



Optical-System-Aware Feature Extraction for Lithography Hotspot Detection

Masahiro Yamamoto, **Masato Inagi**, Shinobu Nagayama
Graduate School of Information Sciences,
Hiroshima City University, Japan
Presenter: Masato Inagi, E-mail: inagi@hiroshima-cu.ac.jp

Masato Inagi – Short Resume

- Education

- B.E. in Computer Science, Tokyo Inst. of Tech. (TokyoTech), Japan, 2000
- M.E. in Integrated Systems and Communications, TokyoTech, 2002
- Ph.D. in Engineering, TokyoTech, 2008

- Professional Experience

- Researcher, University of Kitakyushu, Japan (2005-2008)
- Research Associate, Hiroshima City University, Japan (2008-2018)
- Lecturer (Assistant Prof.), Hiroshima City University (2018-present)

- Research Interests

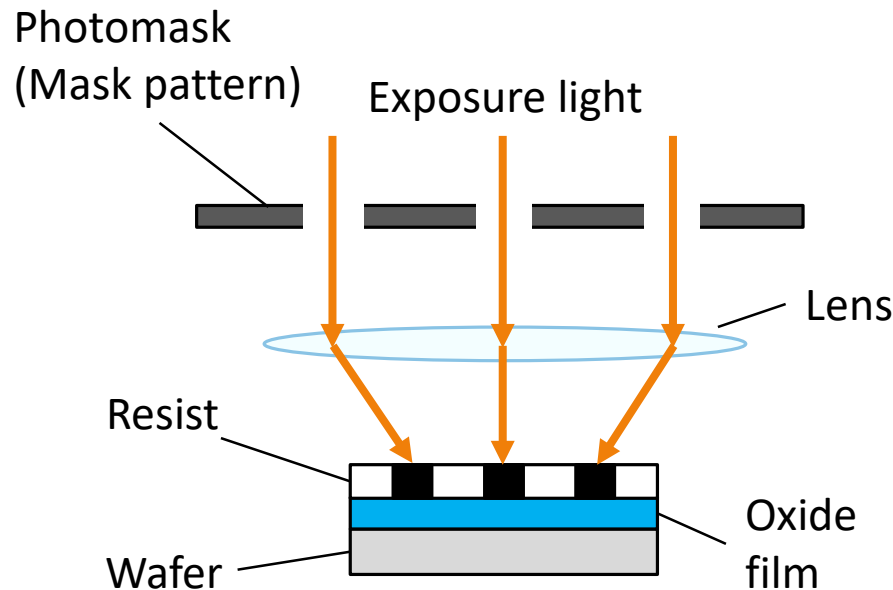
- VLSI Design
- Combinatorial Algorithms for VLSI Design Automation



Background

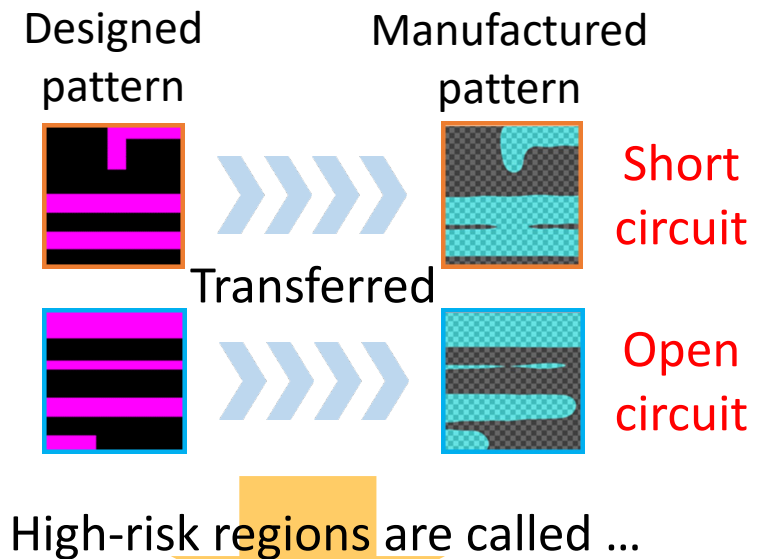
Lithography

transfers circuit patterns onto silicon wafers using light and photomasks.



As technology scales down ...

Pattern fidelity worsens due to increased light diffraction.

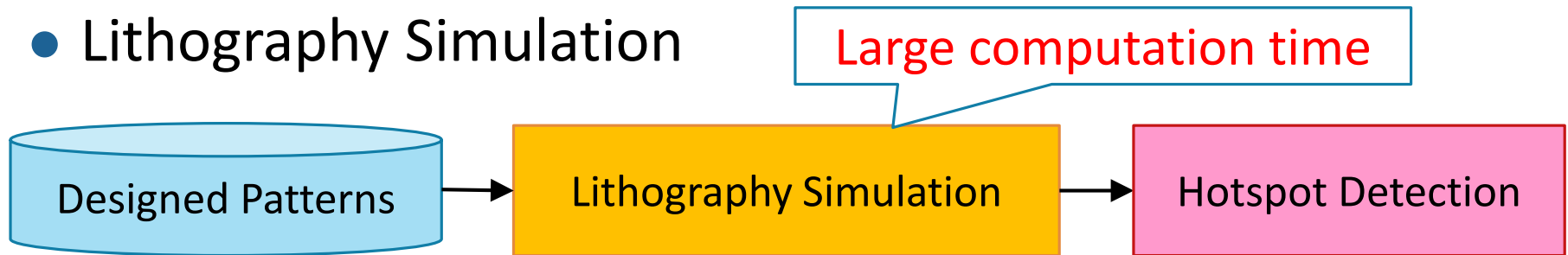


Lithography hotspots

Fast Hotspot Detection and Our Approach

Hotspot detection methodologies before manufacturing

- Lithography Simulation



- Existing approaches for faster hotspot detection
 - **Machine learning-based** (ML-based) methods have been proposed.
 - Various **features** for ML-based methods have been introduced to capture **the characteristics of hotspots**.

