EMOGen: Enhancing Mask Optimization via Pattern Generation

Su Zheng¹, Yuzhe Ma², Bei Yu¹, Martin Wong³

¹Chinese University of Hong Kong
²Hong Kong University of Science and Technology (GZ)
³Hong Kong Baptist University
Outline

1 Introduction
2 Preliminaries
3 Method
4 Experiments
Introduction
Motivation

DNN-based Layout Pattern Generation

Original Pattern $T$  Generative DNN $G(T)$  Generated Patterns $T'$

DNN-based Inverse Lithography Technique

Target Image $Z_T$  DNN ILT Model $F(Z_T)$  Optimized Mask $M^*$
Contributions

- **EMOGen**: co-evolution of layout generation and mask optimization
  - Use layout generation to improve DNN-based ILT methods
- ILT-aware training and legalization schemes
  - Discover the weaknesses of the DNN-based ILT model
- Extensive experiments verify the effectiveness of EMOGen
  - 39% enhancement in DNN-based ILT
  - 34% improvement in pattern legalization
Preliminaries
Efficient representation of the layout patterns

- A topology matrix + two geometry vectors

\[
\begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
\end{bmatrix}
\]

\[
\Delta_x = [10, 20, 30, 20]
\]

\[
\Delta_y = [10, 30, 20, 20]
\]
Autoencoder-based approaches

\[ T' = f_{\text{dec}} \left( f_{\text{enc}}(T) + \epsilon \mathcal{N}(0, I) \right). \]  

- Map the topology matrices to a latent space
- Generate new patterns by perturbing latent features
Mask Optimization

- ILT for Mask Optimization

- Lithography simulation: \( I = H(M) = \sum_{k=1}^{K} \mu_k |h_k \otimes M|^2 \).

- Objectives: 
  \[
  L2(Z_{\text{nom}}, Z_T) = \|Z_{\text{nom}} - Z_T\|^2, \\
  PVB(Z_{\text{max}}, Z_{\text{min}}) = \|Z_{\text{max}} - Z_{\text{min}}\|^2.
  \]
Method
Mask optimization problem

\[ M^* = f_M(Z_T | \theta_M). \]  

- \( Z_T \) is the target image.
- \( M^* \) represents the optimized masks given by the ILT model.
Mathematical Formulation

- Pattern generation problem

\[ T', \Delta'_x, \Delta'_y = f_P(T, \Delta_x, \Delta_y | \theta_P). \]  \hspace{1cm} (3)

- \( f_P(\cdot) \) denotes the pattern generation model with the parameters \( \theta_P \).
- \( T', \Delta'_x, \Delta'_y \) represent the generated topology and geometry.
Co-optimization problem → two players competing with each other

\[
\min_{\theta_M} \max_{\theta_P} L_{ILT} (f_M(X|\theta_M), X) \quad \text{s.t.} \quad X = r (f_P(T, \Delta x, \Delta y|\theta_P)) . \tag{4}
\]

- \( L_{ILT} \) is the loss function of ILT.
- \( r(\cdot) \) converts the generated pattern to ILT input.
Legalization of the generated patterns

\[ \sum_{k \in k_x} \Delta'_{x,k} \geq \text{Space}_{\text{min}}, \sum_{k \in k_y} \Delta'_{y,k} \geq \text{Space}_{\text{min}}, \forall k_x, k_y \in S_{\text{min}}, \]  
\[ \sum_{l \in l_x} \Delta'_{x,l} \geq \text{Width}_{\text{min}}, \sum_{l \in l_y} \Delta'_{y,l} \geq \text{Width}_{\text{min}}, \forall l_x, l_y \in W_{\text{min}}, \]  
\[ \sum_{(i,j) \in p} \Delta'_{x,i} \Delta'_{y,j} \in [\text{Area}_{\text{min}}, \text{Area}_{\text{max}}], \forall \text{ polygon } p. \]
Pattern Generation Model
ILT Models

- **GAN-OPC**\(^1\) It follows the design of generative adversarial network (GAN).
- **Neural-ILT**\(^2\) A UNet is utilized in Neural-ILT to predict the optimized mask.
- **DAMO**\(^3\) It improves the GAN for ILT with the backbone based on UNet++ and a multiscale discriminator.
- **CFNO**\(^4\) Combining the basic principles of Vision Transformer (ViT) and Fourier Neural Operator (FNO).

---


\(^3\) Guojin Chen et al. (2020). “DAMO: Deep agile mask optimization for full chip scale”. In: Proc. ICCAD.

Combining Pattern Generation and ILT Models

• Make it differentiable

(c)

(d)
Overview of EMOGen Training

Step 1: Pattern Generation Optimization

DNN-based Pattern Generation

Pattern $Z_0$ → Squish Patterns $T^1, \Delta_1, \Delta_0^1$ → New Patterns $T^2, \Delta_2, \Delta_0^2$ → DNN $f_P(T, \Delta_x, \Delta_y|\theta_P)$ → New Patterns

Step 2: ILT Optimization

DNN-based ILT

Target Image $Z_T$ → DNN $f_M(Z_T|\theta_M)$ → Optimized Mask $M^*$

$L_{Gen} = KL_T + KL_A - L_{ILT}\theta_P = \theta_P - \frac{\partial L_{Gen}}{\partial \theta_P}$

$L_{ILT} = L^2 + PVB\theta_M = \theta_M - \frac{\partial L_{ILT}}{\partial \theta_M}$
Experiments
Comparison Between ILT Models With and Without Co-evolution

- Better ILT performance.
Comparison on the Legalization of Generated Patterns

- Better pattern generation: effectively deteriorate the ILT performance.
- Better pattern legalization: the generated results have a smaller average number of design rule violations.

<table>
<thead>
<tr>
<th>Metric</th>
<th>μν</th>
<th>L2</th>
<th>PVB</th>
<th>EPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Legalization</td>
<td>0.094</td>
<td>51777</td>
<td>57149</td>
<td>22.0</td>
</tr>
<tr>
<td>Design Rules Only</td>
<td>0.070</td>
<td>50782</td>
<td>57335</td>
<td>21.4</td>
</tr>
<tr>
<td>Design Rules + ILT (ours)</td>
<td>0.062</td>
<td>64800</td>
<td>65394</td>
<td>44.1</td>
</tr>
</tbody>
</table>

Examples from the Trained Pattern Generation and ILT Models
Thanks!