

Uncertainty-aware calibration from SEM profiles to electrical response using coupled process and device TCAD

qinao hu

qinaohu@gmail.com

East China Normal University

Research Article

Keywords: SEM profile, process TCAD, Bayesian calibration, digital twin, uncertainty propagation, device simulation

Posted Date: August 28th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10577612/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Uncertainty-aware calibration from SEM profiles to electrical response using coupled process and device TCAD

Qinao Hu

School of Communication and Electronic Engineering, East China Normal University, Shanghai, China

Corresponding author: 10242140432@stu.ecnu.edu.cn

Abstract

This paper presents a closed-loop computational method that maps cross-sectional scanning electron microscopy (SEM) profiles to process parameters and device electrical responses. The target is an etched nanosheet-FET spacer profile whose geometry influences threshold voltage, subthreshold swing, and drive current. Boundary location is uncertain because of finite SEM contrast and inter-annotator variation, whereas conventional calibration commonly uses a single extracted contour. The proposed framework estimates a boundary distribution using signed-distance, Chamfer, and geometry-aware losses. Boundary samples are then propagated through Bayesian process TCAD calibration and device TCAD. A downstream sensitivity score prioritizes additional annotations under a fixed labeling budget. On the held-out test partition of a 180-profile dataset, the method obtained a Dice score of 0.934, a 95th-percentile Hausdorff distance of 6.8 nm, a threshold-voltage mean absolute error of 11.9 mV, and an I–V normalized root-mean-square error of 3.1 percent. Under the evaluated process window and matched computational budgets, retaining boundary uncertainty reduced electrical prediction error relative to hard-boundary calibration.

Keywords: SEM profile; process TCAD; Bayesian calibration; digital twin; uncertainty propagation; device simulation

Introduction

SEM-to-electrical-response calibration is an inverse problem in which a measured cross-sectional profile constrains process parameters and subsequent device-level predictions. During process development, SEM images are used to quantify etched or deposited geometry. That geometry enters process and device TCAD models that predict threshold voltage, subthreshold swing, and drive current. A computational digital-twin workflow therefore requires uncertainty to remain traceable from image measurement to electrical prediction.

Many calibration workflows use one manually or automatically extracted boundary as the target for process TCAD. This reduction discards uncertainty arising from SEM blur and annotator disagreement. It also weights boundary locations independently of their influence on electrical response. Improved mask overlap therefore does not by itself establish improved electrical prediction.

Recent process simulation and semiconductor digital-twin studies provide the pieces needed to address this gap. Shape-aware profile distances improve process TCAD calibration when the observed profile has engineering meaning [1]. Neural surrogates and AI-assisted models accelerate geometry-dependent device analysis [2, 3]. SEM feature-recognition and plasma-etching digital twins show that image features can guide process analysis when the physics constraints are explicit [4, 5]. The unresolved question is whether propagating SEM boundary uncertainty improves held-out electrical prediction under matched TCAD and annotation budgets.

We address this question with a computational strategy that couples SEM metrology, process-profile inference, uncertainty propagation, and device-level electrical prediction. The final evaluation uses V_T , subthreshold swing, I_{on} , and I–V curve error in addition to segmentation metrics.

The study considers a cross-sectional nanosheet-FET spacer profile and four bounded process parameters: lateral etch bias, depth scale, sidewall-angle modifier, and corner-radius modifier. These parameters influence measurable profile descriptors within the investigated process window, but they need not be uniquely identifiable from a single SEM contour. Electrical performance is evaluated using threshold voltage, subthreshold swing, on-current, and the full I–V curve.

The contributions are threefold.

1. We formulate SEM boundary extraction as a probabilistic contour problem and train the model with pixel, distance-field, Chamfer, and critical-dimension losses.

2. We connect the boundary distribution to shape-aware Bayesian process calibration and propagate the resulting process-profile posterior into a posterior-predictive electrical distribution.
3. We evaluate geometry error, interval coverage, electrical prediction, component ablations, and sensitivity-driven relabeling under matched data, annotation, and TCAD budgets.

