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

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

Transfer Learning for Lithography Modeling

Source Domain
Train
Source Model
\( f : X \rightarrow Y \)

Target Domain
Fine-tune
Target Model
\( f : X \rightarrow Y \)

Case I: From N10 to N7

<table>
<thead>
<tr>
<th>Case I</th>
<th>Case II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge Transfer</td>
<td>N10→N7</td>
</tr>
<tr>
<td>Dataset Similarity</td>
<td>Medium</td>
</tr>
<tr>
<td>Best k</td>
<td>q/0</td>
</tr>
</tbody>
</table>

Active Learning with Clustering

Unlabeled data
Labeled dataset

Random Selection
Active Selection

Intuition: close features have close labels

Theorem
\[ \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} (\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{X} (\mathbf{x}_i - \mathbf{x}_j) \leq \epsilon \]

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
Source Model Training
Target Model Training

From N10 to N7

<table>
<thead>
<tr>
<th>N10</th>
<th>N7</th>
<th>N7'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design Rule</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Optical Source</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Resist Material</td>
<td>A</td>
<td>A</td>
</tr>
</tbody>
</table>

- Improve data efficiency
- Reduce turn-around time
- Increase model accuracy