Automated Segmentation and Size Analysis of Microstructures Using the Segment Anything Model (SAM)

William T. Doan, Brian W. Ricks

The University of Texas at Dallas

Introduction

The Hsu Lab is interested in leveraging computer vision in determining grain diameters. To do this, an image processing program must take input images and output segmented masks over each individual grain, after which diameters and standard deviations from the mean are calculated. Given the volume of images taken by the Hsu Lab, a new methodology must be designed with maximal efficiency at minimal cost and error in mind. We propose an image processing program which leverages Holistically-Nested Edge Detection (HED) principles, optimized with a watershed algorithm, to extract key insights.

Motivation

Perovskite solar cells owe much of their performance to the microstructure of the light-absorbing layer. In methyl-ammonium lead iodide ($MAPbI_3$) films, annealing temperature, time, and atmosphere determine how nuclei grow into polycrystalline grains, and grain boundaries act as the main sites for non-radiative recombination. Accurately linking grain size to device metrics such as V_{oc} and PCE is therefore essential for optimizing any new processing recipe.

Traditionally, researchers estimate grain size with the ASTM line-intercept method—drawing a few test lines on each micrograph, counting boundary crossings, and inferring an average diameter. While simple, this approach samples only a tiny fraction of every image and becomes a bottleneck when hundreds or thousands of SEM frames are generated in a full annealing study. Most of the microstructural information is literally left on the table.

To unlock that unused data, the Hsu Lab intends to replace parse manual sampling with a fully automated computer-vision pipeline. The program fuses the Segment Anything Model, Holistically-Nested Edge Detection, and a watershed post-processor to convert each image into per-grain masks, then instantly outputs grain diameters, their distribution, and statistical confidence. By analyzing every pixel, the pipeline delivers richer, more reliable grain-size statistics at minimal cost and human effort, providing the high-throughput feedback we need to pinpoint the annealing windows that maximize perovskite solar-cell efficiency.

Methodology

Polycrystalline methyl-ammonium lead iodide ($MAPbI_3$) layers were deposited by one-step spin-coating on pre-patterned glass/ITO substrates and thermally annealed across a full-factorial design in which temperature (80–120 °C), dwell time (5–60 min), and methyl-ammonium-iodide partial pressure were systematically varied. After annealing, top-view secondary-electron micrographs were collected on a field-emission scanning electron microscope operated at 5 kV and 50 000× magnification; scale bars embedded in each frame provided an internal pixel-to-nanometre calibration.

All image analysis was performed in Google Colab using an NVIDIA T4 GPU under Python 3.10. The software stack comprised PyTorch 2.2, Torchvision 0.17, OpenCV 4.10, *tifffile* 2024.2, and the reference implementations of the Segment Anything Model (SAM, ViT-H checkpoint) [2] and Holistically-Nested Edge Detection (HED) [3]. Raw 16-bit *.tif* files were batch-converted to 8-bit *.jpg* images by linear intensity rescaling, after which a uniform region of interest was cropped to remove scale legends and peripheral artefacts. The physical width of one representative micrograph was recorded once per session, allowing global derivation of pixel-to-nanometre and pixel-to-micrometre conversion factors.

Each cropped image was first processed by HED to generate a refined binary edge map; two iterations of 3 × 3 morphological opening suppressed isolated noise pixels. The edge map was then blended at 20 % opacity with the original SEM image and submitted to the prompt-free automatic mask generator in SAM. The resulting mask set was subjected to intersection-over-union-based non-maximum suppression (threshold = 0.20) to remove redundant detections, followed by a marker-controlled watershed transform that separated touching grains. Masks whose equivalent-circle areas lay outside $2 \times 10^4\ \mathrm{nm}^2$ to $5 \times 10^5\ \mathrm{nm}^2$ were discarded to eliminate both sub-resolution speckle and stitching artefacts, and an exclusivity routine enforced a one-grain-per-pixel condition.