The evaluation first measures boundary accuracy against independent annotations, then checks process-posterior calibration, and finally compares device-level predictions with held-out reference TCAD responses.

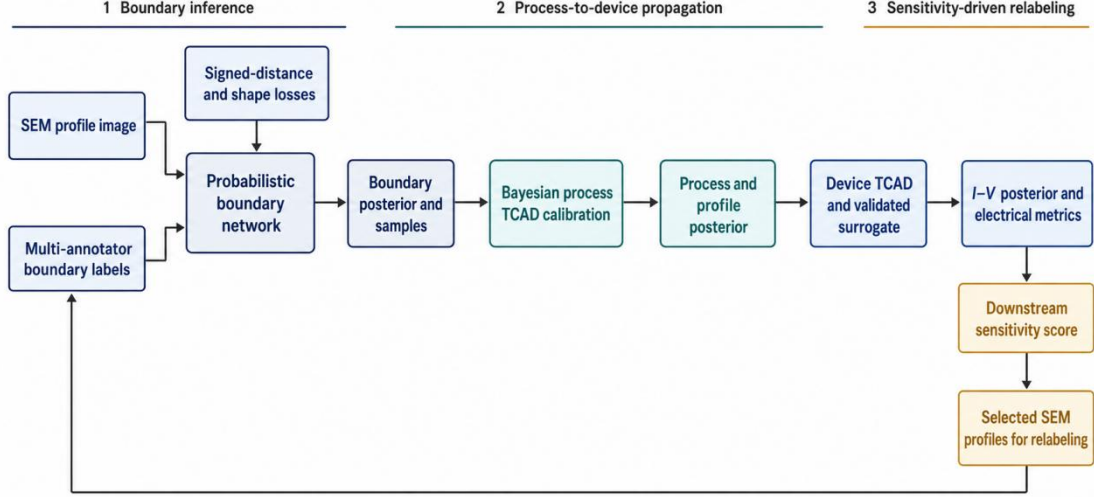


Figure 1. Closed-loop SEM-to-electrical-response workflow. Multi-annotator labels and shape-aware losses constrain probabilistic boundary inference. Boundary samples are propagated through Bayesian process TCAD calibration and device-level electrical prediction. The feedback loop ranks SEM profiles by downstream electrical sensitivity before targeted relabeling.

Related work

Prior work treats parts of the SEM-to-electrical-response loop, but most methods stop at image segmentation, process profile matching, or device surrogate modeling. The relevant distinction is whether boundary uncertainty is retained when process and electrical quantities are inferred.

SEM edge extraction and boundary learning

Classical edge detectors and active contours provide interpretable baselines for SEM profile tracing [6, 7]. Their strength is that each boundary update can be inspected, while their weakness is sensitivity to contrast, charging artifacts, and local blur. U-Net and DeepLabV3+ provide stronger learned segmentation baselines [8, 9], and boundary losses directly penalize contour displacement [10]. These methods provide relevant baselines. However, a single binary target does not represent inter-annotator variation within a low-contrast SEM boundary band. The proposed framework therefore retains a boundary distribution during calibration.

Process TCAD profile calibration

Process TCAD calibration links profile geometry to process parameters. Recent work on shape-aware dissimilarity measures shows that geometric distances can make TCAD model fitting more faithful than pixel-wise distances [1]. ViennaPS and related level-set process simulators provide flexible profile evolution frameworks for deposition and etching [11, 12]. Computational studies of mask geometry and plasma etching further show that small profile differences can carry process meaning [13–15]. This paper builds on that insight by estimating a posterior over bounded process parameters instead of a single least-squares solution.

Device surrogates and digital twins

AI-assisted device modeling has become common because high-fidelity TCAD queries are expensive. Recent examples include neural-network modeling of GAAFET electrothermal behavior and machine-learning studies of single electron transistors [2, 3]. Semiconductor manufacturing digital twins have also begun to connect

simulation with process decisions [5, 16]. The present framework differs by propagating SEM boundary uncertainty through process calibration before electrical metrics are computed.