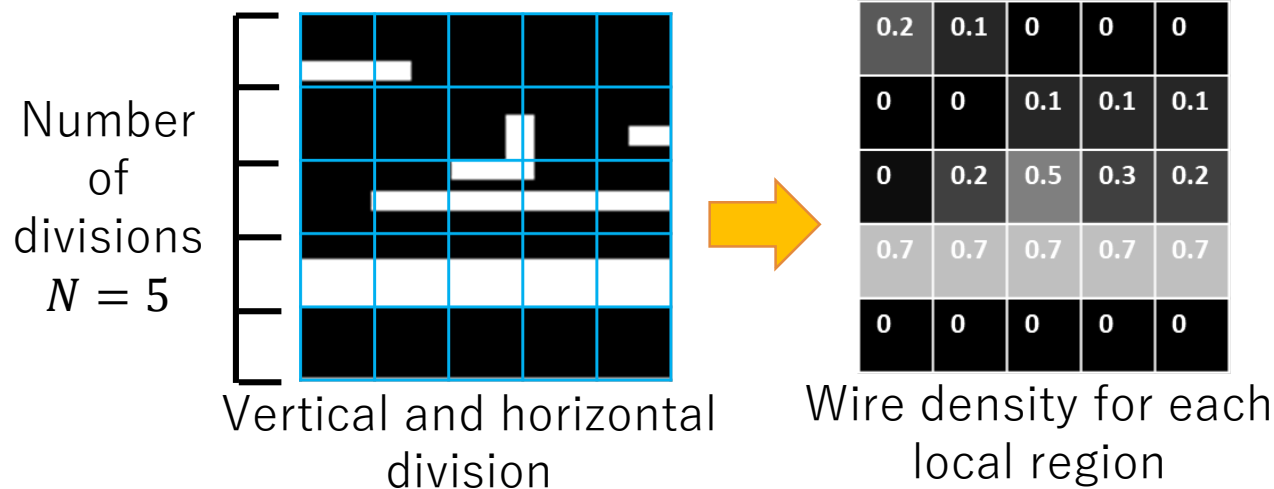
Our approach:

Features incorporating **the optical system characteristics**

Existing Features for Hotspot Detection

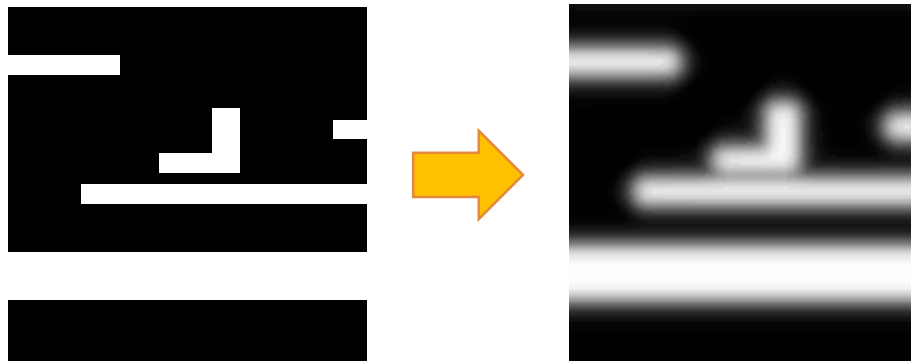
DBLF (Density Based Layout Feature)[1]:

captures local wiring density



HOLP (Histogram of Oriented Light Propagation)[2]:

captures approximate intensity gradient using Gaussian filtering



[1] Yu2015

[2] Tomioka2017

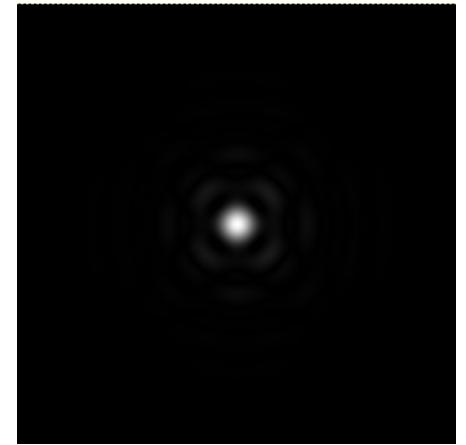
Our Approach: Motivation & Key Idea

- Hotspots depend on the optical conditions of the exposure system.
- Considering the optical characteristics of the exposure system may improve detection accuracy.

