VLSI Layout Hotspot Detection Based on Discriminative Feature Extraction

Hang Zhang, Haoyu Yang, Bei Yu, Evangeline F.Y. Young

Department of Computer Science and Engineering,
The Chinese University of Hong Kong

Oct 28, 2016

Outline

Problem Background

Conventional Methods

Our Method

Results

Outline

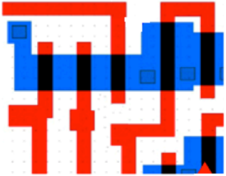
Problem Background

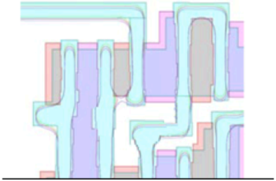
Conventional Methods

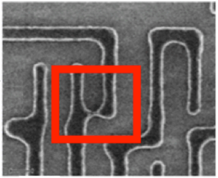
Our Method

Results

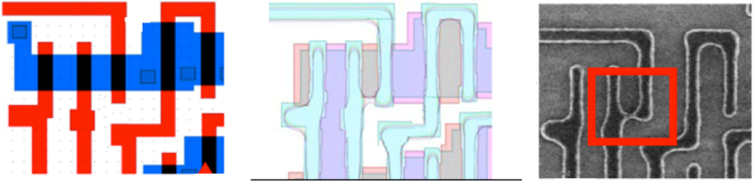
Lithography Hotspot Detection Background





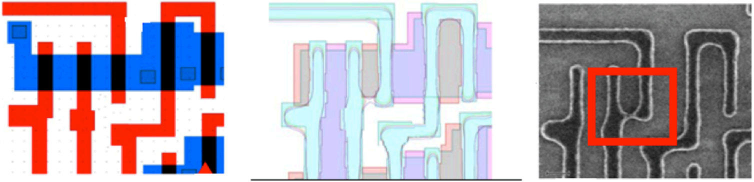


Lithography Hotspot Detection Background



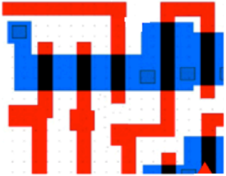
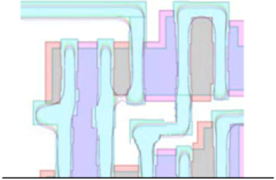
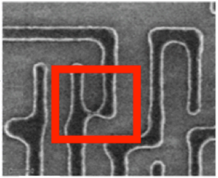
- What you see $\neq$ what you get

Lithography Hotspot Detection Background



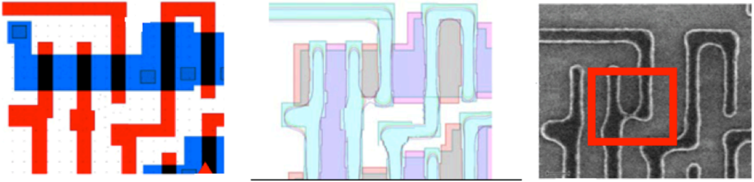
- What you see $\neq$ what you get
- DFM: OPC, RET, MPL

Lithography Hotspot Detection Background

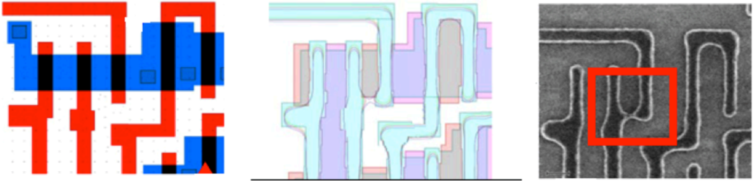
- What you see $\neq$ what you get
- DFM: OPC, RET, MPL
- Still hotspot: low fidelity

Lithography Hotspot Detection Background

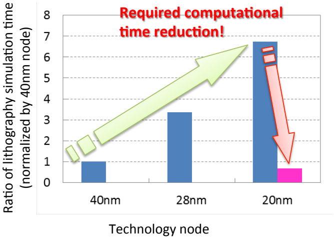


- What you see $\neq$ what you get
- DFM: OPC, RET, MPL
- Still hotspot: low fidelity
- Simulations: extremely CPU intensive