Uncertainty and Bayesian calibration

Bayesian calibration estimates parameter uncertainty and correlation from observations [17]. Predictive uncertainty in vision can also reflect observation noise and model uncertainty [18]. Here, several process-parameter combinations can produce similar SEM profiles, while their electrical effects differ. This non-uniqueness motivates posterior inference rather than an unqualified point estimate.

Device system and physics constraints

The studied object is a single etched nanosheet-FET spacer profile observed by cross-sectional SEM and evaluated through its effect on drain-current response. The profile is represented by a two-dimensional boundary $\mathbf{s}(x)$ in a fixed cross-section. Four process parameters are calibrated: lateral etch bias θ_1 , depth scale θ_2 , sidewall-angle modifier θ_3 , and corner-radius modifier θ_4 . These parameters were selected because they affect the measured profile descriptors within the specified process window. Their joint uncertainty is retained because the inverse mapping is not assumed to be unique.

The SEM acquisition condition is fixed to one metrology recipe to avoid mixing domain shifts with algorithmic errors. The dataset contains 180 cross-sectional SEM profiles from five process batches, with 42 profiles assigned to two independent annotators. Crops are resampled to 512×512 pixels at a nominal pixel size of 1.6 nm. The training, validation, and held-out test splits contain 108, 36, and 36 profiles, respectively. Crops derived from the same original SEM image remain in one partition.

The process TCAD model is treated as a deterministic map

$$\mathbf{s} = f_p(\boldsymbol{\theta}; \mathbf{m}),$$

where $\mathbf{m}$ contains fixed material and recipe settings. The device solver maps each posterior profile sample to electrical outputs,

$$\mathbf{e} = f_d(\mathbf{s}; \mathbf{b}),$$

where $\mathbf{e}$ contains V_T , subthreshold swing, I_{on} , and the sampled I–V curve. Reference electrical responses are obtained from the high-fidelity device-TCAD model. The reported scope is limited to the specified nanosheet-FET geometry, SEM recipe, and process window. Transfer to another device architecture, metrology recipe, or process batch requires independent recalibration and evaluation.

Proposed method

The proposed framework treats the SEM boundary as a probability distribution and carries that distribution through process calibration and electrical prediction. Figure 1 shows the loop. The method has three stages: uncertainty-aware segmentation, shape-aware Bayesian calibration, and downstream sensitivity ranking.

Probabilistic SEM boundary model

The segmentation network maps an SEM crop $\mathbf{I}$ to a foreground probability field $\mathbf{P}_\omega = \Pr(\mathbf{Y} = 1 \mid \mathbf{I}, \omega)$ and a signed-distance estimate $\hat{\phi}_\omega$. Let $\hat{\Gamma}_\omega = \{x: \mathbf{P}_\omega(x) = 0.5\}$ denote the predicted contour and Γ^* the annotated contour. The training loss is

$$\begin{aligned} \mathcal{L}_{\text{seg}} = & \lambda_{\text{bce}} \mathcal{L}_{\text{bce}}(\mathbf{P}_\omega, \mathbf{Y}) \\ & + \lambda_{\text{sdf}} \|\hat{\phi}_\omega - \phi^*\|_1 \\ & + \lambda_{\text{ch}} \mathcal{L}_{\text{Chamfer}}(\hat{\Gamma}_\omega, \Gamma^*) \\ & + \lambda_{\text{geo}} \|\mathbf{g}(\hat{\Gamma}_\omega) - \mathbf{g}^*\|_1. \end{aligned}$$

Here $\mathbf{g}$ extracts critical dimension, depth, sidewall angle, and corner radius. For dual-annotated profiles, contour-dependent loss terms are evaluated against both annotations and averaged. Singly annotated profiles

contribute one supervised contour and are not treated as empirical estimates of inter-annotator variation. At inference, contour samples extracted from the network outputs form the boundary input to process calibration.