We propose a feature vector incorporating the optical system characteristics

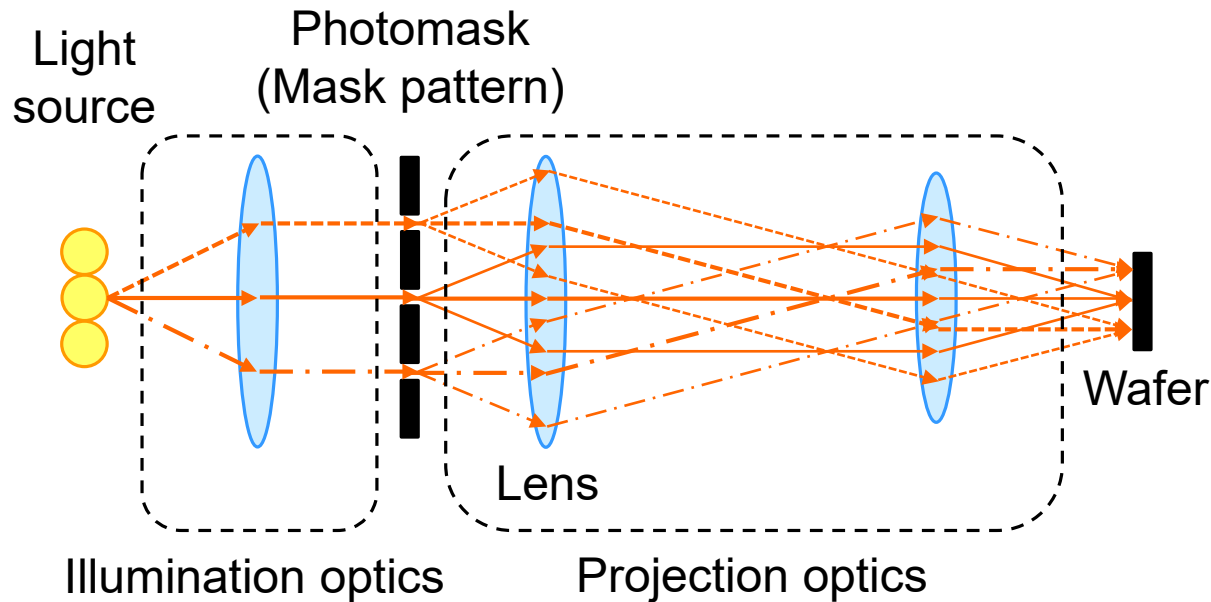
To achieve this:

- Use **SOCS kernel** from optical/lithography simulators
- **SOCS kernel** represents the optical characteristics of the exposure system as a set of 2D matrices.



Example of SOCS kernel
in space domain

Optical System Overview



Light intensity on wafer (SOCS model):

$$I(x, y) = \sum_{j=1}^n \sigma_j |(\boldsymbol{\phi}_j \star \mathbf{M})(x, y)|^2$$

$\mathbf{I}$: simulated intensity image,

$\boldsymbol{\phi}_j$: j -th SOCS kernel component (matrix),

$\mathbf{M}$: layout image

Convolution

Example: 25-order ($n = 25$) kernel

$\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \dots, \boldsymbol{\phi}_{25}$

