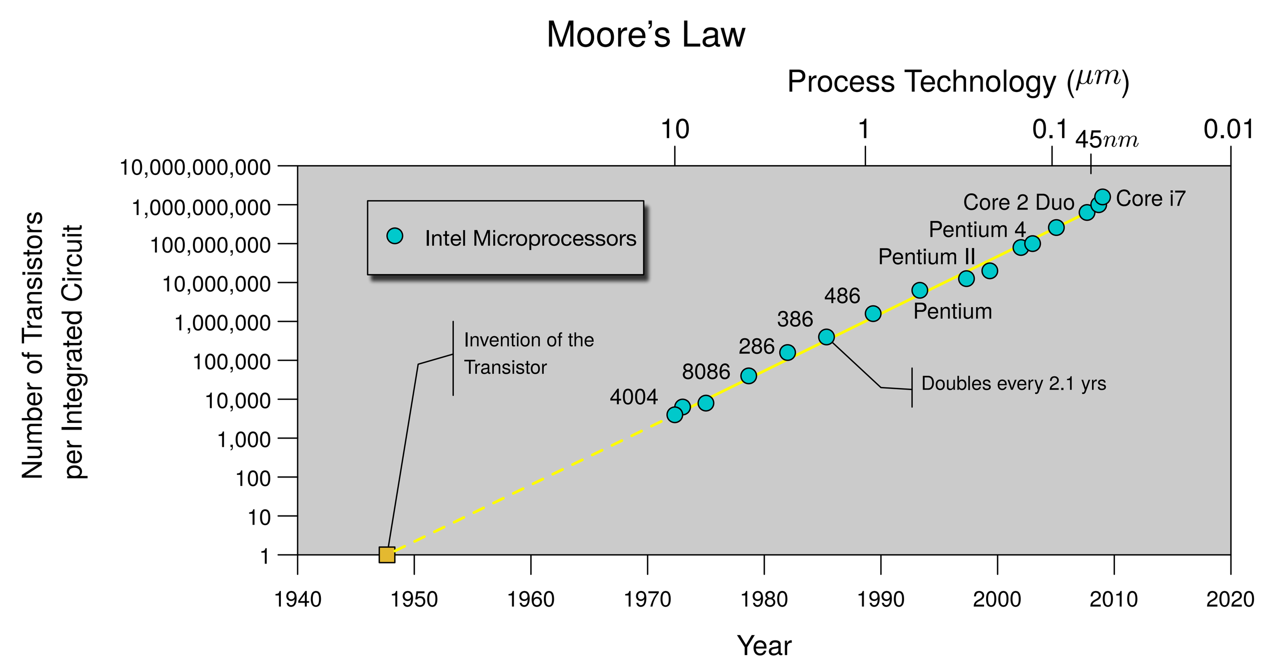
DeePattern: Layout Pattern Generation with Transforming Convolutional Auto-Encoder

Haoyu Yang[1], Piyush Pathak[2], Frank Gennari[2], Ya-Chieh Lai[2] and Bei Yu[1]

[1]Department of Computer Science and Engineering, The Chinese University of Hong Kong
[2]Cadence Design Systems, Inc.

Backgrounds

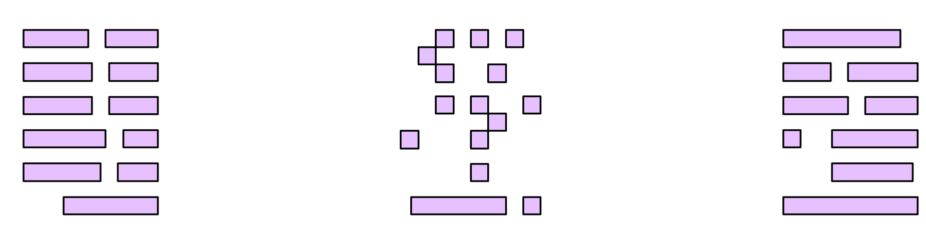
Moore's Law to Extreme Scaling



EUV Brings Challenges in DFM

- Hotspot detection and fix
  - Previous researches show the significance of a diverse and balanced training data set.[Yang+,SPIE'17]
  - Hotspot pattern library covering the design space required by machine learning and pattern matching solutions.
  - Lithographic simulation challenge due to complicated computational lithography model under EUV nodes. [Levinson+,SPIE'18]
- Early technology node development
  - Due to long logic to layout cycle, test layout patterns are not usually available.
  - OPC convergence problem.
  - Patterns are required to massage Design rule, OPC recipe, ...