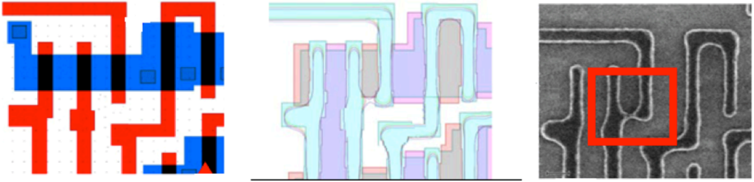
Lithography Hotspot Detection Background



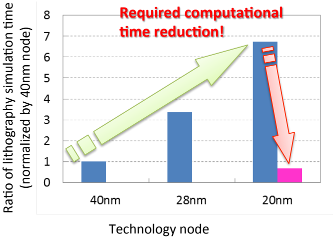
- What you see $\neq$ what you get
- DFM: OPC, RET, MPL
- Still hotspot: low fidelity
- Simulations: extremely CPU intensive



Lithography Hotspot Detection Background



- What you see ≠ what you get
- DFM: OPC, RET, MPL
- Still hotspot: low fidelity
- Simulations: extremely CPU intensive



- Pattern Matching or Machine Leaning?

Key Issue in Hotspot Detection?

Key Issue in Hotspot Detection?

- The first problem in machine learning based hotspot detection is?

Key Issue in Hotspot Detection?

- The first problem in machine learning based hotspot detection is?
- Definitely layout pattern feature extraction.

Key Issue in Hotspot Detection?

- The first problem in machine learning based hotspot detection is?
- Definitely layout pattern feature extraction.
- We need discriminative pattern information to detect hotspot.

Outline

Problem Background

Conventional Methods

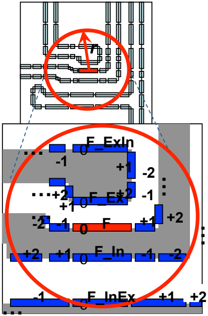
Our Method

Results

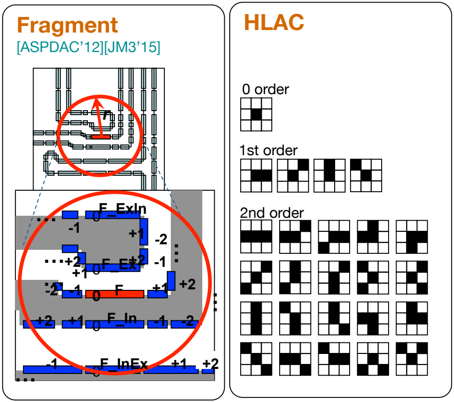
Conventional Feature Extraction

Fragment

[ASPDAC'12][JM3'15]

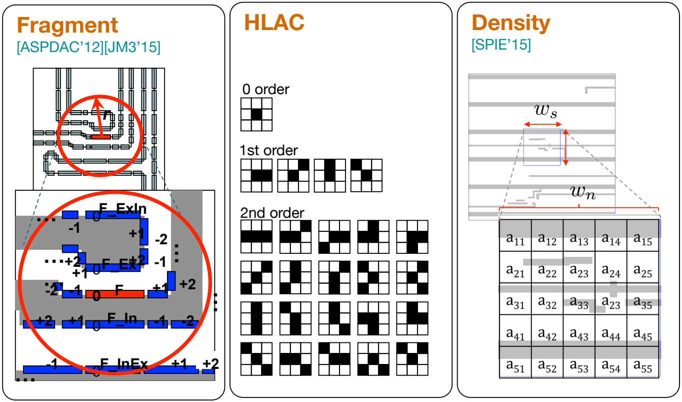


Conventional Feature Extraction



- Fragment feature is very complicated, which leads to over-fitting.

