

Deep learning-based perspective distortion correction for outdoor photovoltaic module images

Yun Li^{*}, Brendan Wright, Ziv Hameiri

School of Photovoltaic and Renewable Energy Engineering, The University of New South Wales, Sydney, Australia



ABSTRACT

To address the growing demand for photovoltaic inspection, an increasing volume of photoluminescence, electroluminescence, and infrared image datasets is being produced for research and analysis. Most of these images exhibit perspective distortions, which introduce challenges in subsequent analysis, such as defect detection and classification. Hence, distortion correction is required during the preprocessing stage. However, manually correcting each dataset to eliminate distortions is labour-intensive. This paper addresses this issue by developing a convolutional neural network model to estimate the camera parameters of yaw, pitch, and roll. A heuristic method was referenced as a comparative benchmark, demonstrating that the developed deep learning-based approach is equally accurate and significantly faster. The findings underscore the potential of deep learning techniques in automating and refining image analysis in photovoltaic diagnostics, offering substantial improvements over traditional methods in terms of speed and scalability.

1. Introduction

Addressing global warming has intensified the need for sustainable, clean energy sources, thereby accentuating the significance of advancements in photovoltaic (PV) technologies [1]. In this realm, maximising both the efficiency and lifespan of solar modules is crucial for the economic viability and overall effectiveness of PV systems [2].

Regular and accurate diagnostics are important for maintaining the optimal performance of PV modules [3]. Such diagnostics are fundamental for ensuring that the modules operate at their peak efficiency throughout their nominal service life [4]. In the context of these diagnostic processes, luminescence imaging [5] stands out as a critical inspection tool [6]. It plays a significant role in the evaluation of the quality of PV modules [7]. However, the manual interpretation of luminescence images is a challenging task [8] that can limit the scalability and accessibility of fault detection [9]. Automated image analysis methods eliminate these limitations through enhancements in speed, accuracy, and throughput [10]. Pre-processing of the source images is a key requirement for any of these automated methods [11]; it primarily involves the correction of distortions introduced by the camera [11]. Two common types of camera distortion are *perspective distortion* and *lens distortion* [12]. Perspective distortion [13] occurs due to the way humans perceive depth and spatial relationships in a three-dimensional world when they are projected onto a two-dimensional medium like a photograph, while lens distortion [14] refers to aberrations directly caused by the lens optics, such as barrel distortion (where images bow

outwards) [15] and pincushion distortion (where images bow inwards) [15]. Therefore, the distortion correction step is indispensable for accurate information extraction from luminescence images, thereby augmenting the quality and reliability of the diagnostics derived from them [16,17].

Correction of image distortions (perspective) is essential for various imaging-based applications [18,19]. Mantel et al. [20] developed mathematical-based correction algorithms that detect lines and corners by rotating the binary image at different angles. These algorithms then employ projective sum methods to compute the angular parameters and eventually utilise matrices for distortion correction. Kölblin et al. [21] then improved this process by using an iterative Hough transform [22] for line structure detection. In a similar vein, Yubin et al. [23] employed a minimum residual method to ascertain the optimum inverse distortion matrix, further enhancing the accuracy of the correction process. Another approach relies on the OpenCV image processing library [24] and the calibration processes developed by Zhang et al. [25] and Heikkilä et al. [26]. This approach utilises a chessboard grid of control points as the calibration pattern, which accounts for radial distortions of multiple orders and decentring distortion.

Machine learning (ML)-based approaches have recently been used for distortion corrections. Compared with traditional methods, ML methods offer benefits such as reduced computational cost, improved accuracy, and the ability to learn and adapt to diverse imaging conditions. Zhao et al. [27] introduced a learning-based method using deep neural networks to undistort images, particularly useful in enhancing

* Corresponding author.

E-mail address: yun.f.li@unsw.edu.au (Y. Li).