Related Works on Pattern Generation

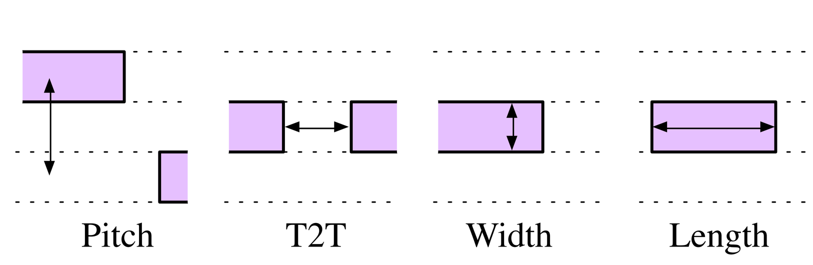


- Transferring from previous technology node. (not applicable for large technology node gap) [Zhuang+,ICSICT'16]
- Randomly placing patterns according to certain constraints. (limited diversity)
- Generative machine learning models. (violating design rules) [Alec+,ICLR'16]

Preliminaries

Pattern Generation Challenges

- *7nm* EUV metal layer unidirectional on-track shapes.
- *Pitch*, denoted as p, measures the distance between two adjacent tracks that contain shapes.
- *T2T*, denoted as t, measures the line-end-to-line-end distance between two adjacent shapes in a track.
- Wire *length* l and *width* w measure the shape size along and against the design track.



Evaluation of Pattern Library

All shape edges in a fixed-size window are aligned with x-axis and y-axis. If we extend all horizontal and vertical edges infinitely into scan lines, more non-overlapping scan lines always come with more complex patterns. We hence define the complexity of a layout pattern as follows.

- **Pattern Complexity**. The complexity of a pattern in x and y directions (denoted as c_x and c_y) are defined as the number of scan lines subtracted by one along x-axis and y-axis, respectively.

We also introduce the concept of *pattern diversity* (denoted as H) to measure how are the pattern complexities distributed in a given library. A larger H implies the library contains patterns that are more evenly distributed, as in the following definition.

- **Pattern Diversity**. The diversity of a pattern library is given by the *Shannon Entropy* of the pattern complexity sampled from the library,

$$\mathrm{H} = -\sum_i \sum_j P(c_{xi}, c_{yj}) \log P(c_{xi}, c_{yj}),$$

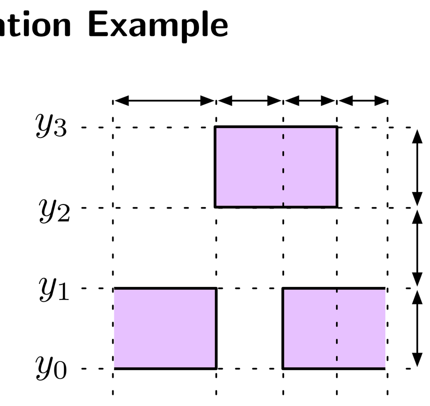
where $P(c_{xi}, c_{yj})$ is the probability of a pattern sampled from the library has complexities of c_{xi} and c_{yj} in x and y directions respectively.

Problem (Pattern Generation)

Given a set of layout design rules, the objective of pattern generation is to generate a pattern library such that the pattern diversity and the number of unique DRC-clean patterns in the library is maximized.

Methods

Squish Representation Example



$$T = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix}$$
$$\delta_x = \begin{bmatrix} x_1 - x_0 & x_2 - x_1 & x_3 - x_2 & x_4 - x_3 \end{bmatrix}$$
$$\delta_y = \begin{bmatrix} y_1 - y_0 & y_2 - y_1 & y_3 - y_2 \end{bmatrix}$$