Conventional Feature Extraction

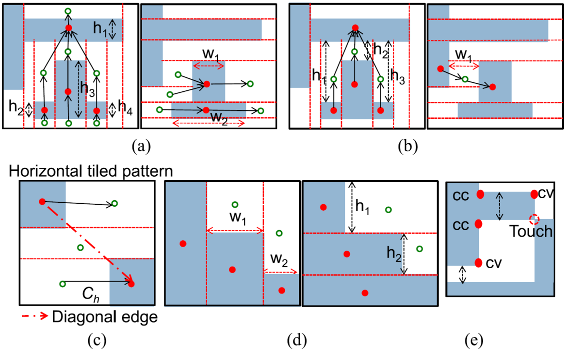


- Fragment feature is very complicated, which leads to over-fitting.
- High order local correlation (HLAC) is only efficient in some image processing task.

Conventional Feature Extraction

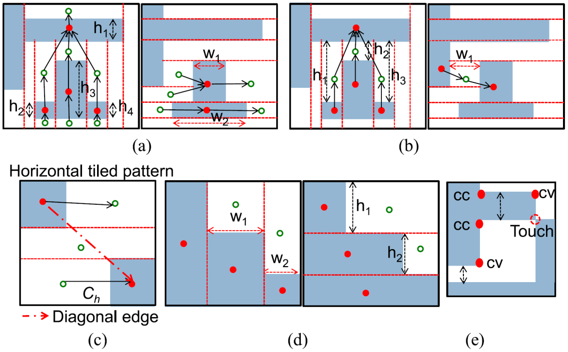
- Fragment feature is very complicated, which leads to over-fitting.
- High order local correlation (HLAC) is only efficient in some image processing task.
- Density based feature loses some important pattern information.

Feature in [TCAD'15]



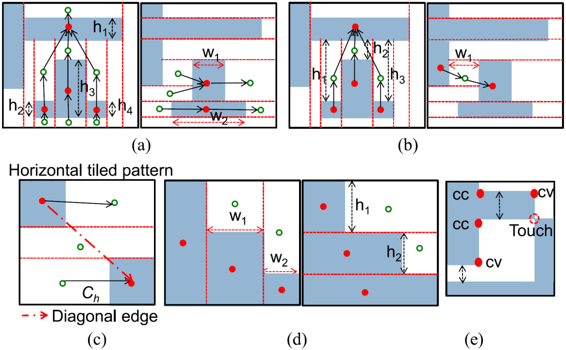
- (a) Internal feature

Feature in [TCAD'15]



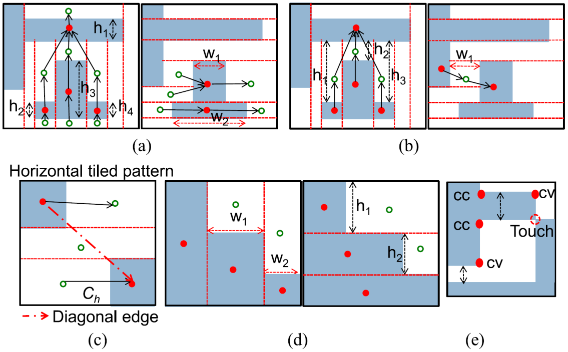
- (a) Internal feature
- (b) External feature

Feature in [TCAD'15]



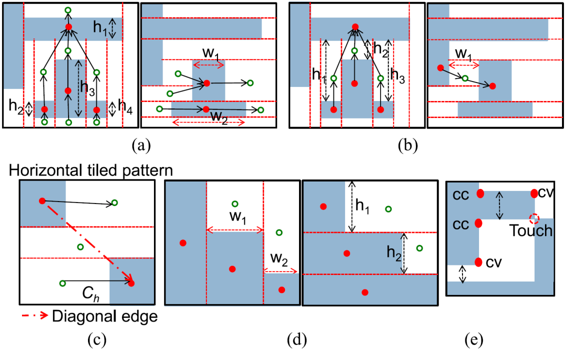
- (a) Internal feature
- (b) External feature
- (c) Diagonal feature

Feature in [TCAD'15]



- (a) Internal feature
- (b) External feature
- (c) Diagonal feature
- (d) Segment feature