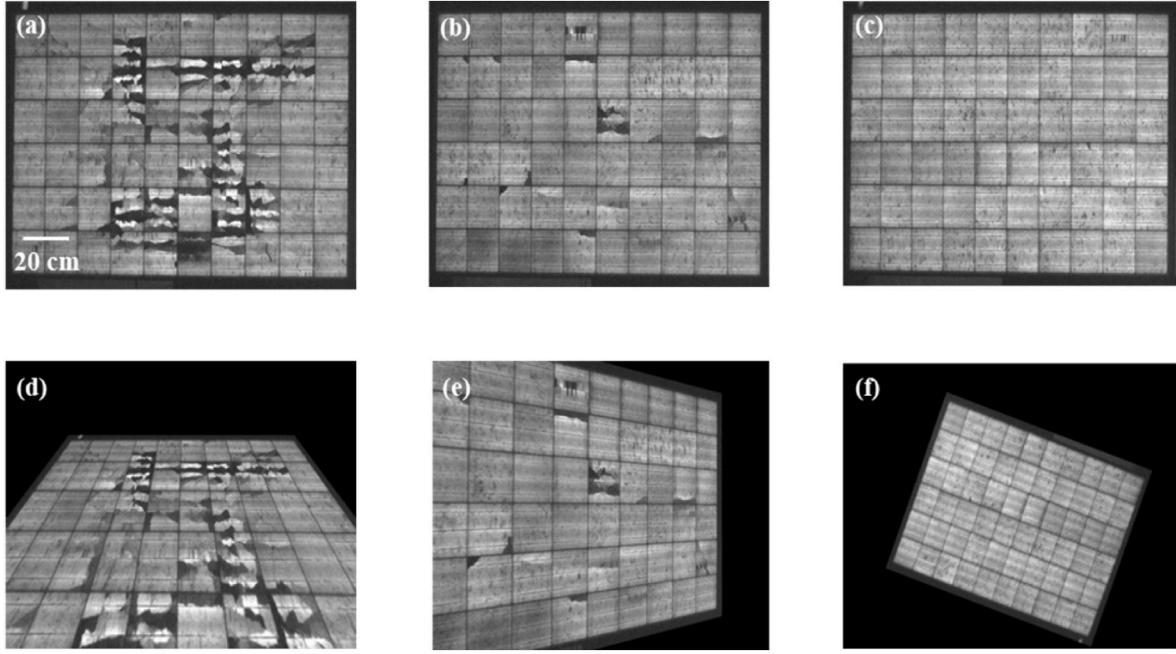


Fig. 1. (a–c) Representative source images and transformed images by (d) yaw, (e) pitch, and (f) roll.

the visual quality of portraits. Further developments were made by Del Gallego et al. [28] and Wang et al. [29], who developed methods using parallel convolutional neural networks and a perspective-aware 3D generative adversarial network (GAN), respectively, to automatically correct distortions in digital imaging. These approaches not only improve the accuracy of diagnostics but also contribute to the scalability of image processing techniques, for example, for effective blind correction techniques in processing a wide range of images in real time, as discussed by Del Gallego et al. [28]. Wu et al. [30] then demonstrated how adaptive algorithms and spectral analysis can be utilised to further refine the image correction process by developing a self-adaptive method that can be adjusted based on the detected distortion levels, aiming to optimise processing time and resource usage. However, none of these methods were designed considering the automation and robustness required for outdoor PV module inspection.

This paper introduces a novel approach for distortion correction, utilising supervised regression models. The developed correction model has been trained on a simulated dataset that closely mimics electroluminescence (EL) images. The model was then validated using measured outdoor distorted EL images. It is shown that the model predicts the three distortion parameters with an error smaller than 10 %, a clear demonstration of its effectiveness.

2. Methodology

The perspective distortion is impacted by the orientation of the camera in three-dimensional (3D) space [12]. It is characterised by three angles: yaw, pitch, and roll [31]. Yaw is the rotation around the z -axis, determining the camera's left-right orientation. Pitch is the rotation around the y -axis, controlling the up-down tilt. Roll is the rotation around the x -axis and defines the camera's tilt to its sides. These angles are translated into a combined 3D rotation matrix:

$$R = R_z(\gamma)R_y(\beta)R_x(\alpha) = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix} \quad (1)$$

where R is rotation, γ is the roll angle, β is the pitch angle, and α is the yaw angle.

Our proposed method uses an ML model, ResNet18 [32], to predict these three angles. The extracted angles are then used with the above rotational matrices [33] to correct the distortions.

2.1. Image dataset

Sixty EL images were used as ‘source’ images. Utilising randomly selected 3D rotation matrices (each with different yaw in the range of -30° – 30° , pitch in the range of -30° to 30° , and roll in the range of -90° to 90°), distorted images were created using the OpenCV library [24] and its in-built function (“get_perspective”). A training dataset of approximately 10,000 images was generated using 48 source images, while the additional 12 distinct source images were utilised to create a separate testing set of 1000 images. Fig. 1 presents representative source and transformed images.