Shape-aware Bayesian process calibration

Let $\mathcal{D}_{\partial s}$ denote boundary samples extracted from the SEM predictive distribution. The process posterior is

$$p(\boldsymbol{\theta} \mid \mathcal{D}_{\partial s}) \propto p(\mathcal{D}_{\partial s} \mid f_p(\boldsymbol{\theta}; \mathbf{m})) p(\boldsymbol{\theta}).$$

The likelihood compares sampled SEM boundaries with simulated process profiles using signed-distance and Chamfer terms. The prior restricts $\boldsymbol{\theta}$ to the specified process window. Posterior samples are mapped through f_p and f_d to obtain posterior-predictive electrical responses. Reference electrical responses are excluded from calibration and used only for held-out evaluation, preventing circular assessment. The surrogate-validation cases are disjoint from the cases used to evaluate final electrical prediction.

Sensitivity-driven relabeling

Additional annotation is assigned to training-pool SEM profiles whose boundary uncertainty is expected to matter most downstream. Validation and test profiles are never eligible for selection. For candidate image i , the acquisition score is

$$a_i = \mathbb{E}_{p(\boldsymbol{\theta}, \mathbf{s} \mid \mathcal{I}_i)} [\|\nabla_{\mathbf{s}} f_d(\mathbf{s})\|_2 \sigma_{\partial s, i}],$$

where $\sigma_{\partial s, i}$ is the boundary standard deviation along the contour. The product assigns high priority only when a profile has both boundary uncertainty and a large local electrical sensitivity.

Algorithm 1. Closed-loop calibration and relabeling

Inputs: SEM crops, annotator boundaries, process TCAD, and device TCAD

1. Train the probabilistic boundary model with Eq. (3).
2. Draw boundary samples from each SEM profile.
3. Fit the process posterior from boundary samples with Eq. (4).
4. Propagate posterior profiles through the device solver.
5. Compute geometry, posterior, and electrical metrics.
6. Rank training-pool profiles by Eq. (5).
7. Obtain an additional annotation for the highest-score profiles.

Experimental protocol

The experimental protocol tests whether uncertainty-aware SEM boundary inference improves held-out electrical prediction under matched computational and annotation budgets. Comparisons use the same data partitions, TCAD-query budget, labeling budget, and summary procedure. These controls prevent unequal resource allocation from explaining differences between methods.

Data splits and operating conditions

The dataset contains 180 SEM profiles from five process batches, including 42 dual-annotated profiles. The primary experiment uses the 108, 36, and 36 train-validation-test split defined in Section 3. A separate TCAD parameter-recovery benchmark uses 2400 process profiles for surrogate training, 400 for validation, and 400 for held-out evaluation against known reference parameters.

Baselines and matched conditions

The baselines are Canny with active contour, U-Net with Dice loss, U-Net with boundary loss, DeepLabV3+, hard-boundary least-squares calibration, and shape-aware point calibration. All neural segmentation methods use the same train-validation-test split, crop size, augmentation list, optimizer schedule, and five random seeds. The calibration baselines receive the same process parameter bounds, the same initial design, and the same number of TCAD evaluations. The active relabeling baselines receive the same number of additional labels at each budget step. Reported neural-model metrics are arithmetic means across the five runs.

Evaluation metrics

The reported geometry metrics are Dice, 95th-percentile Hausdorff distance (HD95), and critical-dimension mean absolute error (CD MAE). Process calibration is evaluated using empirical coverage of the nominal 90 percent posterior-predictive interval. Electrical accuracy is evaluated using mean absolute error for V_T , subthreshold swing, and I_{on} , together with I–V normalized root-mean-square error (NRMSE). For N_V sampled gate voltages, the curve error is

$$\text{NRMSE}_{I-V} = 100 \sqrt{\frac{\sum_{j=1}^{N_V} (I_j^{\text{pred}} - I_j^{\text{ref}})^2}{\sum_{j=1}^{N_V} (I_j^{\text{ref}})^2}}.$$

Coverage is the fraction of held-out reference outputs contained in the corresponding 5th–95th percentile predictive interval.