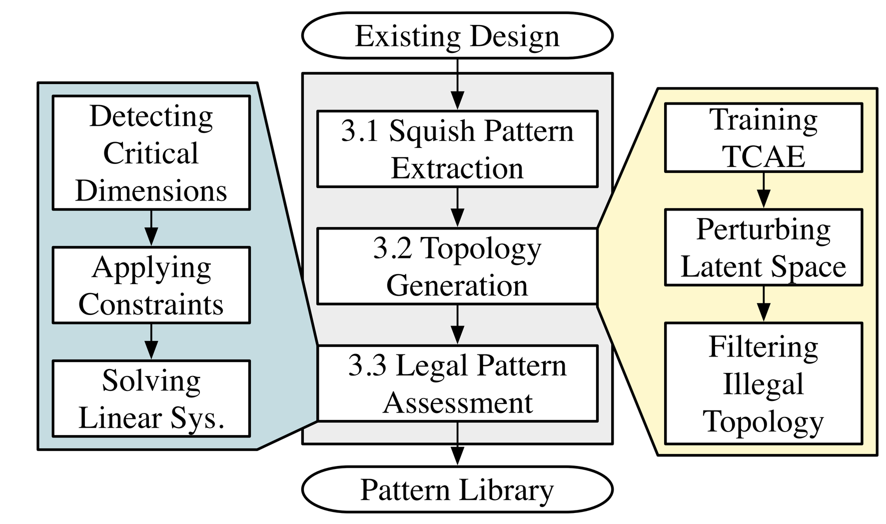
- Scan line-based representation, naturally supports easy computation of pattern complexity.
- Lossless feature representation.
- Easily feed into convolutional neural networks.

Problem Simplification

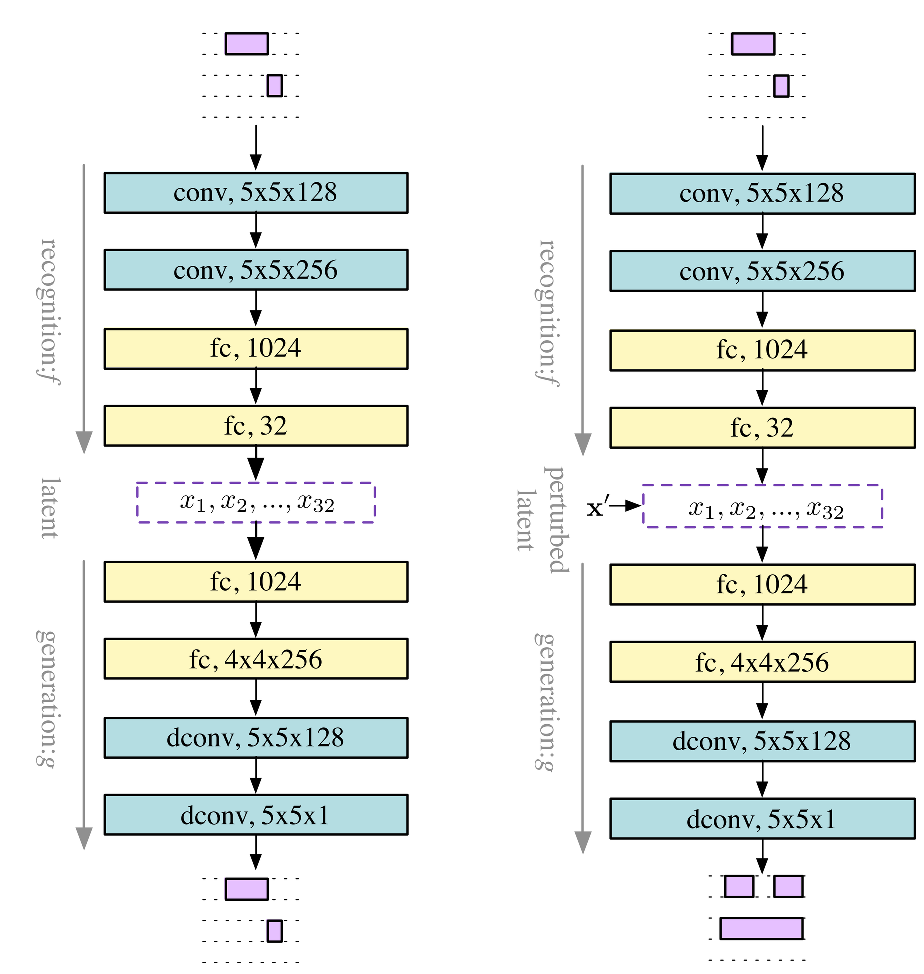
- Legal topology generation.
- Solving geometry constraints for DRC clean patterns.