2.2. Machine learning model

The developed ML-based approach utilises the architecture of ResNet18, characterised by convolutional and fully connected layers, complemented by suitable activation functions and pooling operations. ResNet18 has been designed to address the problem of vanishing gradients, a challenge that often occurs when training deep neural networks [34]. It addresses this issue through an innovative use of residual connections, which allow the network to learn identity functions effectively, ensuring that deeper layers can perform at least as well as shallower layers. To monitor the model’s progression, checkpoints were implemented every five epochs, providing insights into the model’s development throughout the training phase. The Adam optimiser [35], known for its efficient convergence properties, was selected for training,

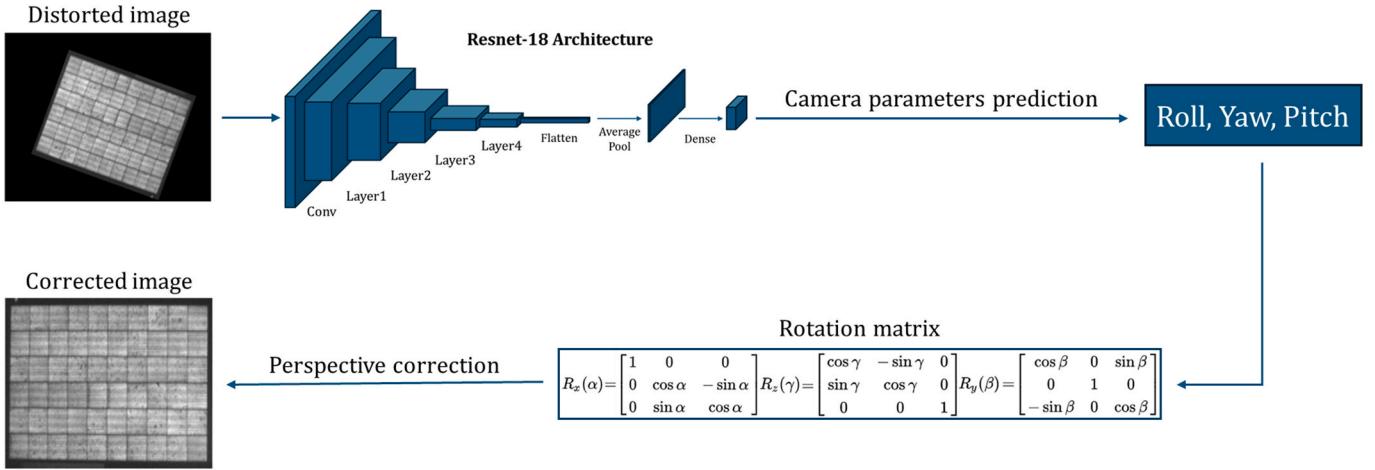


Fig. 2. Schematic of the used machine learning pipeline.

with a learning rate of 0.01. The root mean squared error (RMSE) was selected to serve as the loss function, offering a dependable measure of the model's predictive accuracy [36]. Pre-processing was performed to standardise the input data by resizing the images to 440×360 pixels. The model's performance was validated using a separate validation set. The schematic of the used ML pipeline is shown in Fig. 2.

2.3. The heuristic method

To validate the performance of the proposed ML model, a heuristic method [37] was utilised. The heuristic method was selected for its robustness in perspective correction across different PV luminescence image datasets. It performs a Radon transform [38] to the input images resulting in rotation maps or sinograms. From the image sinogram, the