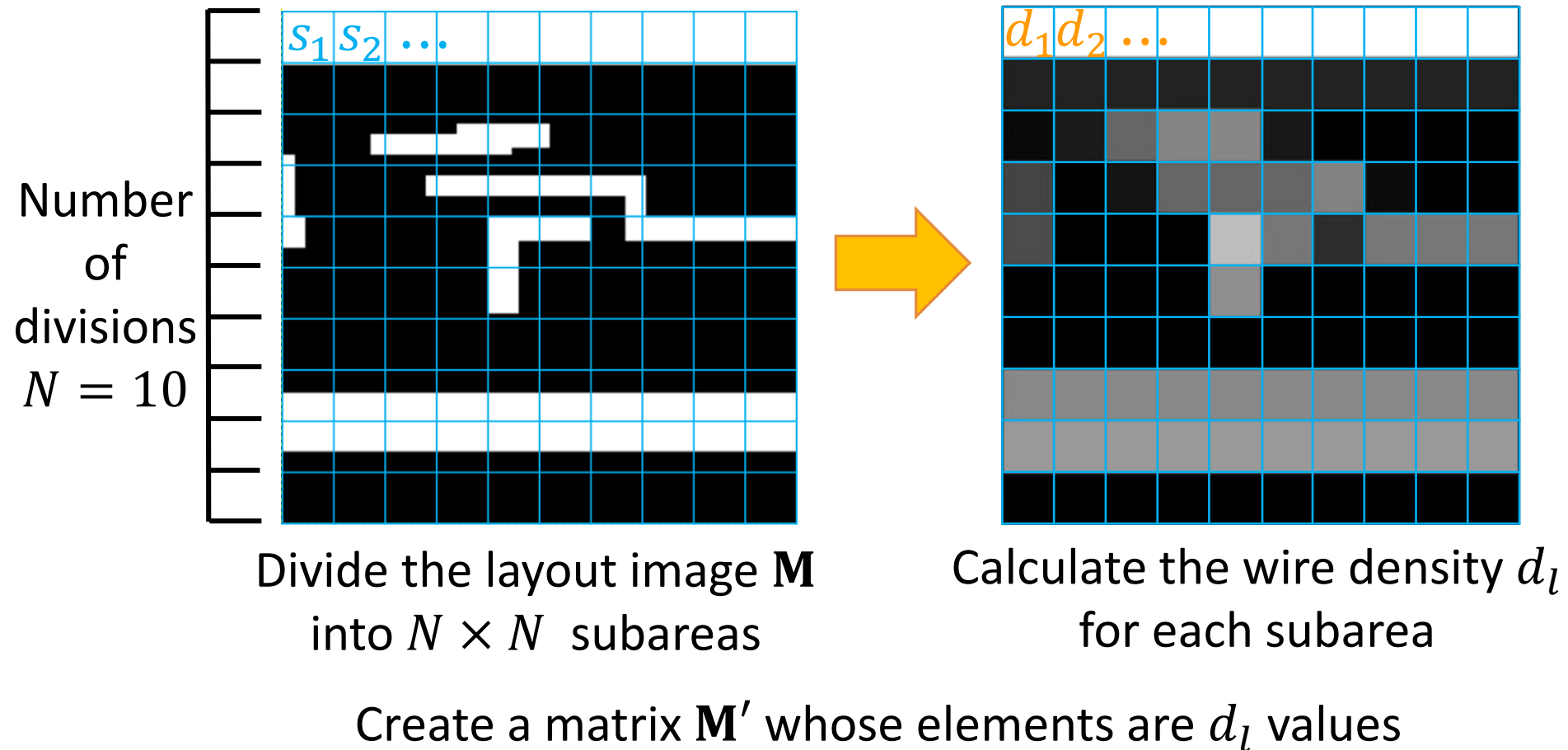
Representation of the optical system

Intensity = sum of squared
kernel-layout convolutions

[3] Cobb1995

Proposed Feature: Step 1

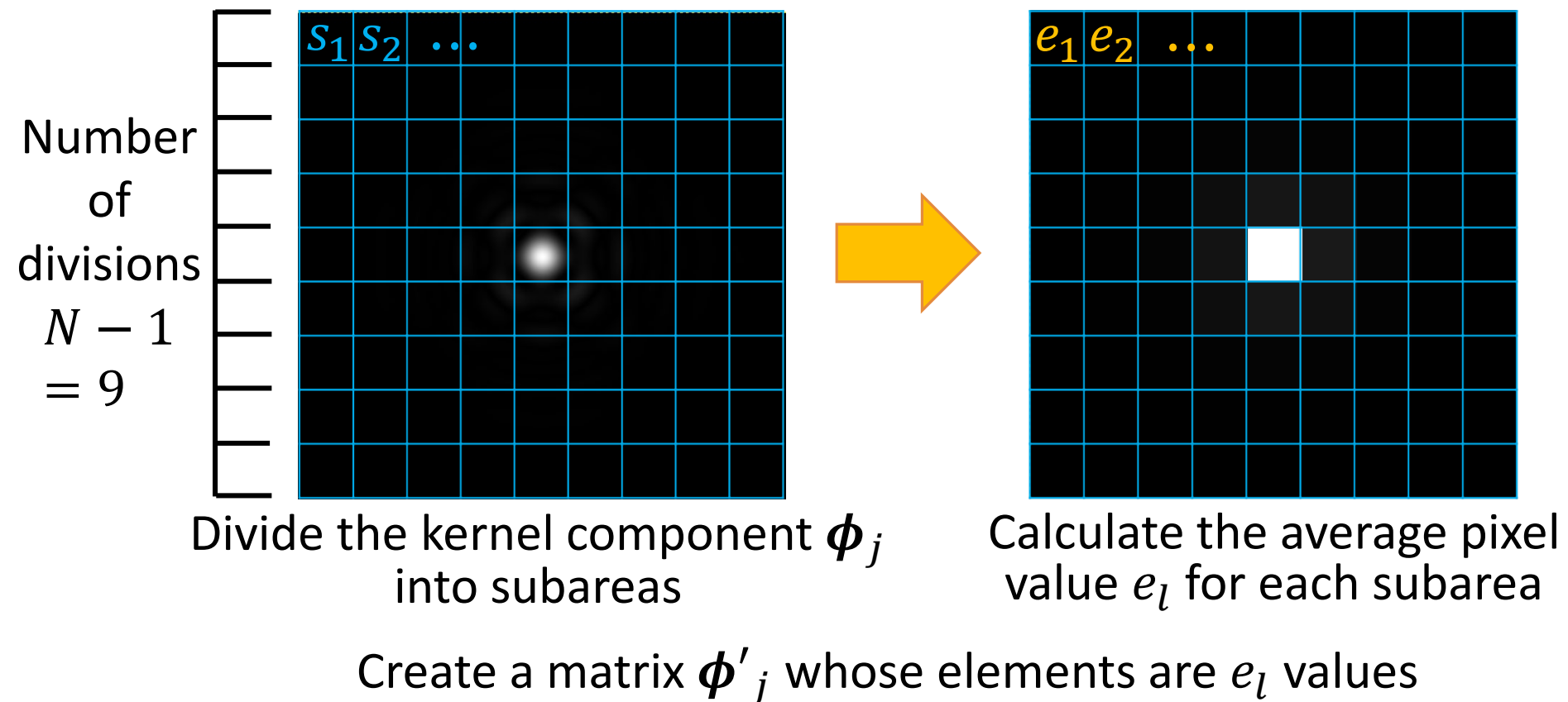
Divide the layout image into subareas and calculate the density of wiring for each subarea (local region in the paper) (same as DBLF)



