SCIENCE CHINA
Information Sciences

• **Supplementary File** •

Litho-AsymVnet: Super-Resolution Lithography Modeling with an Asymmetric V-net Architecture

Qing Zhang[1,†], Yuhang Zhang[1,†], Wei Lu[1], Huajie Huang[1], Zheng Zhong[2], Congshu Zhou[2] & Yongfu Li[1,*]

1. Department of Micro-Nano Electronics and MoE Key Lab of Artificial Intelligence, Shanghai Jiao Tong University, Shanghai 200240, China
2. Primarius Technologies Co., Ltd., China

Appendix A Experimental Results

Appendix A.1 Evaluation Metrics

We quantify and evaluate the lithography model's performance using the following metrics [1].
Pixel Accuracy (Pixel Acc.): The ratio of correctly predicted pixels to total pixels in an image.

$$Pixel\ Acc. = \frac{TN\# + TP\#}{TN\# + TP\# + FP\# + FN\#}, \tag{A1}$$

where $TN\#$, $FN\#$, $TP\#$, and $FP\#$ are the number of true negative, false negative, true positive, and false positive pixels, respectively.
Class Accuracy (Class Acc.): The average ratio of correctly predicted and actual pixels for each class in the image.

$$Class\ Acc. = \frac{1}{2}(\frac{TN\#}{TN\# + FP\#} + \frac{TP\#}{TP\# + FN\#}). \tag{A2}$$

Mean Intersection Over Union (Mean IOU): The average of IOU of all classes, where IOU for each class is the ratio of the pixels in both true and predicted images and pixels in either of them.

$$Mean\ IOU = \frac{1}{2}(\frac{TN\#}{TN\# + FP\# + FN\#} + \frac{TP\#}{TN\# + FP\# + FN\#}). \tag{A3}$$

Appendix A.2 Validating the Effectiveness of Trimming Method

To validate the effectiveness of the "trimming" method, various target dimensions (h) from 224 to 1344 pixels (224, 448, 672, 896, 1120, and 1344) are explored to determine the best trade-off between accuracy and speed. Table A1 provides detailed experimental metrics based on validation datasets. Figure A1 illustrates the trendlines and bar chart of the experimental results. The worst performance is the baseline model (h = 1344 pixel), in which the validation *Pixel Acc.*, *Class Acc.*, and *Mean IOU* are 99.3%, 99.1%, and 98.0%, respectively. When the dimension decreases to 1120, the validation *Pixel Acc.*, *Class Acc.*, and *Mean IOU* are improved by 0.2%, 0.2%, and 0.6%, respectively, and the runtime is 1.2× that of 1344. When the target dimension further decreases, the prediction accuracy remains the same, while the runtime continues to increase and reaches 21.3× slower when using the target dimension of 224. Thus, we have chosen a target dimension, h of 1120, which enables us to achieve a high prediction accuracy with acceptable runtime overhead.

Appendix A.3 Effectiveness of Litho-AsymVnet Model

To verify the effectiveness of the Litho-AsymVnet framework, we have conducted experiments to benchmark with the popular network architectures, including SegNet [2], fully convolutional networks (FCN) [3], and U-net [4] based on the same dataset. The experimental results are summarized in Table A2 and the distribution of the prediction errors (pixel error, class error, and IOU error) are illustrated in Figure A2. Note that the Litho-AsymVnet model takes in a lower resolution dimension of 224 × 224 and produces a higher resolution contour resist image with a dimension of 1344 × 1344, while the input and output dimensions of remaining networks are 224 × 224. Among the three popular networks (U-Net, FCN, and SegNet), U-Net achieves relatively high performance with *Pixel Acc.*, *Class Acc.*, and *Mean IOU* of 99.0%, 98.7%, and 97.3%, respectively. Compared to U-net, Litho-AsymVnet has further enhanced the performance with 0.3%, 0.4%, and 0.7% improvement in *Pixel Acc.*, *Class Acc.*, and *Mean IOU*, respectively. Overall, the Litho-AsymVnet model outperforms other popular network architectures with the highest *Pixel Acc.*, *Class Acc.*, and *Mean IOU* of 99.3%, 99.1%, and 98.0%, respectively.

* Corresponding author (email: yongfu.li@sjtu.edu.cn)
† Qing Zhang and Yuhang Zhang have contributed equally to this work.