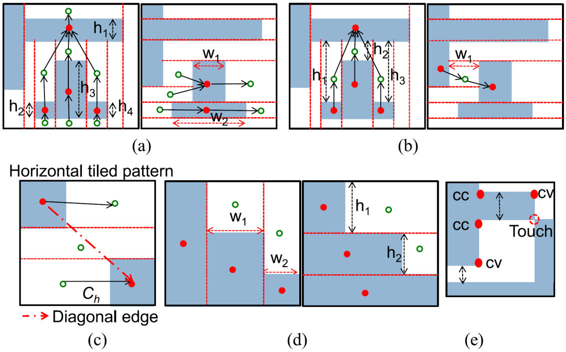
Feature in [TCAD'15]



- (a) Internal feature
- (b) External feature
- (c) Diagonal feature
- (d) Segment feature

- Pros: easy and fast to extract.

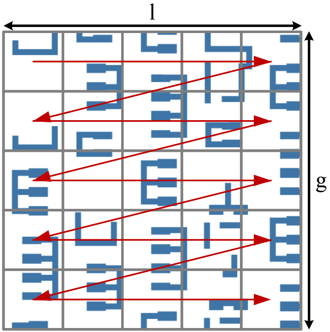
Feature in [TCAD'15]



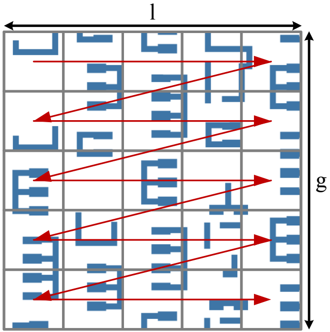
- (a) Internal feature
- (b) External feature
- (c) Diagonal feature
- (d) Segment feature

- Pros: easy and fast to extract.
- Cons: still complicated, hard to detect new patterns.

Density based Feature [SPIE'15]

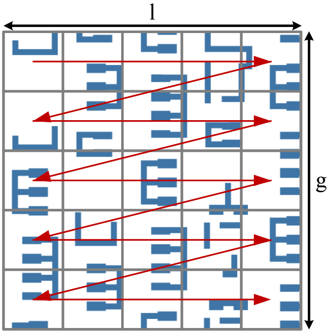


Density based Feature [SPIE'15]



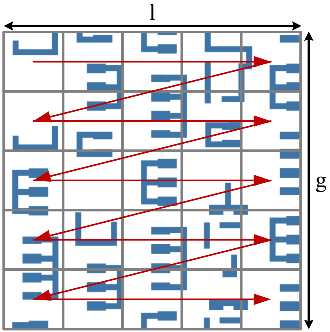
- Side length l, grid number g.

Density based Feature [SPIE'15]



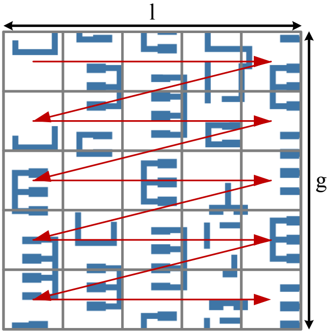
- Side length l, grid number g.
- a_{ij} denotes the density value of the grid in i^{th} row and the j^{th} column.

Density based Feature [SPIE'15]



- Side length l, grid number g.
- a_{ij} denotes the density value of the grid in i^{th} row and the j^{th} column.
- Feature vector: $X = \{a_{11}, a_{12} \ldots, a_{54}, a_{55}\}$

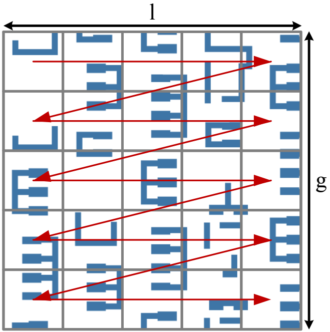
Density based Feature [SPIE'15]



- Side length l, grid number g.
- a_{ij} denotes the density value of the grid in i^{th} row and the j^{th} column.
- Feature vector: $X = \{a_{11}, a_{12}..., a_{54}, a_{55}\}$

- Pros: simple and efficient compared to previous methods.

