Hemker, H. Kung, M. Mayo, R. Mitra, and H. V. Swygenhoven, *MRS Bull.* **24**, 44 (1999).
- <sup>3</sup>D. W. Boukhvalov and M. I. Katnelson, *Nano Lett.* **8**, 4373 (2008).
- <sup>4</sup>M. A. Meyers, A. Mishra, and D. J. Benson, *Prog. Mater. Sci.* **51**, 427 (2006).
- <sup>5</sup>J. Li, L. Porter, and S. Yip, *J. Nucl. Mater.* **255**, 139 (1998).
- <sup>6</sup>A. Alkauskas, M. D. Mccluskey, and C. G. Van de Walle, *J. Appl. Phys.* **119**, 181101 (2016).
- <sup>7</sup>J. Gil Sevillano, P. van Houtte, and E. Aernoudt, *Prog. Mater. Sci.* **25**, 69 (1980).
- <sup>8</sup>R. Z. Valiev, R. K. Islamgaliev, and I. V. Alexandrov, *Prog. Mater. Sci.* **45**, 103 (2000).
- <sup>9</sup>M. Li, C. Zhang, M. Li, F. Liu, L. Zhou, Z. Gao, J. Sun, D. Han, and J. Gong, *Chem. Eng. J.* **429**, 132450 (2022).
- <sup>10</sup>G. Müller and J. Friedlich, *J. Cryst. Growth* **1–3**, 1 (2004).
- <sup>11</sup>H. J. Möller, C. Funke, M. Rinio, and S. Scholz, *Thin Solid Films* **487**, 179 (2005).
- <sup>12</sup>S. Castellanos, K. E. Ekstrom, A. Autruffe, M. A. Jensen, A. E. Morishige, J. Hofstetter, P. Yen, B. Lai, G. Stokkan, C. del Canizo, and T. Buonassisi, *IEEE J. Photovoltaics* **6**, 632 (2016).
- <sup>13</sup>B. Sopori, P. Rupnowski, V. Mehta, V. Budhraja, S. Johnston, N. Call, H. Mountinho, M. Al-Jassim, A. Shaikh, M. Seacrist, and D. Carlson, *ECS Trans.* **18**, 1049 (2009).
- <sup>14</sup>V. Kveder, M. Kittler, and W. Schroter, *Phys. Rev. B* **63**, 115208 (2001).
- <sup>15</sup>B. Ryninen, G. Stokkan, M. Kivambe, T. Ervik, and O. Lohne, *Acta Mater.* **59**, 7703 (2011).
- <sup>16</sup>D. Oriwol, M. Trempa, L. Sylla, and H. S. Leipner, *J. Cryst. Growth* **463**, 1 (2017).
- <sup>17</sup>Y. Fukuda, K. Kutsukake, T. Kojima, and N. Usami, *CrystEngComm* **24**, 1948 (2022).
- <sup>18</sup>H. Kato, S. Kamibeppu, T. Kojima, T. Matsumoto, H. Kudo, Y. Takeuchi, K. Kutsukake, and N. Usami, *IEEE Trans. Electr. Eng.* **17**, 1685 (2022).
- <sup>19</sup>K. Hara, T. Kojima, K. Kutsukake, H. Kudo, and N. Usami, *APL Mach. Learn.* **1**, 026113 (2023).
- <sup>20</sup>Y. LeCun, Y. Bengio, and G. Hinton, *Nature* **521**, 436 (2015).
- <sup>21</sup>H. Shin, H. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. Summers, *IEEE Trans. Med. Imaging* **35**, 1285 (2016).
- <sup>22</sup>M. Paoletti, J. Haut, J. Plaza, and A. Plaza, *ISPRS J. Photogramm. Remote Sens.* **145**, 120 (2018).
- <sup>23</sup>R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, *Int. J. Comput. Vision* **128**, 336 (2020).
- <sup>24</sup>R. Guidotti, A. Monreale, S. Ruggieri, F. Turin, F. Giannotti, and D. Pedreschi, *ACM Comput. Surv.* **51**, 93 (2018).
- <sup>25</sup>A. B. Arrieta, N. Diaz-Rodriguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, *Inf. Fusion* **58**, 82 (2020).
- <sup>26</sup>J. Schmidt, M. R. G. Marques, S. Botti, and M. A. L. Marques, *npj Comput. Mater.* **5**, 83 (2019).
- <sup>27</sup>A. P. Bradley, *Pattern Recognit.* **30**, 1145 (1997).
- <sup>28</sup>J. Huang and C. Ling, *IEEE Trans. Knowl. Data Eng.* **17**, 299 (2005).
- <sup>29</sup>K. Hara, T. Kojima, K. Kutsukake, H. Kudo, and N. Usami, Usami-CREST/DC\_prediction, 2023, GitHub [https://github.com/UsamiCREST/DC\\_prediction](https://github.com/UsamiCREST/DC_prediction).
- <sup>30</sup>C. Shorten and T. M. Khoshgoftaar, *J. Big Data* **6**, 60 (2019).
- <sup>31</sup>D. A. van Dyk and X. L. Meng, *J. Comput. Graphical Stat.* **10**, 1 (2001).