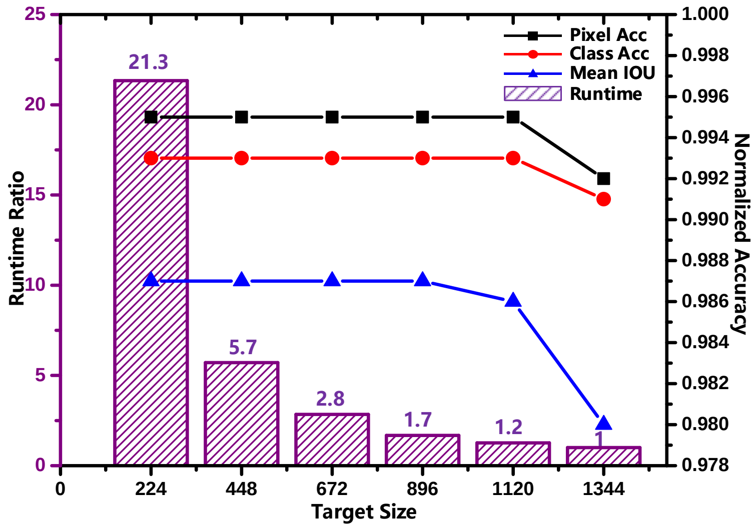

Figure A1 Trimming method performance analysis with different target dimension (h) on validation dataset.

Table A1 Performance comparison with different target dimensions on validation dataset

h	*Pixel Acc.* (%)	*Class Acc.* (%)	*Mean IOU* (%)	Runtime (s)	Normalized
224	**99.5**	**99.3**	**98.7**	131.4	21.3 ×
448	**99.5**	**99.3**	**98.7**	35.1	5.7 ×
672	**99.5**	**99.3**	**98.7**	17.5	2.8 ×
896	**99.5**	**99.3**	**98.7**	10.4	1.7 ×
1120	**99.5**	**99.3**	98.6	7.8	1.2 ×
1344	99.3	99.1	98.0	**6.2**	1 ×

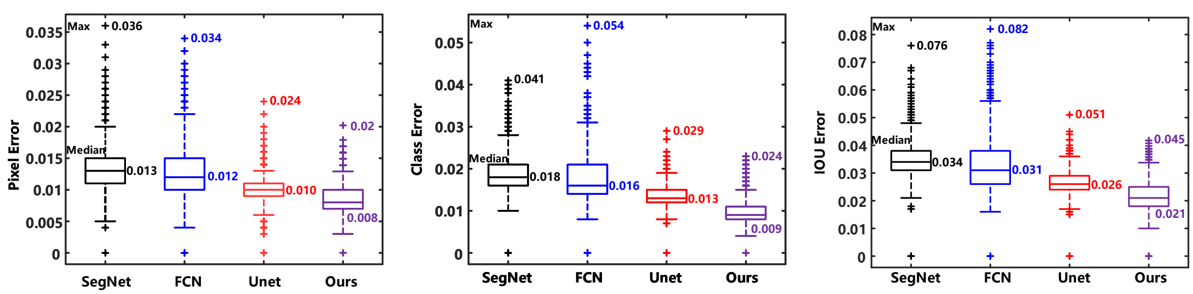

Figure A2 Experimental comparison between SegNet, FCN, Unet, and Litho-AsymVnet.

Table A2 Performance comparison with different models

Model	*Pixel Acc.* (%)	*Class Acc.* (%)	*Mean IOU* (%)
SegNet [2]	98.5	98.0	96.2
FCN [3]	98.7	98.2	96.7
U-net [4]	99.0	98.7	97.3
Litho-AsymVnet	**99.3**	**99.1**	**98.0**

Appendix A.4 Comparison with the State-of-the-art Work

LithoGAN [1] has reported a runtime reduction of 1800× and 190× compared with the rigorous simulation and other prior work [5], respectively. However, it is much more challenging to achieve high performance using a large area ratio. To verify our observation, we implement and evaluate the LithoGAN model with the two datasets (2D polysilicon and contact layers). **Poly/Metal Layer**

Table A3 Performance comparison with the state-of-the-art work [1]

Dataset	Model	h (nm)	*Pixel Acc.* (%)	*Class Acc.* (%)	*Mean IOU* (%)
Polysilicon 2D Lines	LithoGAN [1]	1344	98.3	97.7	95.5
	Litho-AsymVnet w/o trimming	1344	99.3	99.1	98.0
	Litho-AsymVnet with trimming	1120	**99.4**	**99.3**	**98.5**
Contact	LithoGAN [1]	672	99.5	94.5	90.1
	Litho-AsymVnet w/o trimming	672	99.1	98.2	97.1
	Litho-AsymVnet with trimming	560	**99.8**	**98.6**	**97.6**