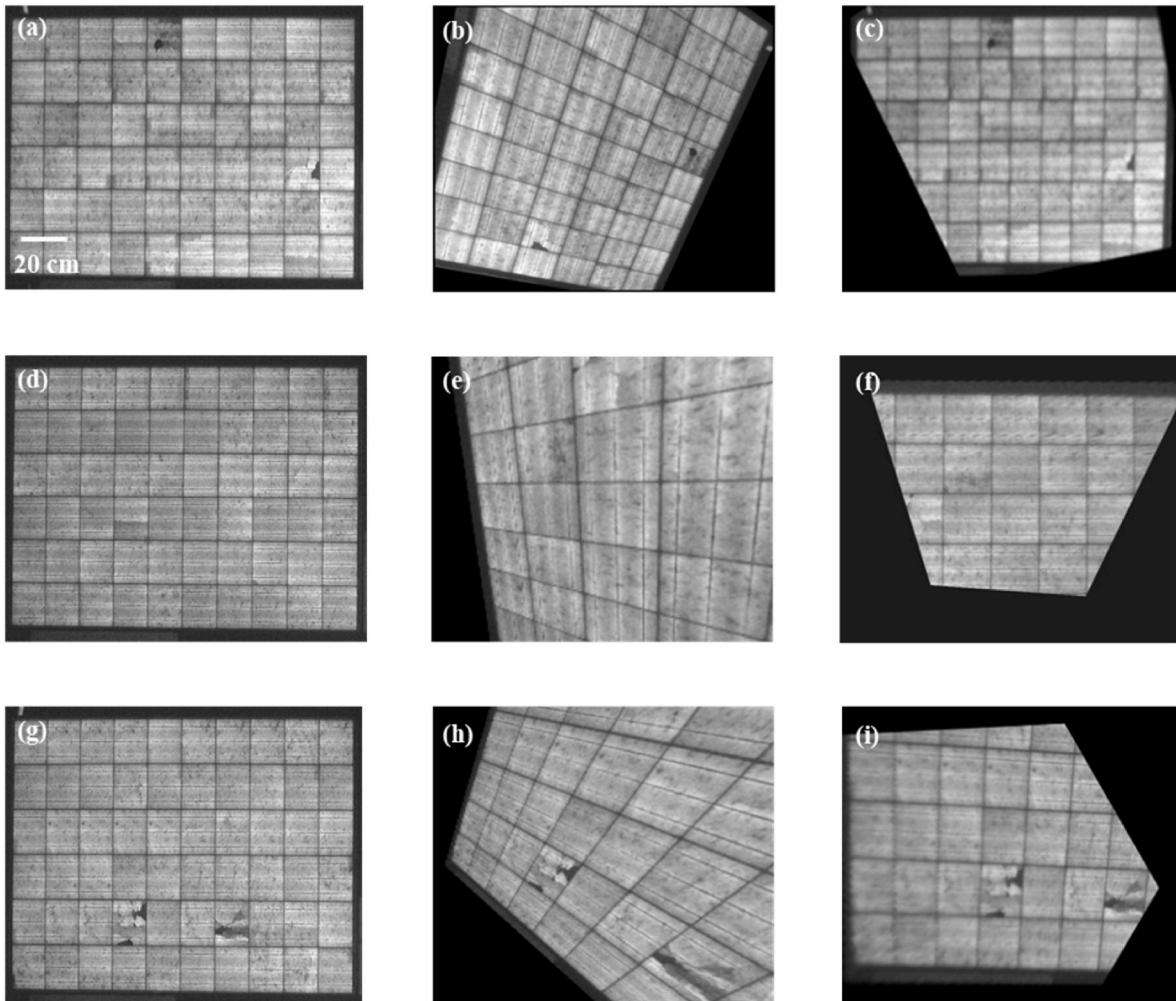


Fig. 3. (a, d, g) Representative source images; (b, e, h) corresponding distorted images; and (c, f, i) corresponding corrected images by the proposed method.

in-plane module perspective can be obtained through morphological operations [39], such as gradients, followed by smoothing techniques like Gaussian blurring [40]. Otsu's thresholding method [41] is then utilised to isolate significant structural features from the transformed image data. The final analysis stages involve labelling and filtering the detected features to isolate distinct lines. This process uses connected component labelling to identify contiguous edge points, followed by statistical filtering to remove outliers. Clustering techniques combined with linear regression models are then applied to these points to determine the linear relationships and angular orientations of the edges. With this methodology, module-plane angles are calculated in pixel coordinates, which can then be used to obtain the four angles of interest. The robustness of the heuristic method serves as a good benchmark for validating the proposed ML model (providing the 'ground truth'). However, its application to very large datasets is limited due to its computational expense from necessary dataset calibration and intensive matrix calculations. The proposed ML method aims to perform similarly effective corrections while decreasing the processing time.

3. Results and discussion

Fig. 3 demonstrates the effectiveness of the developed ML-based model by presenting representative original images (a, d, g) and their corresponding distorted (b, e, h) and corrected (c, f, i) images. The missing parts in the corrected images (c, f, i) are due to the large distortion and magnification used in the distorted images (b, e, h). As can be seen, the noticeable perspective distortions (the edges of the cells are misaligned with the image borders) have been corrected in **Fig. 3(c)** and (f) when compared to **Fig. 3(b)** and (e). The outcome underlines the potential of the proposed algorithm to rectify distortions, enhancing the accuracy of subsequent image analyses. **Fig. 3(h)** represents an extreme distortion of pitch and yaw (30° or -30°); however, even in this case, the images are accurately corrected, although some of the corners in **Fig. 3(i)** appear slightly blurry.

Fig. 4 presents the predicted versus the simulated ('true') angles for the entire dataset. As can be seen, across most of the angles, the model performs very well, accurately predicting the three angles. The RMSEs of both the pitch and yaw are very low (3.2° and 2.9° , respectively), with most data points within the region $y = x \pm 10\%$. The RMSE associated with the roll is larger (11.3°), mainly due to cases where the angle is above 85° or below -85° , values that approach the extreme limits which are inherently more difficult to predict due to the similarity between 90° clockwise and anticlockwise rotations. If all values above 85° and below -85° are removed, the RMSE of the roll is reduced to 2.8° , similar to the pitch and yaw.