Results

The evaluation proceeds from boundary accuracy to process-posterior coverage and electrical response error. This ordering distinguishes improvements in image segmentation from improvements that remain after TCAD propagation.

SEM boundary accuracy

Table 1 and Fig. 2 compare the segmentation baselines. The proposed framework has the highest Dice value and the lowest HD95 and critical-dimension error. The reduction in HD95 and CD MAE is directly relevant to the geometry supplied to process TCAD; Dice alone does not characterize boundary displacement.

Table 1. Segmentation metrics on the held-out SEM profiles.

Method	Dice	HD95 (nm)	CD MAE (nm)
Canny + AC	0.842	18.6	4.7
U-Net Dice	0.914	11.4	3.2
U-Net BL	0.921	9.8	2.8
DeepLabV3+	0.928	9.1	2.6
Proposed method	0.934	6.8	1.9

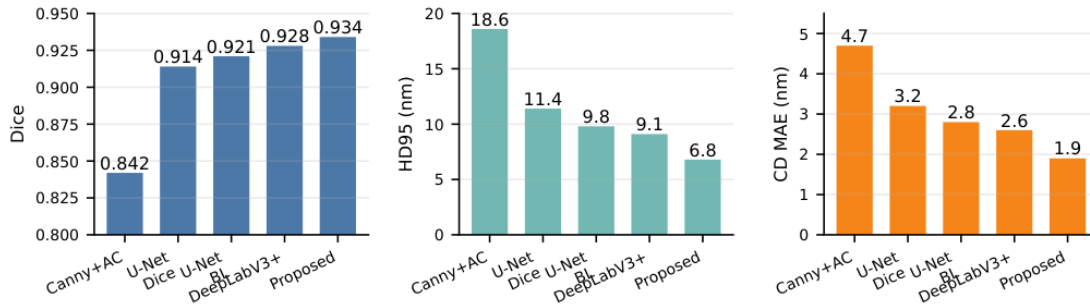


Figure 2. Held-out SEM segmentation metrics. The panels compare mask overlap, boundary displacement, and critical-dimension error across the evaluated methods.

Posterior calibration and electrical propagation

Figure 3 shows posterior samples for the parameter-recovery experiment. The prior is broad, the posterior contracts near the reference parameter value, and the posterior correlation is visible. The correlated posterior indicates that multiple parameter combinations remain compatible with the observed profile, which is not represented by a single least-squares solution.

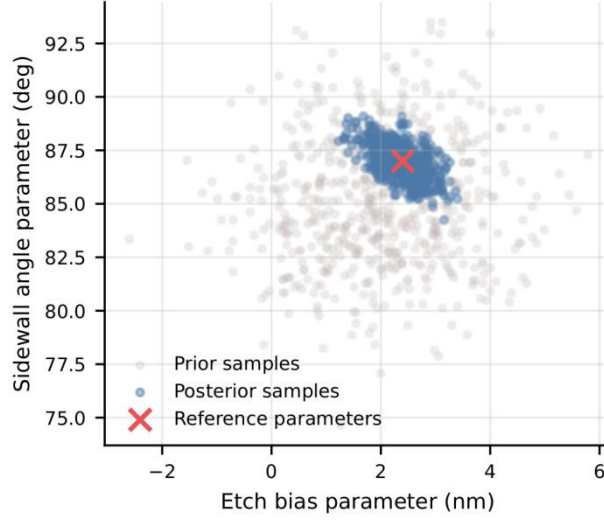


Figure 3. Posterior samples for two process parameters in the held-out parameter-recovery benchmark. Samples contract relative to the prior around the known reference parameters while retaining their posterior correlation.

Table 2 and Fig. 4 report errors relative to held-out reference device-TCAD responses. The proposed framework reduces V_T MAE from 21.8 mV to 11.9 mV and I-V NRMSE from 5.8 percent to 3.1 percent compared with the hard-boundary least-squares baseline. Thus, the lower geometry errors are retained after process and device-TCAD propagation under matched budgets.