Density based Feature [SPIE'15]



- Side length l, grid number g.
- a_{ij} denotes the density value of the grid in i^{th} row and the j^{th} column.
- Feature vector: $X = \{a_{11}, a_{12} \ldots, a_{54}, a_{55}\}$

- Pros: simple and efficient compared to previous methods.
- Cons: Severe layout pattern information loss.

Outline

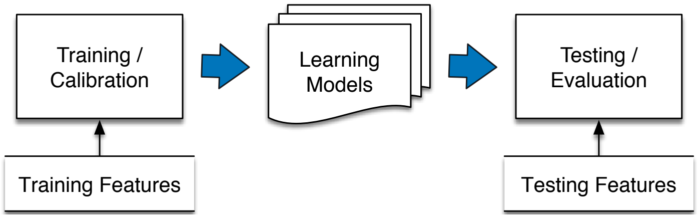
Problem Background

Conventional Methods

Our Method

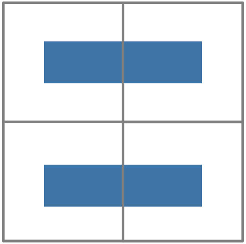
Results

Learning Framework

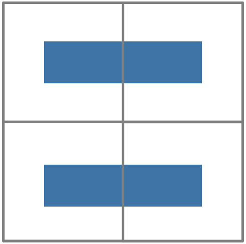


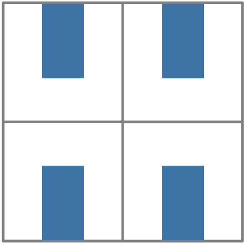
- Training stage $\rightarrow$ models.
- Testing stage
- Learning models: Decision-tree, ANN, SVM...

Major Drawbacks of Conventional Density Based

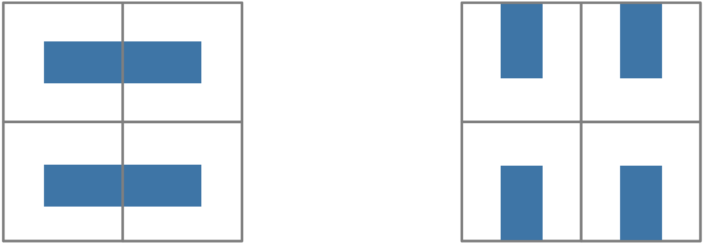


Major Drawbacks of Conventional Density Based



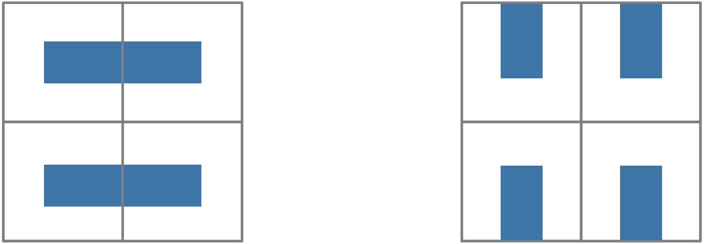


Major Drawbacks of Conventional Density Based



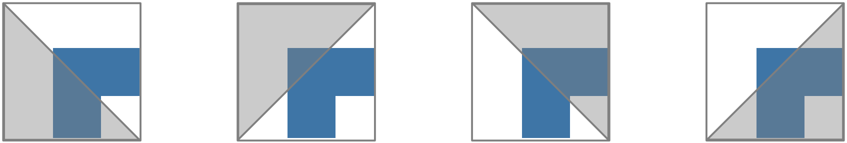
- For both patterns, we can only get the same feature vector.

Major Drawbacks of Conventional Density Based

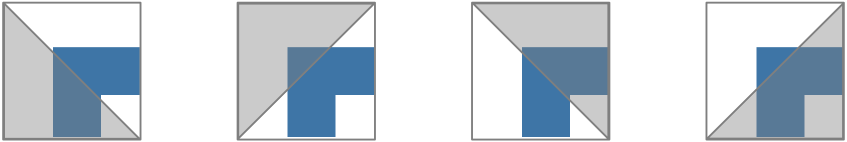


- For both patterns, we can only get the same feature vector.
- However, their contributions to the hotspot formation are different.

