

Data Efficient Lithography Modeling with Transfer Learning and Active Data Selection

Yibo Lin¹, Meng Li¹, Yuki Watanabe², Taiki Kimura², Tetsuaki Matsunawa², Shigeki Nojima², David Z. Pan¹

¹University of Texas at Austin, ²Toshiba Memory Corporation

Data Efficient Lithography Modeling

Optical Source
Mask
Contact Mask
Aerial Image (Light intensity map)
Resist Pattern

Deep Neural Networks for Lithography Modeling

Neural networks are getting **deeper** for higher accuracy
AlexNet-8, VGG-19, ResNet-101, ResNet-1202

"You just keep on adding layers, until the test error doesn't improve anymore." — Yoshua Bengio

- Extend 5-layer CNN to 10-layer ResNet
- Solve gradient vanishing with shortcut connections

CNN-5 [Watanabe+, SPIE'17]
ResNet-10

Case I: From N10 to N7

Testing RMS Error vs N7a Training Set Size (%)

Case II: From N7_a to N7_b

Testing RMS Error vs N7b Training Set Size (%)

	Case I	Case II
Knowledge Transfer	N10 → N7 _a	N7 _a → N7 _b
Dataset Similarity	Medium	High
Best k	0/4	8

Modeling Photoresist

Intensity
x
y

Intensity
x
y

Resist model $f: X \rightarrow Y$

Predicted Patterns
Match
Manufactured Patterns

Problem formulation

Source Domain (x_s, y_s) Train Source Model $f_s: X \rightarrow Y$

Target Domain (x_t, y_t) Fine-tune Target Model $f_t: X \rightarrow Y$

Starting point

Transfer Learning for Lithography Modeling

Challenges in Lithography Modeling

Rigorous simulation

- Physics-level simulation
- e.g., Synopsys Sentaurus Lithography

Compact model

- e.g., Mentor Graphics Calibre, machine learning models

	Rigorous	Compact
Accuracy	High	Medium
Prediction Efficiency	Low	High
Data Demands	Medium	High

Prediction Efficiency
Data Demanding

For 1K 2x2um² clips

CD RMS Error (nm) vs Training Set Size (%)

High Accuracy → big training data
Expensive to prepare data

- Time consuming
- Manufacturing cost

10nm Data
7nm Data

Data amount ↑

Training
Knowledge Transfer
Training

Litho Model
Generality ↑ Accuracy ↑

TF_k scheme

Source Domain Input
Target Domain Input

Knowledge Transfer
Fix k layers
Fine-tune

Source Domain Output
Target Domain Output

Technology Transition from N10 to N7

Contact Layer Design Rules [Liebmann, SPIE'15]		
	N10	N7
Patterning	LELE	LELELE
Pitch (nm)	64	45
Mask pitch (nm)	128	135
Litho-target (nm)	60	60

Optical Sources
N10
N7
Resist Materials
Resist A
Resist B
Different dissolution slopes

	N10	N7 _a	N7 _b
Design Rule	A	B	B
Optical Source	A	B	B
Resist Material	A	A	B

Active Learning with Clustering

Target Domain Dataset D
 $\forall(x, ?) \in D$
Making x is cheap
Querying ? is expensive

Unlabeled data
Labeled dataset s

Random Selection
Active Selection

Intuition: close features have close labels

Theorem

$$\frac{1}{n} \sum_{i \in D} Loss_i \leq \frac{C}{n} \sum_{j \in S} \sum_{i=1}^{k_j} \|x_i - x_j^c\| + \epsilon$$

K-Medoids Clustering
(A variation of K-Means Clustering)

Transfer & Active Learning Flow

Labelled Source Domain Data
Unlabelled Target Domain Data

Active Data Selection
Label Querying
Data Augmentation
Knowledge Transfer
Source Model Training
Target Model Training

From N10 to N7

Testing RMS Error vs N7a Training Set Size (%)

Training Data vs CD RMS Error (nm)

- Improve data efficiency
- ~3~10X reduction of training data
- Reduce turn-around time
- Increase model accuracy