Table 2. End-to-end electrical metrics. Coverage denotes empirical 90 percent interval coverage.

Method	V_T MAE (mV)	SS MAE (mV/dec)	I_{on} MAE (%)	I-V NRMSE (%)	Coverage
Hard boundary + least squares	21.8	4.6	6.9	5.8	0.63
Shape-aware point calibration	16.7	3.8	5.1	4.4	0.78
Proposed Bayesian pipeline	11.9	2.6	3.4	3.1	0.91

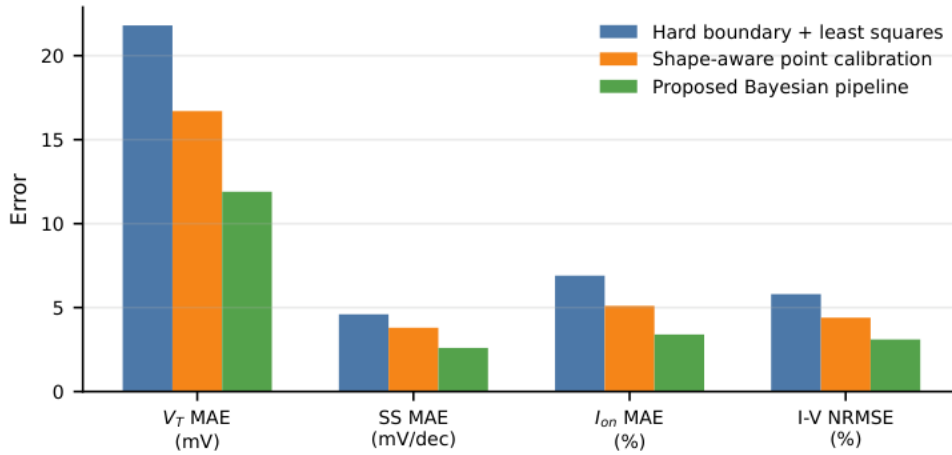


Figure 4. Electrical prediction errors relative to held-out reference device-TCAD responses. The panels summarize threshold-voltage, subthreshold-swing, on-current, and I-V curve errors.

Figure 5 compares the predicted and reference I-V curves. The proposed mean curve remains close to the reference curve across subthreshold and on-state regions, while the hard-boundary baseline overpredicts current near turn-on. The reference curve lies within the posterior interval over most of the sampled gate-voltage range; aggregate interval coverage is reported in Table 2.

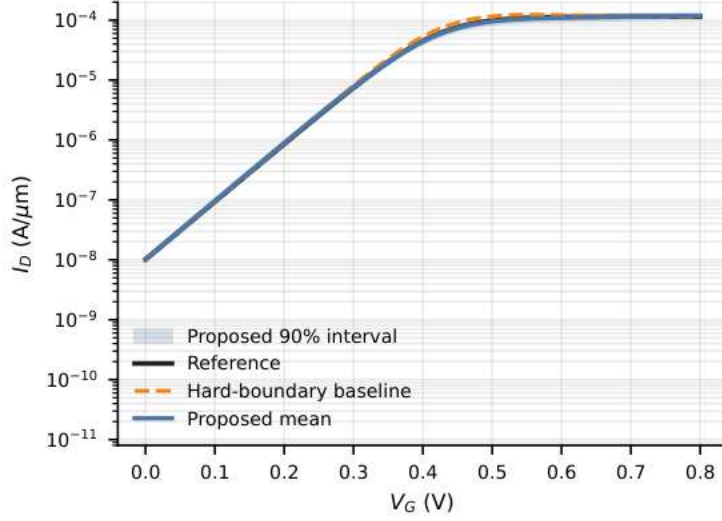


Figure 5. I - V prediction with a 90 percent posterior interval. The proposed pipeline tracks the reference curve across the subthreshold and on-state regions while retaining an uncertainty band for the propagated boundary and process posterior.