Methods

The Overall Flow



Transforming Convolutional Auto-Encoder



Input pattern to latent space,

$$\boldsymbol{I} = f(\boldsymbol{T}; \boldsymbol{W}_f)$$

Topology reconstruction,

$$\boldsymbol{T}' = g(\boldsymbol{I} + \Delta \boldsymbol{I}; \boldsymbol{W}_g)$$

Training objective,

$$\min_{\boldsymbol{W}_f, \boldsymbol{W}_g} ||\boldsymbol{T} - \boldsymbol{T}'||, \text{ s.t. } \Delta \boldsymbol{I} = \boldsymbol{0}$$

- Inspired from TAE. [Hinton+,ICANN'11]
- Feature instantiation attains data set domain properties.
- All capsules contribute together to produce variations of any input objects.
- The transformation in our framework applies directly on the latent vector space that promises a much larger diversity of the generated patterns compared to the limited transformation on the coordinate system only in TAEs.
- Identity mapping in the training phase helps the TCAE capture the design rule properties of existing patterns.

Perturbing the Latent Space

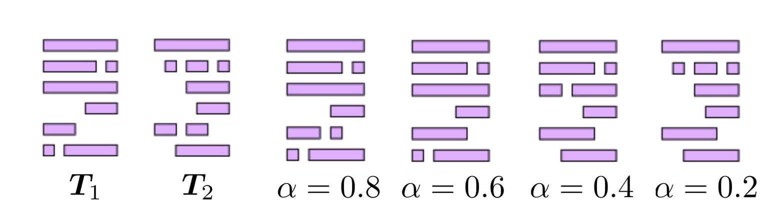
TCAE-Combine

- Generalization from existing topologies

$$\boldsymbol{T}_g = g(\sum_i \alpha_i f(\boldsymbol{T}_i)),$$

where $0 < \alpha_i < 1, \forall i$ are combination coefficients and satisfy $\sum_i \alpha_i = 1$.

- Sample results



TCAE-Random

- Introducing perturbation from certain distribution randomly

$$\boldsymbol{T}_g = g(f(\boldsymbol{T}_i) + \Delta \boldsymbol{x}),$$

where $\Delta \boldsymbol{x} \sim \mathcal{N}$.

- **Feature Sensitivity**. Let $\boldsymbol{I} = [l_1 \; l_2 \; \dots \; l_n]^\top$ be the output of the layer associated with the latent vector space. The sensitivity s_i of a latent vector node l_i is defined as the probability of reconstructed pattern being invalid when a perturbation $\Delta l_i \in [-t, t]$ is added up on l_i with everything else unchanged.

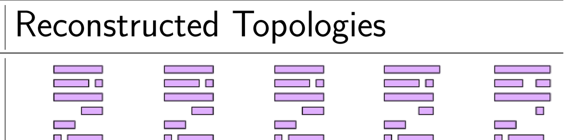
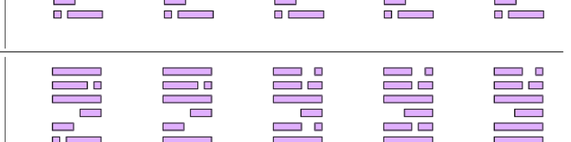
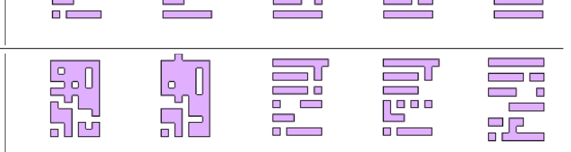
Legal Pattern Assessment

Creating DRC constraints for legal $\boldsymbol{\delta}_x$s and $\boldsymbol{\delta}_y$s,