Proposed Feature: Step 2

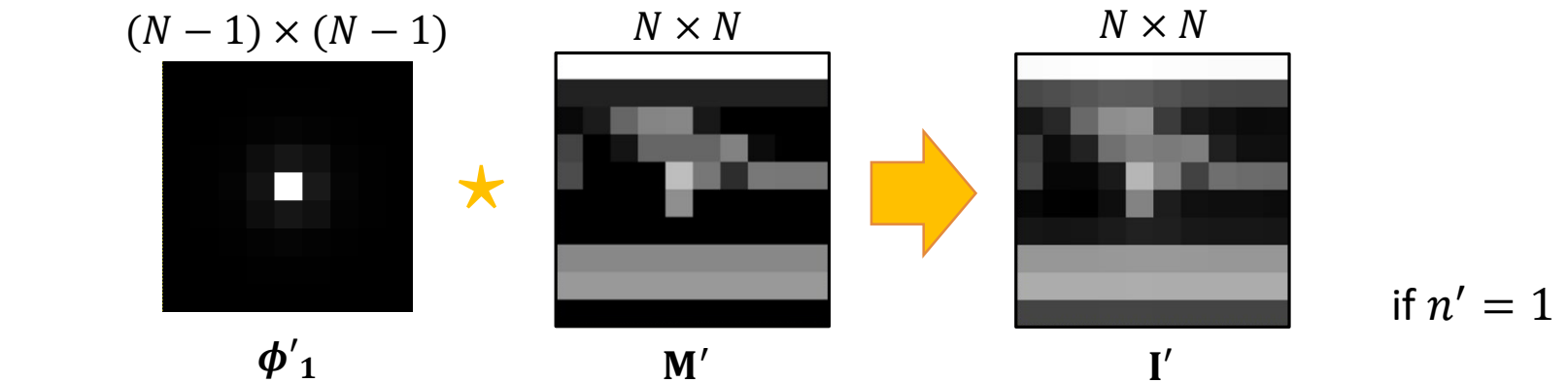
Divide the kernel into $(N - 1) \times (N - 1)$ subareas (if N is even) and compute the average of the pixel values for each subarea

Note: We assume that the layout image and the kernel have the same size. Peripheral kernel parts are clipped when N is even.



Proposed Feature: Step 3

Compute the convolution between the reduced layout matrix $\mathbf{M}'$ and the reduced kernel component ϕ'_j



Simplified light intensity:

$$I'(x, y) = \sum_{j=1}^{n'} \sigma_j \left| (\phi'_j * \mathbf{M}') (x, y) \right|^2 \quad n' \ll n$$

Extract the central $N \times N$ submatrix $\mathbf{C}$ from $\mathbf{I}'$

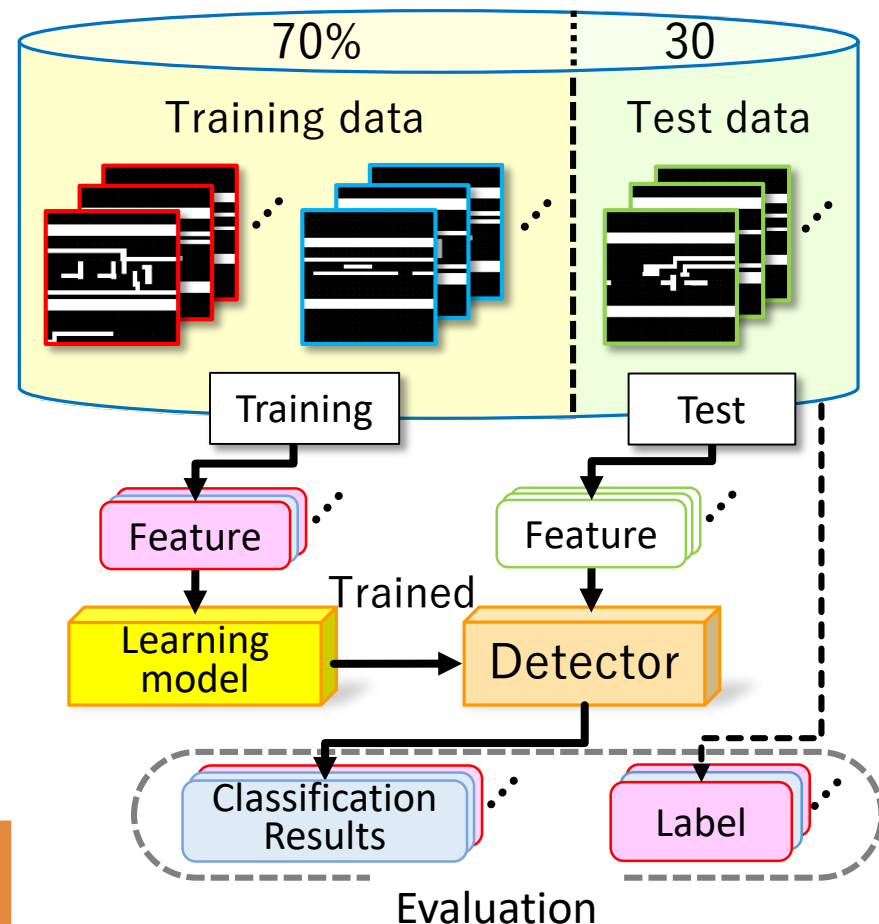
Our feature vector is

$$\mathbf{F}_{\text{OSAM}} = (C_{1,1}, C_{1,2}, \dots, C_{N,N-1}, C_{N,N})$$

Feature Evaluation Experiment

- Goal:
 - Evaluate hotspot detection accuracy
- Dataset:
 - ICCAD2012 [4] (data1-5)
 - re-labeled using ICCAD2013 optical simulator [5]
- Kernel: from ICCAD2013 simulator
- ML model: Real AdaBoost
 - decision trees as weak learners
- Hyperparameters:
 - all combinations tested

