Lithography hotspot detection with ResNet network

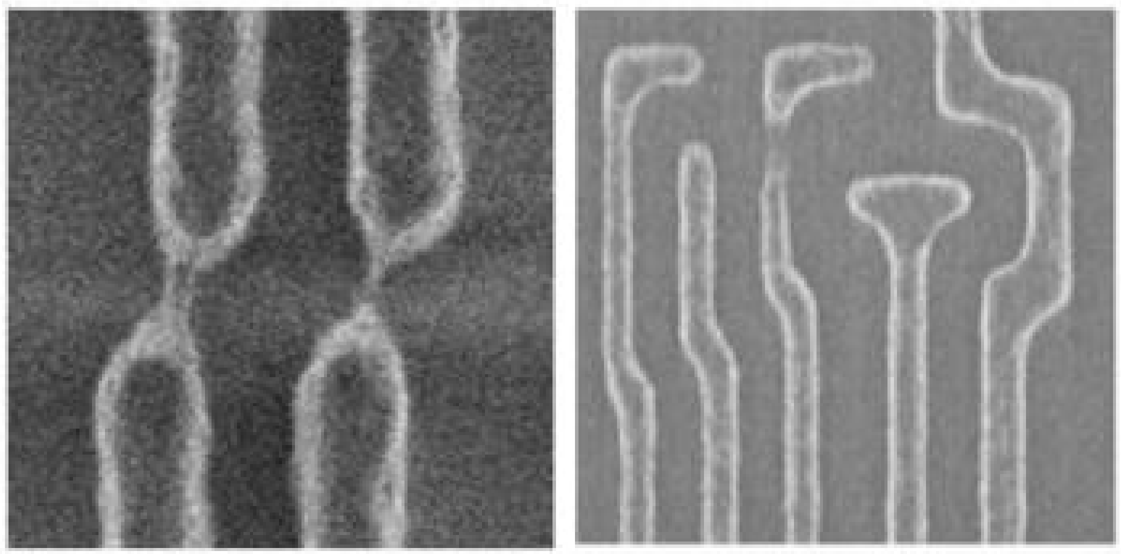
Yijiang Shen

CONTENTS

01 Introduction

Introduction

Due to the backward development of the light source technology, the mismatch between the wavelength of the light source and the characteristic size of the integrated circuit leads to the optical proximity effect (OPE), which can cause the distortion of the lithographic image. In order to improve the yield rate, various resolution enhancement technologies (RET) have been proposed in the industry. However, due to improper mask design or the limitations of RET technology itself, some parts of the circuit may still be broken or bridged, which is called photolithography hotspot.



Methods for detecting lithographic hotspots

lithography simulation

pattern matching

machine learning

deep learning

Our contributions

one	two	three
• residual network	• data enhancements	• loss function

02 Preliminaries

Preliminaries

In this section, we will mainly introduce the terms used and difficulties encountered in the task of hot spot detection in lithography.

terms

Accuracy: Correctly classify graphs with hotspots into hotspots category;

False alarm: Wrong classification of non hotspot graphics into hotspot category;