Fig. 5 compares the heuristic and proposed methods using representative images. Despite the differences in the size and aspect ratio of the images due to the different implementations, the ML-corrected images (c, f) look very similar to the images corrected by the heuristic method (b, e). After correction, both the horizontal and vertical grids are

parallel. Importantly, while the heuristic method requires 42 s to correct an image, it only took 0.1 s for the ML-based method to complete the same process using identical hardware.

To validate the ML model's predictions against the heuristic method using more quantitative metrics, a random selection of 150 distorted images was utilised. The location of the four corners of each image was recorded. The heuristic method was then used to correct the images, while the ML-based method was implemented to predict the roll-pitch-yaw angles. Using OpenCV and the obtained angles, the heuristic method-corrected images were then transformed back into their distorted state. The location of the four corners was recorded again and compared to the four corners of the initial distorted images. If the ML model performs optimally, the rotation matrix will yield four corner points that perfectly align with the original points. Hence, the proximity of the corresponding points (the distance between them) provides a quantitative metric for validating the precision of the ML model's predictions. Figure A-1 in the Appendix presents representative examples of initial and predicted corners.

Fig. 6 presents the histogram of the average distance between the corner points of the original images and the predicted points by the ML-based method (transformed images). As can be seen, in the majority of cases, the prediction is within the range of 5 to 35 pixels (with a median of 27 pixels) compared to the image size of 720×880 pixels, with a corner-to-corner distance of 1137 pixels. Hence, the performance of the ML-based approach is almost identical to that of the heuristic method, although it is more than two orders of magnitude faster when using identical software. The higher errors are assumed to be related to cases of extreme distortions.

The results demonstrate the effectiveness of the ML-based approach, while also identifying areas for potential improvement. The model's performance on cases of extreme distortion, particularly for roll angles approaching $\pm 90^\circ$, presents an opportunity for refinement. Future work could consider expanding the training dataset to include a higher proportion of severely distorted images. Exploring alternative neural network architectures or ensemble methods might also enhance the model's generalisation capabilities.

The potential application of this approach to photoluminescence (PL) and infrared (IR) image datasets warrants further investigation. The structural similarities between EL, PL, and IR images of PV modules, such as cell grid patterns and module boundaries, suggest that transfer learning techniques could be applicable. However, fine-tuning on smaller, modality-specific datasets may be necessary. This approach could extend the utility of the distortion correction method across multiple PV module inspection techniques.

4. Conclusion

This study presents an ML approach for pre-processing PV module images. The developed model achieves an RMSE below 3° for the three main angles—yaw, pitch, and roll—while increasing processing speed

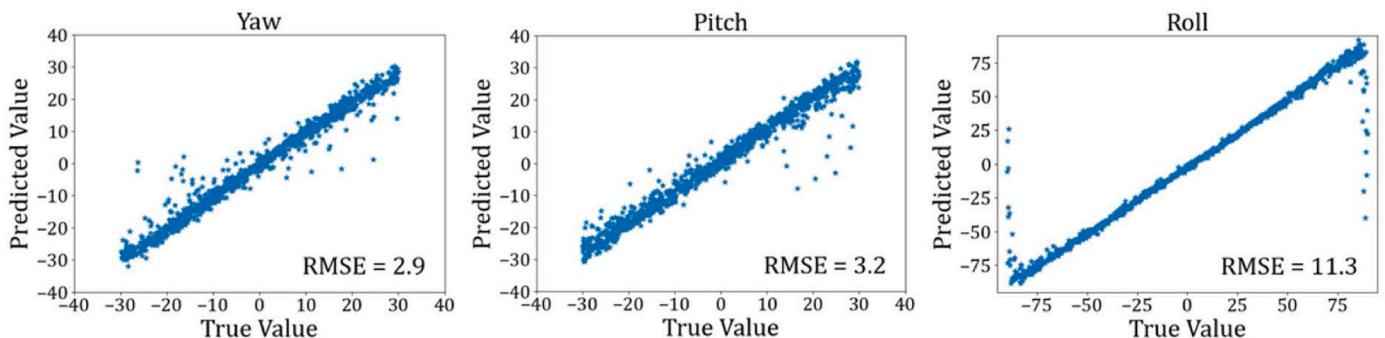


Fig. 4. Predicted vs true camera parameters: (a) pitch, (b) yaw, and (c) roll.