Number of weak learners	Decision tree depth	Learning rate
2, 4, 6, 8, 10, 20, 40, 60, 80, 100, 200, 400, 600, 800, 1000	1, 2, 3, 4	0.95, 0.96, 0.97, 0.98, 0.99



[4] Torres2012
[5] Banerjee2013

Experiment: Evaluation Metrics

- Precision (Pre): ratio of correct hotspot predictions
- Recall (Rec): ratio of detected hotspots
- F1 score (F1): Harmonic mean of Precision and Recall
 - Higher is better

	Actual HS	Actual N-HS
Predicted as HS	TP	FP
Predicted as N-HS	FN	TN

$$\text{Pre} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Rec} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{F1} = 2 \times \frac{\text{Pre} \times \text{Rec}}{\text{Pre} + \text{Rec}}$$

Comparison of Features: DBLF vs OSAM

Compared Features:

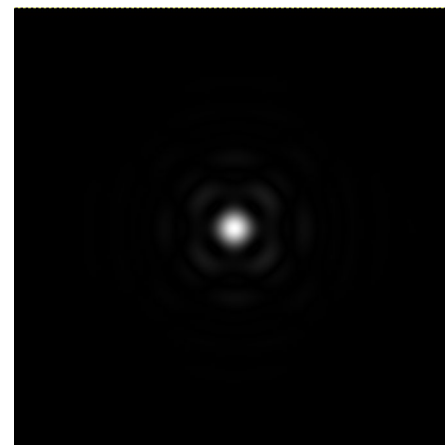
DBLF, OSAM (proposed)

Parameters of **DBLF**

- Number of divisions : 10
- Feature vector dimension: 100

Parameters of **OSAM** (proposed)

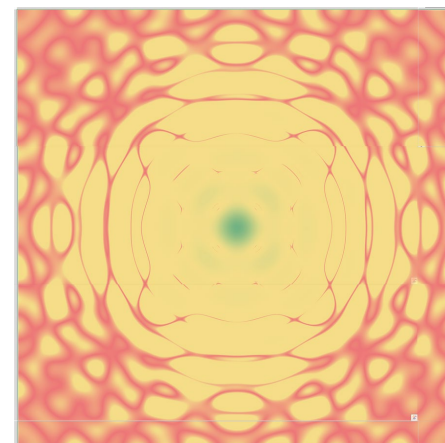
- Number of layout image divisions : 10
- Number of kernel divisions : 9
- Kernel order: 1
- Feature vector dimension 200
 - includes DBLF as part of OSAM



The kernel exhibits high light intensity at the center.

Kernel matrix ϕ_1 [5]

Color-coded by intensity



strong



weak

Kernel intensity shown with color for better visibility

Experimental Results: F1 score Comparison

Comparison between DBLF and OSAM (proposed)

Evaluation metric: F1 score (scaled $\times 10^2$, higher is better)

Dataset	Number of test data	Highest value of all parameter combinations		Average of all parameter combinations	
		DBLF	OSAM	DBLF	OSAM
data1	164	86.40	87.69	81.65	82.71
data2	1393	61.99	60.90	54.06	53.70
data3	1605	83.99	84.56	79.79	81.71
data4	1069	82.81	82.40	67.73	74.75
data5	646	79.93	82.60	67.18	75.15
average		79.02	79.63	70.08	73.60

The proposed feature demonstrates slightly better performance.
Overall F1 scores suggest room for improvement.

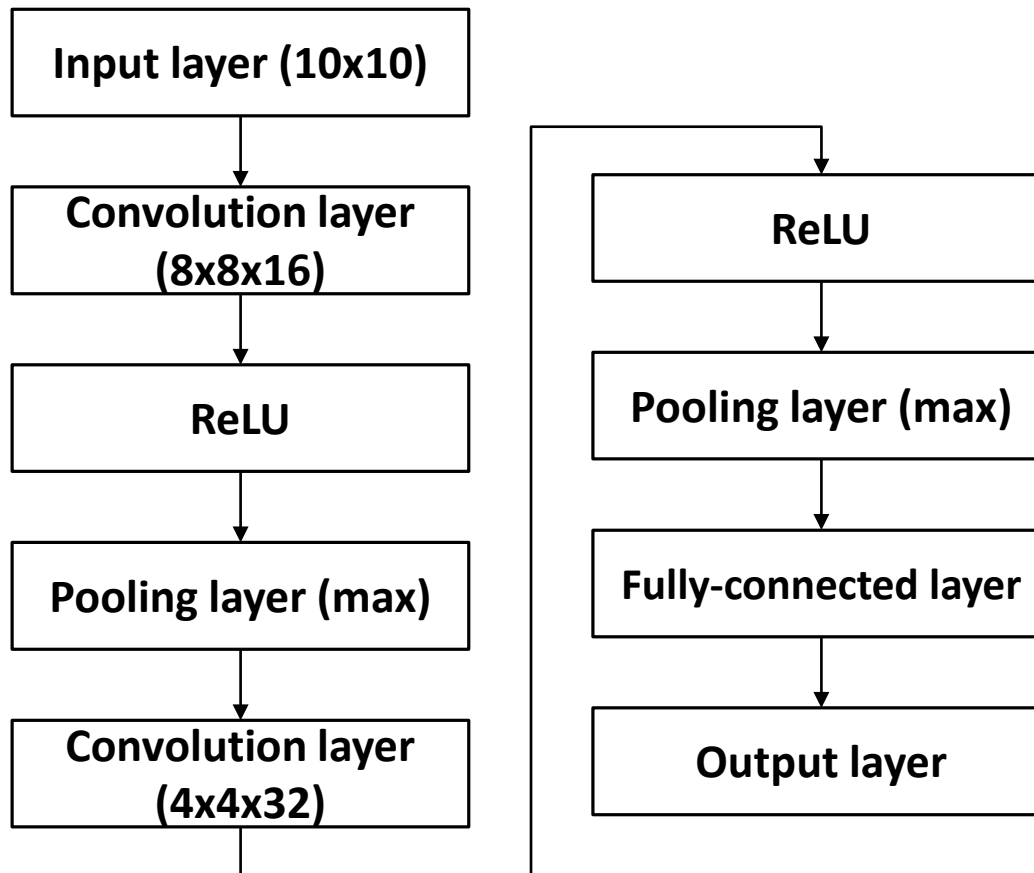
Analysis of Factors behind Low F1 scores