Local Grid Density Differential (LGDD)

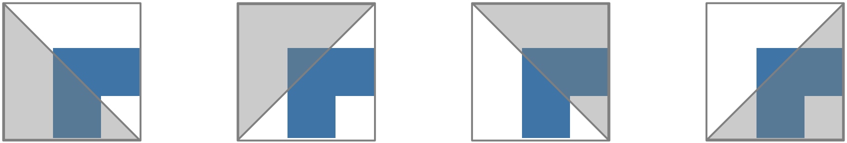


Local Grid Density Differential (LGDD)



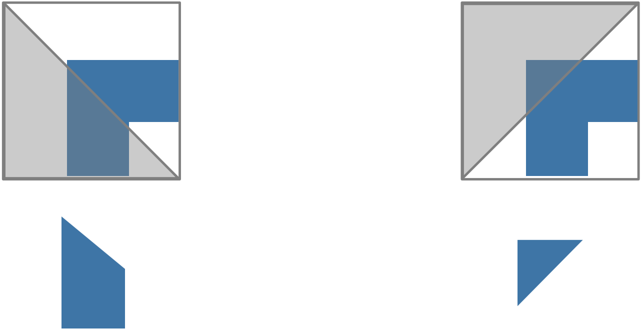
- locally average the density value of a specific area in a grid.

Local Grid Density Differential (LGDD)



- locally average the density value of a specific area in a grid.
- We apply triangle area in this paper.

Examples of LGDD



- The area value of the blue region in the shadow part.

Definitions for Evaluations

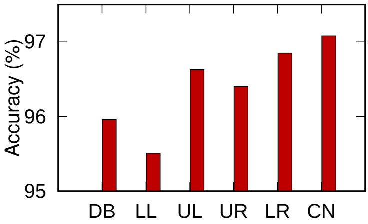
- Accuracy: The rate of correctly predicted hotspots among the set of actual hotspots.
- Extra: The number of falsely detected hotspots.

Effect of LGDD

- Performance comparison between LGDD and conventional density based feature.

Effect of LGDD

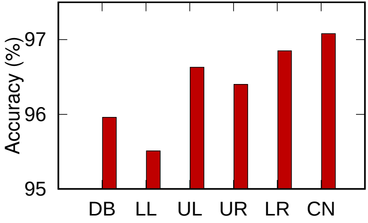
- Performance comparison between LGDD and conventional density based feature.

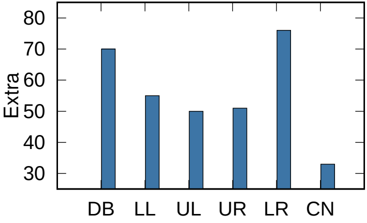


- The impact on accuracy.

Effect of LGDD

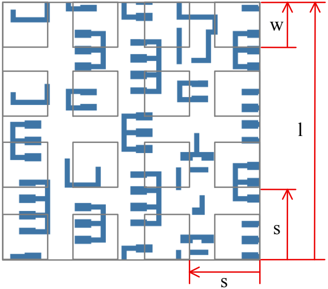
- Performance comparison between LGDD and conventional density based feature.



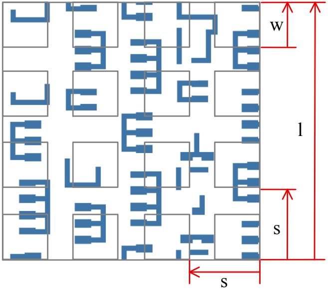


- The impact on accuracy.
- The impact on extra.

Stride Analysis

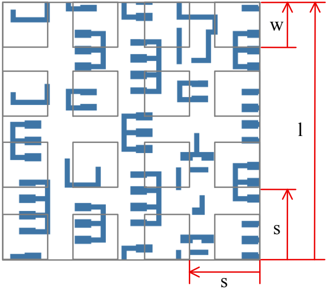


Stride Analysis



- Stride is the spacing between two adjacent grids (horizontally and vertically).