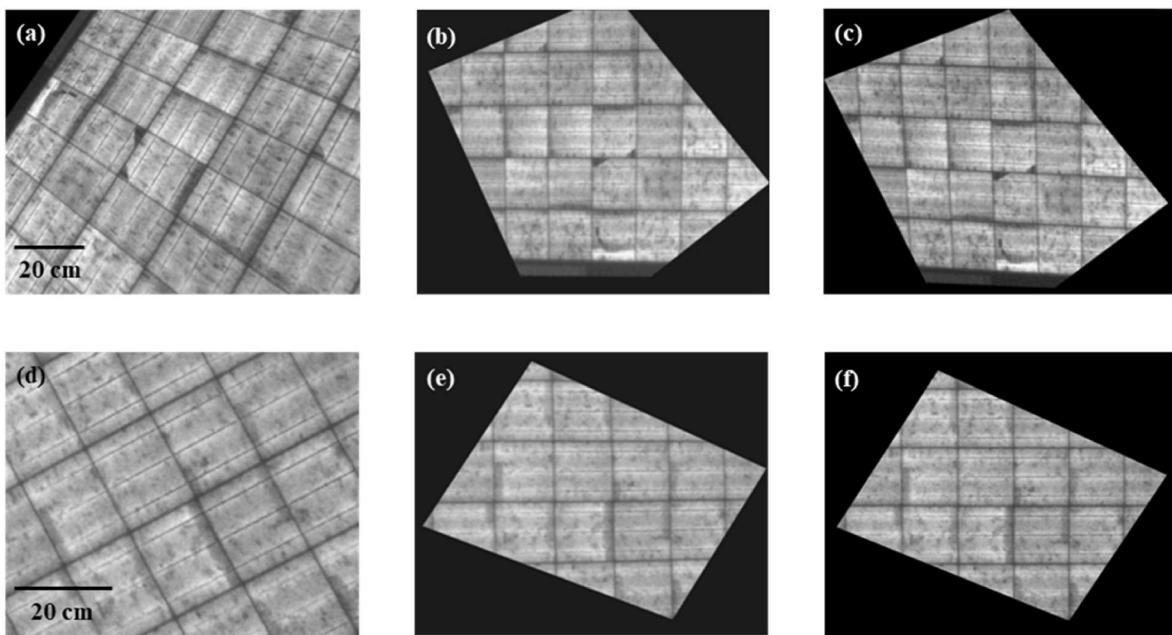


Fig. 5. (a, d) Representative source images (real-world distorted images), (b, e) corresponding corrected images by the heuristic method, and (c, f) corresponding ML-corrected images.

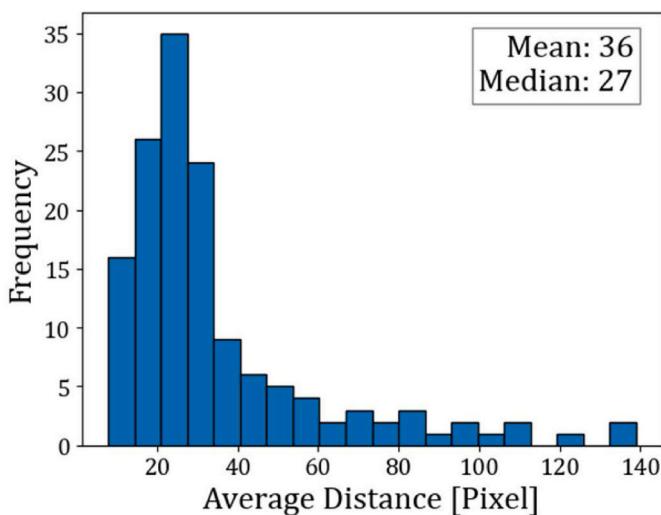


Fig. 6. Average distance between the original and transformed corner points.

by two orders of magnitude compared to the state-of-the-art heuristic method. This method significantly improves the de-skewing process of PV module images, a critical requirement for image-based diagnostics. The application of the proposed method enhances data visualisation and potentially improves performance assessment and prediction in subsequent analysis tasks. Future research will explore the model's applicability to other luminescence images, such as drone-captured EL images, PL images, and IR images. This model represents a significant advancement in PV module image de-skewing and sets a new benchmark for future research and practical applications in this field.

CRediT authorship contribution statement

Yun Li: Writing – original draft, Visualization, Visualization, Software, Methodology. **Brendan Wright:** Writing – review & editing, Visualization, Supervision, Conceptualization. **Ziv Hameiri:** Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Yun Li reports financial support was provided by Australian Renewable Energy Agency. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

The authors would like to thank A. Shakiba (UNSW) for helpful discussions. YL thanks SF for her long-term support. This work was supported by the Australian Government through the Australian Renewable Energy Agency (ARENA, Grant 2020/TRAC002). The views expressed herein are not necessarily the views of the Australian Government, and the Australian Government does not accept responsibility for any information or advice contained.