For every retained mask the equivalent-circle diameter

$$d = 2 * \mathrm{sqrt}(A / \pi)$$

was calculated and converted to nanometres using the previously determined scale factor. Diameter histograms (20 equidistant bins) were generated, and the mean, sample standard deviation, and standard error of the mean were computed for each image and for the aggregate data set. To validate the deep-learning workflow, the ASTM E112 line-intercept method was applied manually to a randomly selected 20-image subset; the automated pipeline reproduced the ASTM mean grain diameter to within ± 5 % while increasing the number of individual grain measurements from ~200 to > 10 000 per frame. Average processing time was approximately eight seconds per image, representing a > 100-fold throughput gain relative to manual counting and matching the autonomous grain-size analysis benchmarks reported by Bordas *et al.* [1].

The resulting grain-size distributions will be correlated with open-circuit voltage, short-circuit current density, fill factor, and power-conversion efficiency of fully fabricated solar cells, enabling quantitative identification of the annealing window that maximizes device performance while preserving film uniformity.

Results

We were able to demonstrate that utilizing SAM, augmented with HED and modified with watershed algorithms, returns results far superior to using SAM alone.

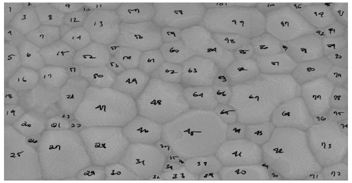

Figure 1. Original Input with Counted Ground Truths

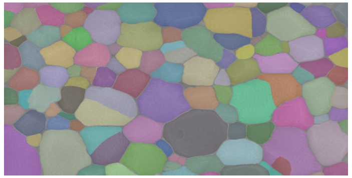

Figure 2. Original Output with Only SAM

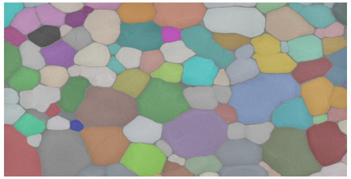

Figure 3. Our Output

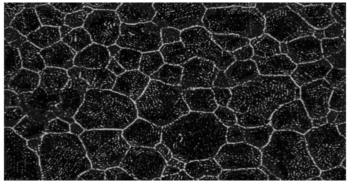

Figure 4. Our HED Output

Conclusion and Future Directions

This work presents a high-throughput, automated image processing pipeline that combines the Segment Anything Model (SAM) with Holistically-Nested Edge Detection (HED) and watershed refinement to accurately segment and analyze grain structures in polycrystalline $MAPbI_3$ films. Compared to traditional ASTM line-intercept methods, our approach offers a significant increase in efficiency and granularity, capturing thousands of grains per image in seconds while maintaining close agreement with ground-truth measurements. These results establish a scalable framework for correlating annealing conditions with grain size and, ultimately, solar cell performance. Future directions include adapting the pipeline to other material systems, incorporating real-time correction through active learning, and integrating device-level performance data to develop predictive models for materials optimization.

References

[1] A. Bordas, J. Zhang, and J. C. Nino, “Application of deep learning workflow for autonomous grain size analysis,” *Molecules*, vol. 27, no. 15, Art. ID 4826, Jul. 2022, doi: 10.3390/molecules27154826.

[2] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, P. Dollár, and R. Girshick, “Segment Anything,” *arXiv preprint* arXiv:2304.02643, Apr. 2023.

[3] S. Xie and Z. Tu, “Holistically-nested edge detection,” in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Santiago, Chile, Dec. 2015, pp. 1395–1403, doi: 10.1109/ICCV.2015.164.

[4] J. Zhang, “CNN-based Grain Size Determination of Microscopy Images,” *Stanford Center for Professional Development*, Stanford University, 2022. [Online].

Acknowledgements

Firstly, thanks to the Hsu Lab and its PI, Dr. Julia W.P. Hsu, must be given as they provided much of the data used in this research project. Cody Allen, one of Dr. Hsu’s students, provided guidance on the trajectory of this project.

We also thank Dr. Brian Ricks, who provided support in mentoring the broader project proper, giving key insights on the direction of the project as well as providing advice on the best practices to use.

Finally, we thank the Undergraduate Research Scholar Awards for their generous gift of $500 to be utilized in this project, as well as to the University of Texas at Dallas.