Stride Analysis



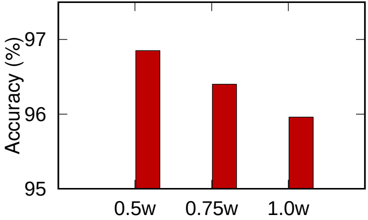
- Stride is the spacing between two adjacent grids (horizontally and vertically).
- Density based feature is a special case with $w = s$ in our stride analysis.

Effect of Strides

- Performance comparison among different strides.

Effect of Strides

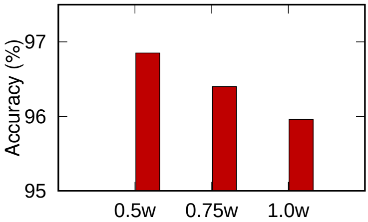
- Performance comparison among different strides.



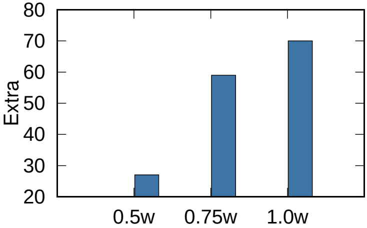
- The impact on accuracy.

Effect of Strides

- Performance comparison among different strides.



- The impact on accuracy.



- The impact on extra.

- The performance raises when shrinking the stride.
- However, after a threshold, the smaller of the stride, the worse of the performance.

Learning Model

Adaboost classifier

Require: $\mathbf{X} = (\mathbf{x_1}, ..., \mathbf{x_n})$, $\mathbf{Y} = (y_1, ..., y_n)$, T.

for $i \leftarrow 1$ to n **do**:

  $D_1(i) = \frac{1}{n}$;

for $t \leftarrow 1$ to T **do**:

  $h_t \leftarrow$ base classifier with small error ϵ_t;

  $\epsilon_t \leftarrow P(h_t(\mathbf{x_i}) \neq y_i) = \sum_{i=1}^{n} D_t(i) I(h_t(\mathbf{x_i}) \neq y_i)$;

  $\alpha_t \leftarrow \frac{1}{2}\log(\frac{1-\epsilon_t}{\epsilon_t})$;

  $Z_t \leftarrow 2[\epsilon_t(1-\epsilon_t)]^{\frac{1}{2}}$;

  for $i \leftarrow 1$ to n **do**:

    $D_{t+1}(i) \leftarrow \dfrac{D_t(i) exp(-\alpha_t y_t h_t(\mathbf{x_i}))}{Z_t}$;

$f \leftarrow \text{sign}(\sum_{t=1}^{T} \alpha_t h_t)$;

return f

- Decision-Tree as weak learner, more details in the paper.

Outline

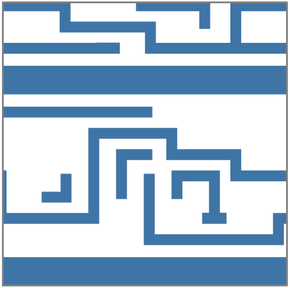
Problem Background

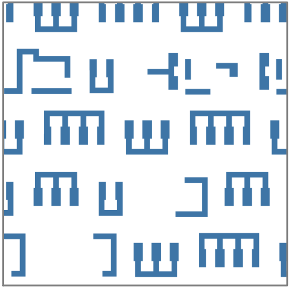
Conventional Methods

Our Method

Results

Benchmark Examples





- ICCAD benchmark.
- Industrial benchmark.

Effect of Our Methods

Table: Comparison with conventional density based method

	Density Based		Our Proposed	
	Extra#	Accuracy	Extra#	Accuracy
ICCAD-1	**0**	99.50%	2	**100.00%**
ICCAD-2	**0**	97.18%	**0**	**98.80%**
ICCAD-3	**0**	97.50%	1	**97.78%**
ICCAD-4	4	82.49%	5	**83.05%**
ICCAD-5	**0**	**95.12%**	**0**	**95.12%**
Industry	70	95.96%	**26**	**97.53%**
Average	12.3	94.63%	**5.6**	**95.38%**

- Consider both LGDD and stride analysis.
- Increase accuracy from 94.63% to 95.38%.
- Reduce the extra number from 12.3 to 6.

End

Thanks and Questions?