VLSI Mask Optimization: From Shallow To Deep Learning

Haoyu Yang\textsuperscript{1}, Wei Zhong\textsuperscript{2}, Yuzhe Ma\textsuperscript{1}, Hao Geng\textsuperscript{1}, Ran Chen\textsuperscript{1}, Wanli Chen\textsuperscript{1}, Bei Yu\textsuperscript{1}

\textsuperscript{1}The Chinese University of Hong Kong
\textsuperscript{2}Dalian University of Technology
Moore’s Law to Extreme Scaling

Moore’s Law

Process Technology ($\mu$m)

Year

Number of Transistors per Integrated Circuit


10,000,000,000 1,000,000,000 100,000,000 10,000,000 1,000,000 100,000 10,000 1,000 100 10 1

- Invention of the Transistor
- Doubles every 2.1 yrs

Intel Microprocessors
Apple Microprocessors

Doubles every 2.1 yrs

A7
A10 A11
A12

Intel Microprocessors
Apple Microprocessors
Challenge 1: Failure (Hotspot) Detection

- RET: OPC, SRAF, MPL
- Still hotspot: low fidelity patterns
- Simulations: extremely CPU intensive
Challenge 2: Optical Proximity Correction (OPC)

Design target

Mask

Wafer

without OPC

with OPC

without OPC

with OPC
Why Deep Learning?

▶ **Feature Crafting v.s. Feature Learning**
   Although prior knowledge is considered during manually feature design, information loss is inevitable. Feature learned from mass dataset is more reliable.

▶ **Scalability**
   With shrinking down circuit feature size, mask layout becomes more complicated. Deep learning has the potential to handle ultra-large-scale instances while traditional machine learning may suffer from performance degradation.

▶ **Mature Libraries**

![TensorFlow](logo_tfl.png) ![Caffe2](logo_caffe.png) ![PyTorch](logo_torch.png)
Outline

Hotspot Detection via Machine Learning

OPC via Machine Learning

Heterogeneous OPC
Outline

Hotspot Detection via Machine Learning

OPC via Machine Learning

Heterogeneous OPC
Hotspot Detection Hierarchy

- **Sampling** (DRC Checking):
  scan and rule check each region

- **Hotspot Detection**:
  verify the sampled regions and report potential hotspots

- **Lithography Simulation**:
  final verification on the reported hotspots

Increasing verification accuracy

(Relative) CPU runtime at each level
Early Study of DNN-based Hotspot Detector

- Total 21 layers with 13 convolution layers and 5 pooling layers.
- A ReLU is applied after each convolution layer.

What Does Deep Learning Learn?

Origin Pool1 Pool2
Pool3 Pool4 Pool5
The Biased Learning Algorithm [DAC’17]

Training Set

MGD: end-to-end training

Update $\varepsilon$

$y_h = [0, 1]$

$y_n = [1 - \varepsilon, \varepsilon]$

Stop Criteria

No

Yes

Trained Model

The AUC objective:

\[ \mathcal{L}_\Phi(f) = \frac{1}{N_+ N_-} \sum_{i=1}^{N_+} \sum_{j=1}^{N_-} \Phi \left( f(x_i^+) - f(x_j^-) \right) . \]

Approximation candidates:

- **PSL**  \( \Phi_{PSL}(z) = (1 - z)^2 \)
- **PHL**  \( \Phi_{PHL}(z) = \max(1 - z, 0) \)
- **PLL**  \( \Phi_{PLL}(z) = \log(1 + \exp(-\beta z)) \)
- **R**  \( \Phi_{R^*}(z) = \begin{cases} 
- (z - \gamma)^p, & \text{if } z > \gamma \\
0, & \text{otherwise}
\end{cases} \)

---

Conventional Clip based Solution

- A binary classification problem.
- Scan over whole region.
- Single stage detector.

- Scanning is time consuming and single stage is not robust to false alarm.
Learning **what** and **where** is hotspot at same time.

Classification Problem -> Classification & Regression Problem.

---
Outline

Hotspot Detection via Machine Learning

OPC via Machine Learning

Heterogeneous OPC
OPC Previous Work

Classical OPC

- Model/Rule-based OPC
  - [Cobb+, SPIE’02][Kuang+, DATE’15]
  - [Awad+, DAC’16][Su+, ICCAD’16]
    1. Fragmentation of shape edges;
    2. Move fragments for better printability.

- Inverse Lithography
  - [Pang+, SPIE’05][Gao+, DAC’14]
  - [Poonawala+, TIP’07][Ma+, ICCAD’17]
    1. Efficient model that maps mask to aerial image;
    2. Continuously update mask through descending the gradient of contour error.

Machine Learning OPC

- [Matsunawa+, JM3’16][Choi+, SPIE’16]
- [Xu+, ISPD’16][Shim+, APCCAS’16]
  1. Edge fragmentation;
  2. Feature extraction;
  3. Model training.
Machine Learning-based SRAF Insertion

SRAF Insertion with Machine Learning [ISPD’16]

Tackling Robustness with Dictionary Learning [ASPDAC’19]


Replace lithography simulation (slow) with machine learning-based EPE predictor (fast) in OPC iterations.

GAN-OPC [DAC’18]

Better starting points for legacy OPC engine and reduce iteration count.

Outline

Hotspot Detection via Machine Learning

OPC via Machine Learning

Heterogeneous OPC
An Observation of Previous OPC Solutions

Machine learning solutions rely on legacy OPC engines

Legacy OPC engines exhibit different performance on different designs

![Diagram showing the comparison between MB-OPC and ILT with MSE values for different design IDs.](image)

MB-OPC
ILT
We design a classification model that can determine the best OPC engine for a given design at trivial cost.
Training on Artificial Designs

- Training data comes from GAN-OPC and is labeled according to results of MB-OPC and ILT.
- Test on 10 designs from ICCAD 2013 CAD Contest.
Experimental Results

Several Benefits

▶ Does not require extremely high prediction accuracy of the classification model.
▶ Take advantages of different OPC solutions on different designs.
Conclusion and Discussion

So Far:
▶ Recent progress of deterministic machine learning model for hotspot detection
▶ State-of-the-art machine learning solutions for OPC and SRAF insertion
▶ A heterogeneous OPC framework guided by a classification engine

Future:
▶ Manufacturability issues.
▶ Classification challenge when more than two OPC engines are available.