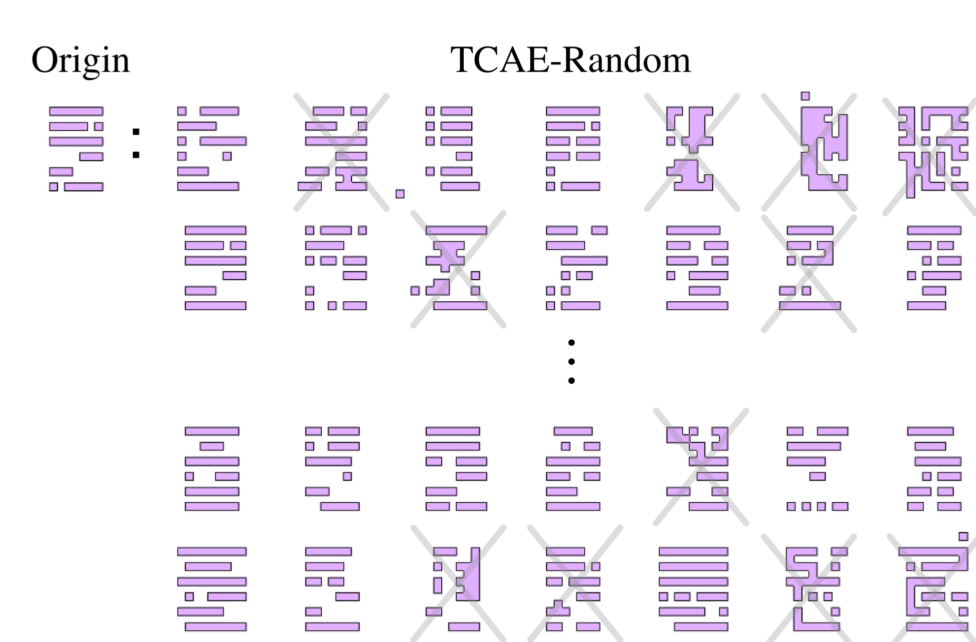
$$\begin{aligned} y_{i+1} - y_i &= \frac{p}{2}, && \forall i, \\ x_i - x_j &= t_{\min}, && \forall (i,j) \in \mathcal{C}_{T2T}, \\ x_i - x_j &= l_{\min}, && \forall (i,j) \in \mathcal{C}_{W}, \\ x_{i+1} - x_i &> 0, && \forall i, \\ x_{\max} - x_0 &= d_x, y_{\max} - y_0 = d_y. \end{aligned}$$

Results and Conclusion

Understanding Features in TCAE

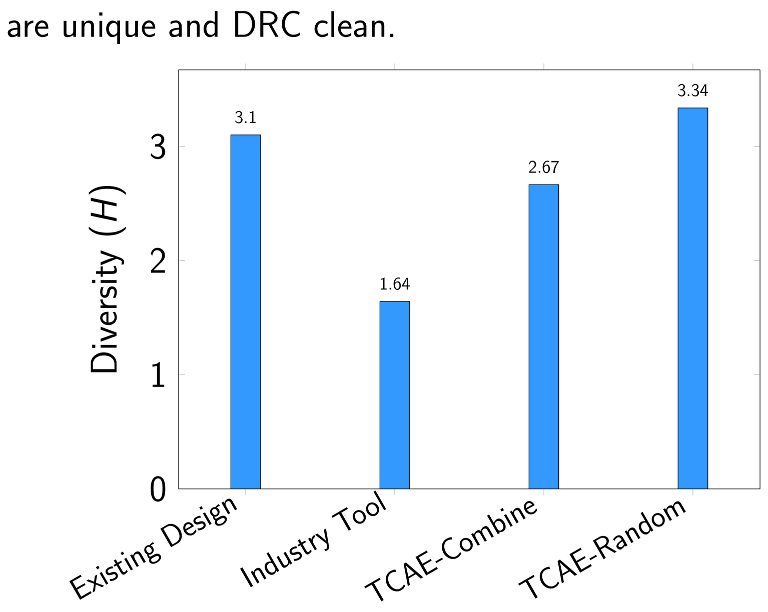
Transformations	Reconstructed Topologies
Extend or pull back line-ends	
Create or destroy shapes	
Control shape directions	

TCAE-Random



Comparison with State-of-the-Art

Perturbation with Gaussian exhibits greatest pattern generation power with around 30% generated patterns are unique and DRC clean.



Conclusion

- Address the pattern library requirements in DFM flows/researches under advanced technology nodes.
- Propose a TCAE framework that can capture layout design rule characteristics.
- We show auto-learned features contribute to layout space locally or globally.