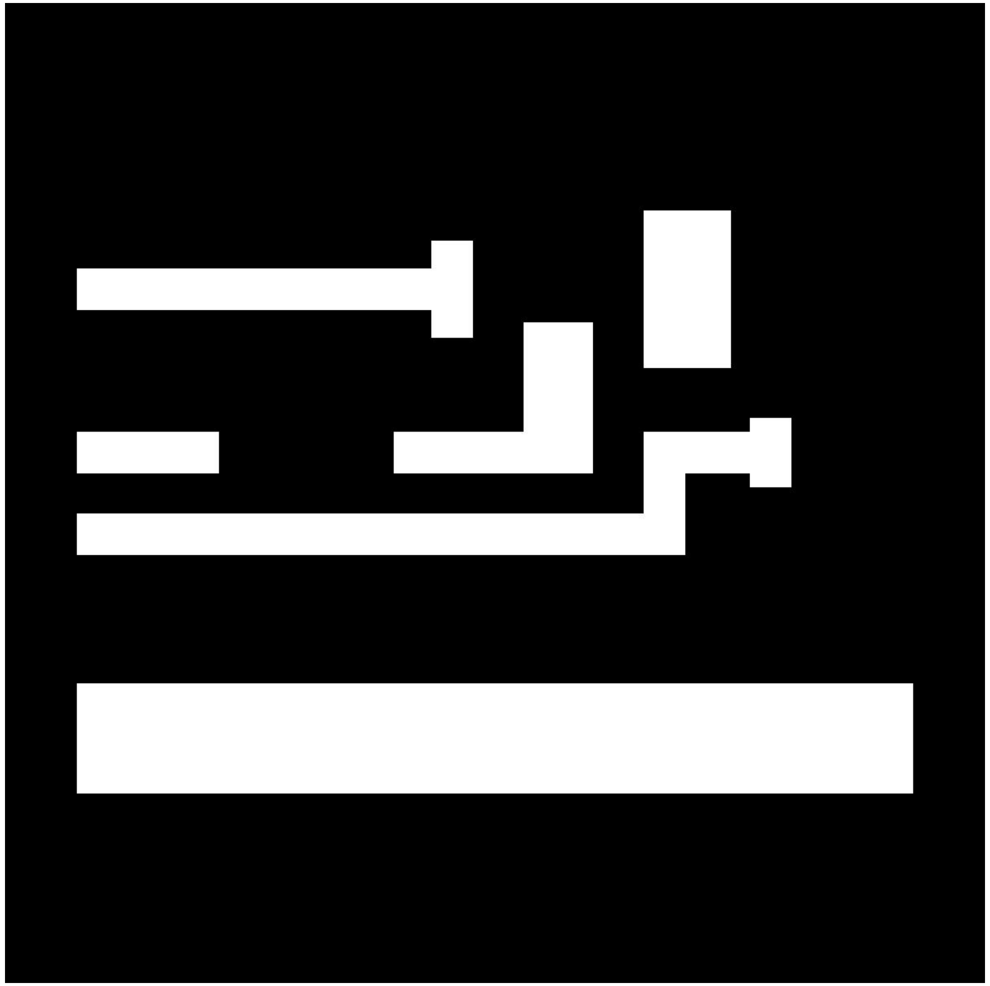
difficulties

issue 1.Data imbalance.
issue 2.The feature similarity

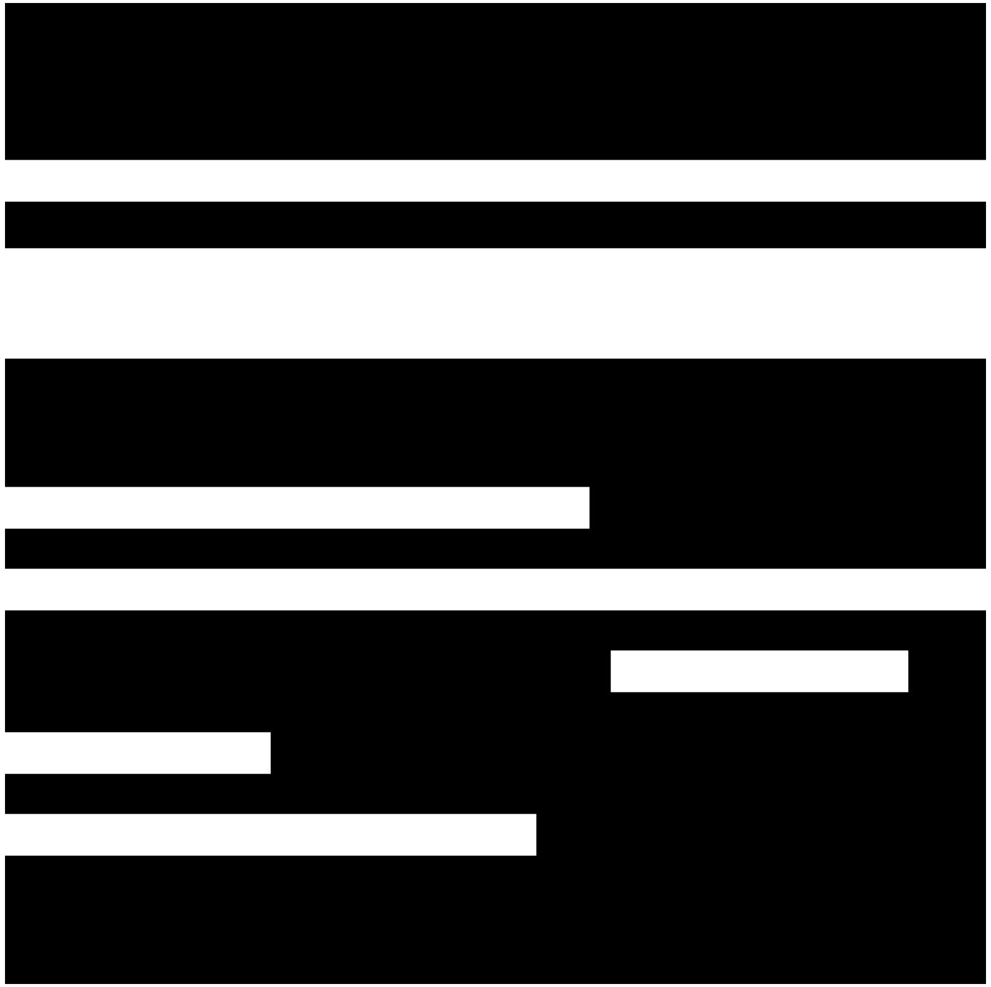
Preliminaries

name	TRAINING		TEST		TECH
	HS	NHS	HS	NHS	
benchmark1	99	340	226	4679	32nm
Benchmark2	174	5285	498	41298	28nm
Benchmark3	909	4643	1808	46333	28nm
Benchmark4	95	4452	177	31890	28nm
Benchmark5	26	2716	41	19327	28nm

Preliminaries



HS



NHS

03

Hotspots detection with ResNet

Hotspots detection with ResNet

01 **Data preparation**

02 **Network**

03 **Loss function**

Hotspots detection with ResNet

Data preparation

(a)

(b)

(e)

(f)

(c)

(d)

(g)

(h)

Network