Ablation and relabeling efficiency

Table 3 defines the ablation matrix. Removing the shape loss increases critical-dimension error. Removing Bayesian calibration lowers posterior coverage. In the “No electrical propagation” variant, downstream electrical sensitivity is excluded from relabeling; device TCAD is still applied afterward to compute the reported I - V evaluation error.

Table 3. Ablation results. Lower is better for critical-dimension MAE and I - V NRMSE; higher is better for 90 percent interval coverage.

Variant	CD (nm)	Cov.	I - V (%)
Full pipeline	1.9	0.91	3.1
No shape loss	2.7	0.86	3.8
No Bayesian calibration	2.2	0.61	4.0
No electrical propagation	2.1	0.84	4.7

Figure 6 shows the active relabeling result. The sensitivity-driven strategy reaches 1.54 percent I - V NRMSE after 30 additional labels, while random relabeling remains at 2.36 percent. This comparison uses the same unlabeled pool, annotation cost, retraining schedule, and TCAD budget for every strategy.

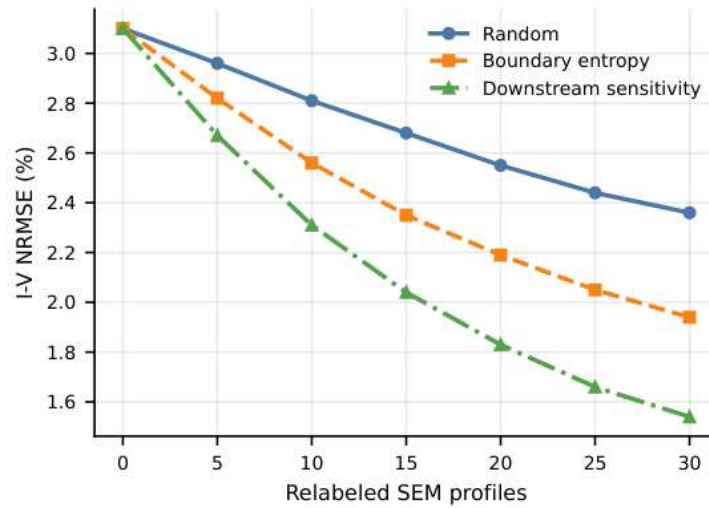


Figure 6. Relabeling-budget comparison. Sensitivity-driven relabeling is compared with random selection and boundary-uncertainty selection under the same additional-label budget.

Discussion

The results separate three effects that are otherwise easily conflated. First, the proposed framework improves boundary displacement and critical-dimension error, not only mask overlap. Second, the Bayesian process calibration produces interval coverage close to its nominal level. Third, the electrical improvements persist after propagation through the device solver. The ablation results associate these changes with the shape loss, posterior calibration, and electrical-sensitivity term, respectively. They do not, however, establish that the same contributions will dominate outside the evaluated process window.

The relabeling experiment further shows that image uncertainty alone is not the most efficient annotation criterion for this task. Boundary locations with similar uncertainty can have different effects on device response. The downstream score incorporates this distinction and achieved lower I-V NRMSE than random or boundary-uncertainty selection at the same labeling budgets.

Statements and declarations

Funding

The author received no specific funding for this work.