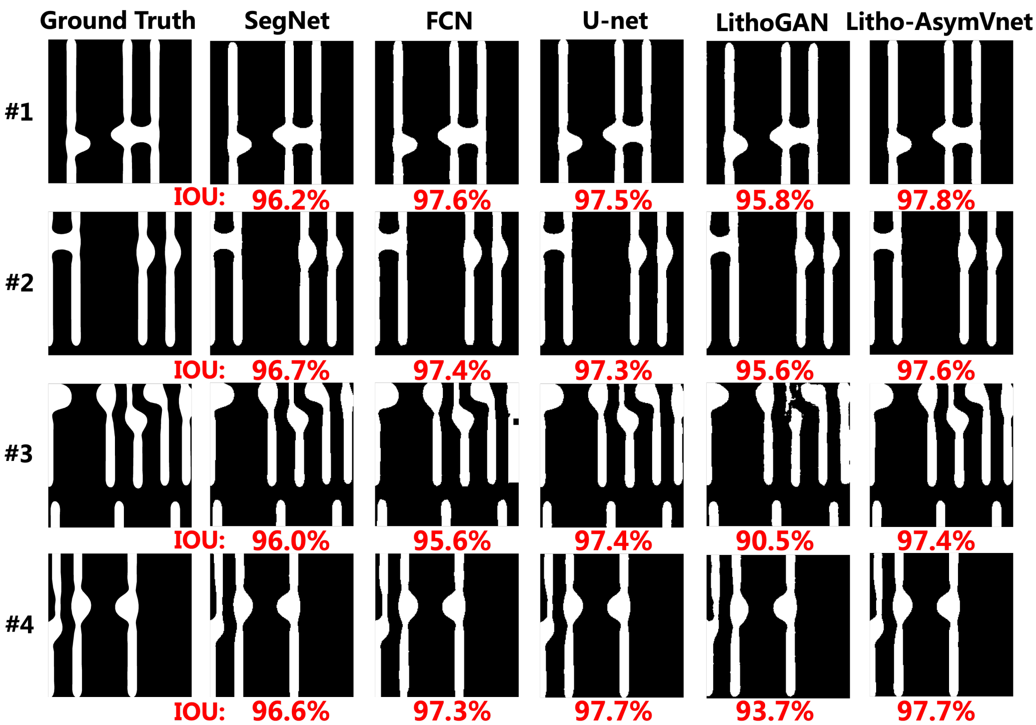

Figure A3 Visualization results of different lithography models, including SegNet, FCN, U-net, LithoGAN, and Litho-AsymVnet.

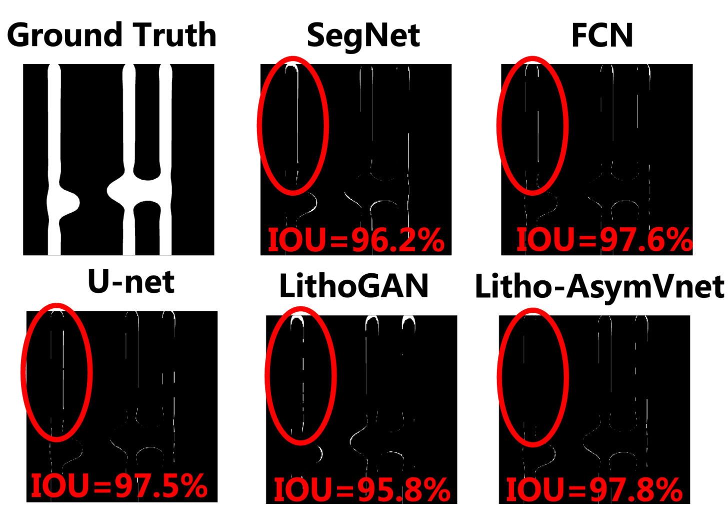

Figure A4 XNOR visualization results of different lithography models, including SegNet, FCN, U-net, LithoGAN, and Litho-AsymVnet.

Benchmark: Table A2 summarizes the comparison results. Compared to LithoGAN [1], our Litho-AsymVnet model with the "trimming" method has achieved 1.1%, 1.6%, and 2.0% improvement in *Pixel Acc.*, *Class Acc.*, and *Mean IOU*, respectively. **Contact Layer Benchmark**: Compared to LithoGAN [1], the Litho-AsymVnet model with "trimming" method has achieved 0.3%, 4.1%, and 7.5% improvement in *Pixel Acc.*, *Class Acc.*, and *Mean IOU*, respectively (See Table A2).

Appendix A.5 Visualization Comparison of Different Lithography Models

To further illustrate the effectiveness of our approach, we have provided several data visualization examples from simulation results of different models, including SegNet [2], FCN [3], U-net [4], LithoGAN [1], and Litho-AsymVnet for comparison. As illustrated in Figure A3, the Litho-AsymVnet model results in smoother contour shapes than the others while achieving the highest simulation accuracy. Besides, we have compared the differences between the simulation results and the ground truth by performing XNOR Boolean operations. The XNOR results are illustrated in Figure A4, where white represents the errors. We observe that the Litho-AsymVnet model achieves higher accuracy and fewer errors than other models.

References

1 Wei Ye, et al. LithoGAN: End-to-end lithography modeling with generative adversarial networks. In: Proceedings of ACM/IEEE Design Automation Conference (DAC), 2019, pages 1–6.
2 Vijay Badrinarayanan, et al. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2017, 39:12, pages 2481–2495.
3 Jonathan Long, et al. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pages 3431–3440.
4 Olaf Ronneberger, et al. U-net: Convolutional networks for biomedical image segmentation. In: Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015, pages 234–241.
5 Yibo Lin. Deep Learning for Mask Synthesis and Verification: A Survey. In: Proceedings of Asia and South Pacific Design Automation Conference (ASP-DAC), 2021, pages 825–832.