- Potential Contributing Factors
 - Factor 1: Problem instances may be inherently challenging.
 - Factor 2: Limited discriminative power of features
 - Factor 3: AdaBoost may be suboptimal for this task.
- Experimental Approach to Investigate Contributing Factors
 - Replacing AdaBoost with CNN for Comparative Evaluation
 - Lower F1 score than AdaBoost  Factor 1 / 2 ?
 - Higher F1 score than AdaBoost  Factor 3 ?

Note: The factors listed are not mutually exclusive and may interact in complex ways.

CNN Architecture Used for Comparison

CNN used in this experiment:



Filter size:

3x3

Number of filters:

1st conv. layer: 16

2nd conv. layer: 32

Pooling layer:

max pooling

Analysis of F1 score Trends Across Models

Comparison of DBLF and OSAM (proposed)

Evaluation metric: F1 score (scaled $\times 10^2$, higher is better)

CNN-Based Evaluation

Dataset	DBLF	OSAM
data1	87.74	86.62
data2	65.52	67.14
data3	85.31	86.97
data4	84.71	84.05
data5	78.80	81.26
average	80.42	81.20

AdaBoost-Based Evaluation

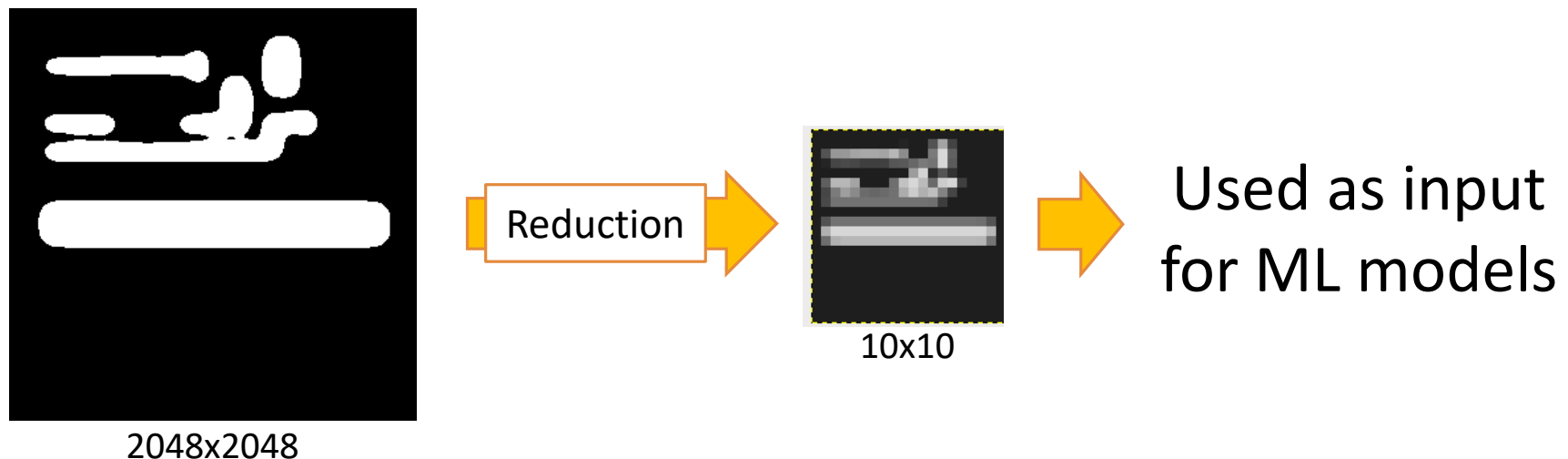
Dataset	DBLF	OSAM
data1	86.40	87.69
data2	61.99	60.90
data3	83.99	84.56
data4	82.81	82.40
data5	79.93	82.60
average	79.02	79.63

The proposed feature performed slightly better in both models.

Hotspot detection remains challenging for both features.

Exploring Improvements to the Proposed Feature

- We conducted experiments to examine whether detection accuracy could be improved by incorporating more optical system information:
 - As a trial, optical simulation images were used directly to maximize optical system information.
 - Models were trained using these optical simulation images.
 - The images were resized to match the dimensions of the proposed feature.



Experimental Results: Assessing Improvement by Incorporating More Optical System Characteristics

We investigated whether incorporating additional optical system information could enhance performance.