Deeper networks are often more expressive and can learn more complex features. However, there are hidden training problems and possible network degradation in the deep level network, so it is obviously necessary to optimize the network. The introduction of residuals can make the network maintain its depth while having the advantages of shallow networks, and can effectively avoid network degradation, thus improving the optimization performance. The mathematical model is as follows:

$$\mathcal{H}(x) = \mathcal{F}(x) + x \Rightarrow \mathcal{F}(x) = \mathcal{H}(x) - x$$

x is the input of the network, H(x) is the output of the network, and F(x) is the learning ability of the network.

Hotspots detection with ResNet

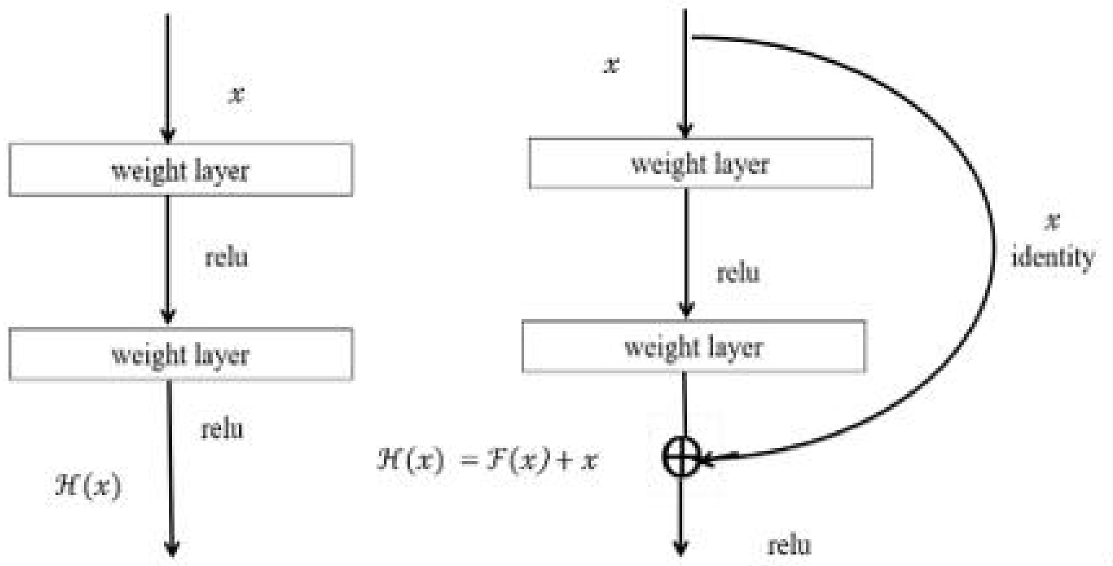


Fig. 2. (a) standard network structure. (b) residual network structure.