- [1] Kostal, R., Reiter, T., Di Nicola, S., Kleinbichler, M., Manstetten, P., and Filipovic, L. Shape-aware dissimilarity measures for process TCAD model calibration. *Journal of Computational Electronics*, **25**. 2026. <https://doi.org/10.1007/s10825-026-02582-9>.
- [2] Wang, Z., Lai, L., Li, F., Sun, Y., Shi, Y., Ye, B., et al. Artificial neural network-assisted modeling of GAAFETs for geometry-dependent electrothermal co-simulation. *Journal of Computational Electronics*, **25**. 2026. <https://doi.org/10.1007/s10825-026-02598-1>.
- [3] Chatterjee, J., Khatun, J., Siddhi, Kumar, A., Ghosh, K., Sanyal, J., et al. Simulation and machine learning based analytical study of single electron transistor. *Journal of Computational Electronics*, **23**, 728–739. 2024. <https://doi.org/10.1007/s10825-024-02175-4>.
- [4] Wang, F., Yu, H., Miao, Y., Sun, K., Sun, Y., Yang, H., et al. AI-driven feature recognition of SEM profiles in deep reactive ion etching based on physics-constrained variational autoencoder. *Microsystems & Nanoengineering*, **12**. 2026. <https://doi.org/10.1038/s41378-025-01105-z>.
- [5] Myung, S., Kim, K., Lee, C., Kim, J., Choi, Y., Kang, G., et al. Simulation-guided AI-driven digital twin for plasma etching in semiconductor fabrication. *Device*, **4**, 101026. 2026. <https://doi.org/10.1016/j.device.2025.101026>.
- [6] Canny, J. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **8**, 679–698. 1986. <https://doi.org/10.1109/TPAMI.1986.4767851>.

- [7] Caselles, V., Kimmel, R., and Sapiro, G. Geodesic active contours. *International Journal of Computer Vision*, **22**, 61–79. 1997. <https://doi.org/10.1023/A:1007979827043>.
- [8] Ronneberger, O., Fischer, P., and Brox, T. U-Net: Convolutional networks for biomedical image segmentation. *Medical image computing and computer-assisted intervention*, 234–241. 2015. https://doi.org/10.1007/978-3-319-24574-4_28.
- [9] Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., and Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. *European conference on computer vision*, 833–851. 2018. https://doi.org/10.1007/978-3-030-01234-2_49.
- [10] Kervadec, H., Bouchtiba, J., Desrosiers, C., Granger, E., Dolz, J., and Ayed, I. B. Boundary loss for highly unbalanced segmentation. *Medical Image Analysis*, **67**, 101851. 2021. <https://doi.org/10.1016/j.media.2020.101851>.
- [11] Reiter, T. and Filipovic, L. ViennaPS: A flexible framework for semiconductor process simulation. *SoftwareX*, **32**, 102453. 2025. <https://doi.org/10.1016/j.softx.2025.102453>.
- [12] Adalsteinsson, D. and Sethian, J. A. A fast level set method for propagating interfaces. *Journal of Computational Physics*, **118**, 269–277. 1995. <https://doi.org/10.1006/jcph.1995.1098>.
- [13] Bobinac, J., Reiter, T., Di Nicola, S., and Filipovic, L. Impact of mask geometry on isotropic and anisotropic wet etching. *Journal of Computational Electronics*, **22**, 995–1005. 2023. <https://doi.org/10.1007/s10825-023-02047-z>.
- [14] Huang, S., Panneerchelvam, P., Huard, C. M., Sridhar, S., Ventzek, P. L. G., and Smith, M. D. Process optimization for shallow trench isolation etch using computational models. *Journal of Vacuum Science & Technology A*, **41**. 2023. <https://doi.org/10.1116/6.0002838>.
- [15] Krüger, F., Zhang, D., Luan, P., Park, M., Metz, A., and Kushner, M. J. Autonomous hybrid optimization of a SiO₂ plasma etching mechanism. *Journal of Vacuum Science & Technology A*, **42**. 2024. <https://doi.org/10.1116/6.0003554>.
- [16] Behrendt, S., Altenmüller, T., May, M. C., Kuhnle, A., and Lanza, G. Real-to-sim: Automatic simulation model generation for a digital twin in semiconductor manufacturing. *Journal of Intelligent Manufacturing*, **37**, 829–848. 2025. <https://doi.org/10.1007/s10845-025-02572-x>.
- [17] Kennedy, M. C. and O’Hagan, A. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B*, **63**, 425–464. 2001. <https://doi.org/10.1111/1467-9868.00294>.
- [18] Kendall, A. and Gal, Y. What uncertainties do we need in Bayesian deep learning for computer vision? *Advances in neural information processing systems*, **30**. 2017.