Evaluation metric: F1 score (scaled $\times 10^2$, higher is better)

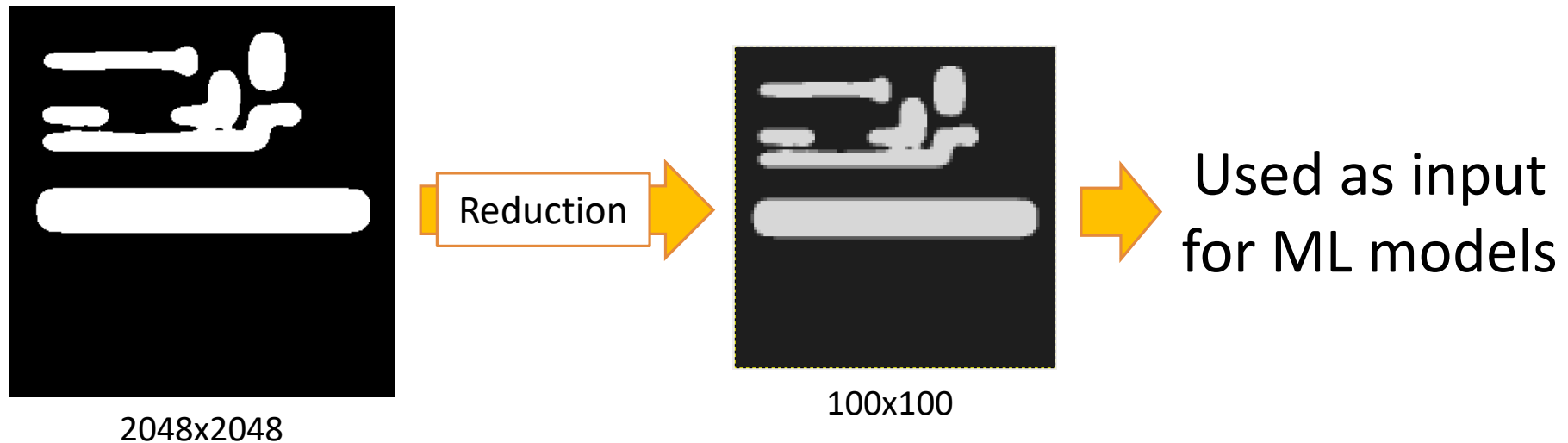
Dataset	AdaBoost		CNN	
	OSAM(10x10)	Sim(10x10)	OSAM(10x10)	Sim(10x10)
data1	87.69	88.39	86.62	87.74
data2	60.90	71.64	67.14	65.83
data3	84.56	87.35	86.97	86.59
data4	82.40	86.11	84.05	85.13
data5	82.60	84.41	81.26	84.86
average	79.63	83.58	81.20	82.03

Better than the
proposed feature

The proposed feature shows
potential for further improvement.

Exploring Improvement Potential by Increasing Feature Dimensions

- We conducted experiments to assess whether increasing the image size could improve performance.
 - Resize simulation images to a larger dimensions than before



Experimental Results: Impact of Increased Feature Dimension on Performance (AdaBoost)

We examined whether increasing feature dimension could improve performance.

Evaluation metric: F1 score (scaled $\times 10^2$, higher is better)

	Simulated image size after reduction					
Dataset	10x10	15x15	20x20	30x30	50x50	100x100
data1	88.39	88.85	88.49	88.48	87.72	87.71
data2	71.64	72.84	71.76	70.03	67.97	67.68
data3	87.35	88.23	87.26	86.88	86.65	86.17
data4	86.11	89.87	88.64	88.23	88.11	87.14
data5	84.41	86.13	83.29	83.72	82.67	82.25
average	83.58	85.18	83.89	83.47	82.63	82.19

F1 scores improved compared to 10x10

Performance tended to decrease as feature dimension increased beyond a certain point.

Experimental Results: Impact of Increased Feature Dimension on Performance (CNN)

We examined whether increasing feature dimension could improve performance.

Evaluation metric: F1 score (scaled $\times 10^2$, higher is better)

Dataset	Simulated image size after reduction		
	10x10	20x20	150x150
data1	87.74	87.01	93.88
data2	65.83	74.77	85.92
data3	86.59	89.72	90.21
data4	85.13	93.56	93.44
data5	84.86	90.65	92.50
average	82.03	87.14	91.19

Improved with increased feature dimensions

CNN outperformed AdaBoost.
Proposed feature shows potential for further improvement.

Conclusions and Future Work

Conclusions:

- Proposed feature OSAM for hotspot detection incorporates optical system.
- OSAM achieved slightly better accuracy than DBLF.

Future Work:

- Optimize kernel order
- Explore higher feature dimensions

References

- [1] Y.-T. Yu, et al., " Machine-learning-based hotspot detection using topological classification and critical feature extraction," in Proc. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol.34, pp.460-470, March 2015.
- [2] Y. Tomioka and T. Matsunawa, " Lithography hotspot detection by two-stage cascade classifier using histogram of oriented light propagation," in Proc. ASP-DAC 2017, pp. 81-86, 2017.
- [3] N. Cobb, "Sum of coherent systems decomposition by SVD," University of California, Berkeley, Technical Report, pp. 1-7, Sep. 1995.
- [4] J.-A. Torres, "ICCAD-2012 CAD contest in fuzzy pattern matching for physical verification and benchmark suite," in Proc. ICCAD 2012, pp. 349-350, Nov. 2012.
- [5] S. Banerjee, Z. Li, and S.-R. Nassif, "ICCAD-2013 CAD Contest in Mask Optimization and Benchmark Suite," in Proc. ICCAD 2013, pp. 271-274, Nov. 2013.