Appendix

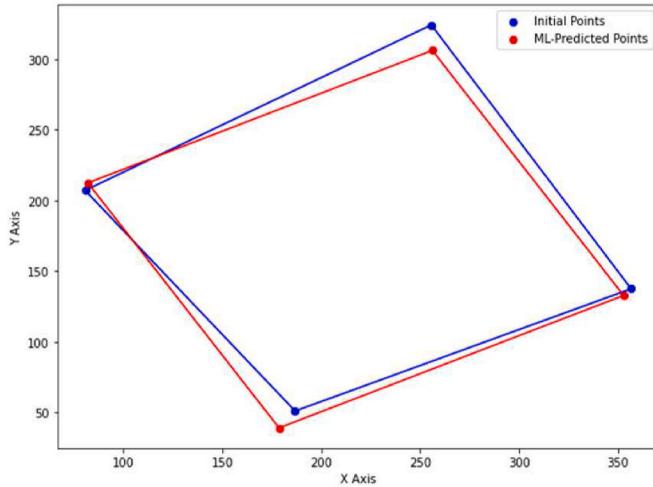


Figure A-1. Representative examples of initial and predicted corners.

References

- [1] Q. Khalid, J. Ikram, N. Arshad, A collaborative approach to operate high powered devices on small-scale PV systems, *Energy Proc.* 111 (2017) 895–903.
- [2] W.C. Since, Development of photovoltaic technologies for global impact, *Renew. Energy* 138 (2019) 911–914.
- [3] Q. Xu, X. Li, C. Feng, R. Wang, Review on fault characterization and diagnosis technique in photovoltaic systems, in: *IEEE Transportation Electrification Conference and Expo*, 2022, pp. 1–6.
- [4] R. Deng, N.L. Chang, Z. Ouyang, C.M. Chong, A techno-economic review of silicon photovoltaic module recycling, *Renew. Sustain. Energy Rev.* 109 (2019) 532–550.
- [5] D. Jaque, F. Vetrone, Luminescence nanothermometry, *Nanoscale* 4 (15) (2012) 4301–4326.
- [6] G. Liu, P. Dwivedi, T. Trupke, Z. Hameiri, Deep Learning model to denoise luminescence images of silicon solar cells, *Adv. Sci.* 10 (18) (2023) 2300206.
- [7] Y. Zhang, R. Wang, F. Wang, D. Zhu, X. Gong, X. Cheng, Electroluminescence as a tool to study the polarization characteristics and generation mechanism in silicon PV panels, *Appl. Sci.* 13 (3) (2023) 1591.
- [8] O. Kunz, J. Schlipf, A. Fladung, Y.S. Khoo, K. Bedrich, T. Trupke, Z. Hameiri, Outdoor luminescence imaging of field-deployed PV modules, *Progress in Energy* 4 (4) (2022) 042014.
- [9] G.A. dos Reis Benatto, et al., Development of outdoor luminescence imaging for drone-based PV array inspection, in: *44th IEEE Photovoltaic Specialist Conference*, 2017, pp. 2682–2687.
- [10] J. Starzyński, P. Zawadzki, D. Harańczyk, Machine learning in solar plants inspection automation, *Energies* 15 (16) (2022) 5966.
- [11] P. Kölblin, A. Bartler, M. Füller, Image preprocessing for outdoor luminescence inspection of large photovoltaic parks, *Energies* 14 (9) (2021) 2508.
- [12] W. Hugemann, Correcting lens distortions in digital photographs, *Ingenieurbüro Morawski+ Hugemann: Leverkusen, Germany* (2010) 20.
- [13] D. Santana-Cedrés, L. Gomez, M. Alemán-Flores, A. Salgado, J. Esclarín, L. Mazorra, L. Alvarez, Automatic correction of perspective and optical distortions, *Comput. Vis. Image Understand.* 161 (2017) 1–10.
- [14] J.M. Hanssen, B. Bourdoncle, J. Gresset, J. Faubert, P. Simonet, Distortion in Ophthalmic Optics: A Review of the Principal Concepts and Models, 2023 *arXiv preprint arXiv:2301.07194*.
- [15] B. Odeleye, G. Loukas, R. Heartfield, G. Sakellari, E. Panaousis, F. Spyridonis, Virtually secure: a taxonomic assessment of cybersecurity challenges in virtual reality environments, *Comput. Secur.* 124 (2023) 102951.
- [16] V.B. Kurukuru, A. Haque, M.A. Khan, Fault classification for photovoltaic modules using thermography and image processing, in: *IEEE Industry Applications Society Annual Meeting*, 2019.
- [17] Y. Zhao, L. Yang, B. Lehman, J.F. de Palma, J. Mosesian, R. Lyons, Decision tree-based fault detection and classification in solar photovoltaic arrays, in: *27th IEEE Applied Power Electronics Conference and Exposition*, 2012, pp. 93–99.
- [18] C. Mantel, F. Villebro, H.R. Parikh, S. Spataru, G.A. dos Reis Benatto, D. Sera, P. B. Poulsen, S. Forchhammer, Method for estimation and correction of perspective distortion of electroluminescence images of photovoltaic panels, *IEEE J. Photovoltaics* 10 (6) (2020) 1797–1802.
- [19] Z. Zhang, A flexible new technique for camera calibration, *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (2000) 1330–1334.
- [20] C. Mantel, S. Spataru, H. Parikh, D. Sera, G.A. dos Reis Benatto, N. Riedel, S. Thorsteinsson, P.B. Poulsen, S. Forchammer, Correcting for perspective distortion in electroluminescence images of photovoltaic panels, in: *7th World Conference on Photovoltaic Energy Conversion*, 2018, pp. 433–437.
- [21] P. Kölblin, A. Bartler, M. Füller, Image preprocessing for outdoor luminescence inspection of large photovoltaic parks, *Energies* 14 (9) (2021) 2508.
- [22] R.O. Duda, P.E. Hart, Use of the Hough transformation to detect lines and curves in pictures, *Commun. ACM* 15 (1) (1972) 11–15.
- [23] Y. Wu, S. Jiang, Z. Xu, S. Zhu, D. Cao, Lens distortion correction based on one chessboard pattern image, *Front. Optoelectron.* 8 (2015) 319–328.
- [24] OpenCV Team, Camera calibration and 3D reconstruction, Available online: http://docs.opencv.org/2.4/modules/calib3d/doc/camera_calibration_and_3D_reconstruction.html, 2019.
- [25] Z. Zhang, A flexible new technique for camera calibration, *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (2000) 1330–1334.
- [26] J. Heikkilä, O. Silvén, A four-step camera calibration procedure with implicit image correction, in: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997, pp. 1106–1112.
- [27] Y. Zhao, Z. Huang, T. Li, W. Chen, C. LeGendre, X. Ren, A. Shapiro, H. Li, Learning perspective undistortion of portraits, in: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 7849–7859.
- [28] N.P. Del Gallego, J. Iiao, M. Cordel, Blind first-order perspective distortion correction using parallel convolutional neural networks, *Sensors* 20 (17) (2020) 4898.
- [29] Z. Wang, H. Liang, X. Wu, Y. Zhao, B. Cai, C. Tao, Z. Zhang, Y. Wang, S. Li, F. Huang, S. Fu, A practical distortion correcting method from fisheye image to perspective projection image, in: *IEEE International Conference on Information and Automation*, 2015, pp. 1178–1183.
- [30] L. Wu, Q. Shang, Y. Sun, X. Bai, A self-adaptive correction method for perspective distortions of image, *Front. Comput. Sci.* 13 (2019) 588–598.
- [31] V.C. Chen, W.J. Miceli, Effect of roll, pitch, and yaw motions on ISAR imaging, *Proc. SPIE-Int. Soc. Opt. Eng.* 3807 (1999) 14–23.
- [32] S. Targ, D. Almeida, K. Lyman, Resnet in Resnet: Generalizing Residual Architectures, 2016 *arXiv preprint arXiv:1603.08029*.
- [33] P.J. Palma, Rotation vector and directional cosine matrix in problems of satellite attitude control, *Int. J. Intell. Syst. Technol. Appl.* 18 (1) (2022) 85–97.
- [34] G. Pang, C. Shen, L. Cao, A.V.D. Hengel, Deep learning for anomaly detection: a review, *ACM Comput. Surv.* 54 (2) (2021) 1–38.
- [35] Z. Zhang, Improved adam optimizer for deep neural networks, in: *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*, 2018, pp. 1–2.
- [36] W. Wang, Y. Lu, Analysis of the mean absolute error (MAE) and the root mean square error (RMSE) in assessing rounding model, in: *IOP Conference Series: Materials Science and Engineering*, vol. 324, 2018 012049.
- [37] B. Wright, A. Shakiba, R. Sharma, Z. Hameiri, Photovoltaic module array luminescence image preprocessing: Heuristic algorithms for perspective correction and cell segmentation, *European Photovoltaic Solar Energy Conference and Exhibition* (2024).
- [38] S.R. Deans, The Radon Transform and Some of Its Applications, Courier Corporation, 2007, Chapter 1, pp. 1–9.
- [39] M. Goyal, Morphological image processing, *Int. J. Comput. Sci. Technol.* 2 (4) (2011) 59.
- [40] A.C. Bergstrom, D. Conran, D.W. Messinger, Gaussian Blur and Relative Edge Response, 2023 *arXiv preprint arXiv:2301.00*.
- [41] X. Xu, S. Xu, L. Jin, E. Song, Characteristic analysis of Otsu threshold and its applications, *Pattern Recogn. Lett.* 32 (7) (2011) 956–961.