If the layer behind the deep network is an identity mapping, the model will degenerate into a shallow network. What needs to be solved now is how the network learns the identity mapping function. If the standard neural network in Fig 2.(a) learns to fit the identity mapping function H(x)= x, it will be more difficult. If the network is designed as the form H(x) = F(x) + x of the residual network in Fig 2.(b), As shown in (1) , an identity map H(x)= x is formed when F(x)=0, and fitting residuals is easier.

Loss function

The loss function is used to measure the degree of deviation between the prediction made by the model and the ground truth. With such reasoning, OHEM is applied to evaluate the deviation degree between the predicted value of the sample input to the network and the ground truth, and select the samples with the largest deviation degree as they intend to have a greater impact on classification and detection. These sample data are trained with stochastic gradient descent to learn the characteristics of hotspots and non-hotspots, with improved convergence of the model.

04 Experimental result

Experimental result

This experiment is implemented with a six core 3.6GHz CPU, Nvidia GTX 1080Ti and 16G memory. The proposed framework is verified in the data set of ICCAD2012 CAD Competition. We use two performance indicators, false alarm rate and recall rate, and used the residual network of resnet34 network to learn to identify lithography hotspot, and the structure of resnet34 is shown in Fig. 3 and Table 2.

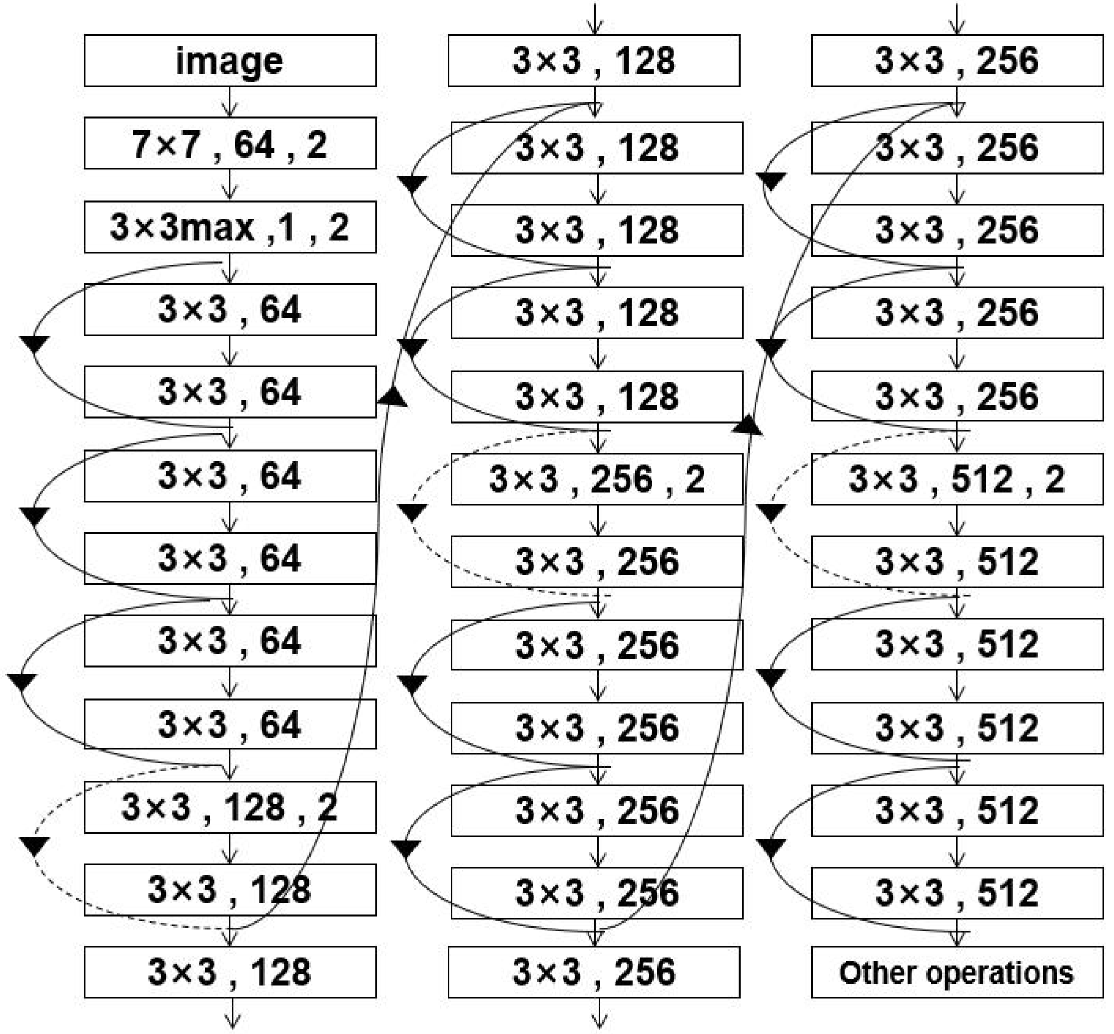


Fig. 3. Structure of resnet34.

TABLE II. Structure of resnet34

layer name	out size	34-layer
conv 1	112×112	7×7, 64, stride 2
conv 2_x	56×56	3×3 max pool, stride 2
		$\begin{pmatrix} 3\times3, & 64 \\ 3\times3, & 64 \end{pmatrix}\times3$
conv 3_x	28×28	$\begin{pmatrix} 3\times3, & 128 \\ 3\times3, & 128 \end{pmatrix}\times4$
conv 4_x	14×14	$\begin{pmatrix} 3\times3, & 256 \\ 3\times3, & 256 \end{pmatrix}\times6$
conv 5_x	7×7	$\begin{pmatrix} 3\times3, & 512 \\ 3\times3, & 512 \end{pmatrix}\times3$
	1×1	average pool ,2-d fc, SoftMax

The training performance on benchmark5 is presented in Fig. 5 and Fig. 6 as an example with the proposed RESNET model. In Fig.4, accuracy, precision and recall when training benchmark5 starts to converge at epoch 4 with a high level. Figure 5 shows that the training loss drops rapidly before epoch 3 and gradually stabilizes at epoch 6. Experimental results performed on benchmarks 1-5 with Yu, Zhang, Yang methods and the proposed method are given in Table 3 , respectively. It is duly observed that while all methods are capable of hotspot detection with comparable high recall rate, the proposed model retrieves low false alarm rate, which effectively reduce the effort of repair process.

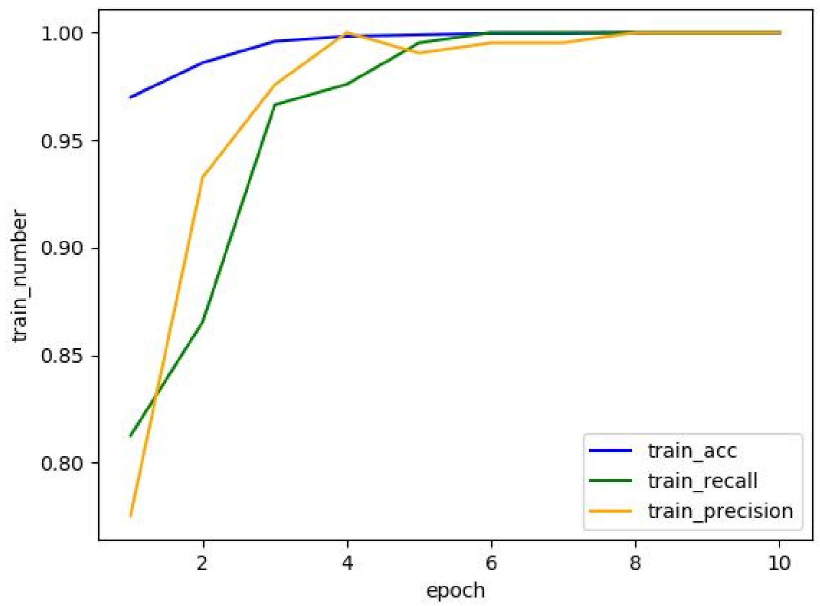


Fig. 5. Training performance on benchamark5.

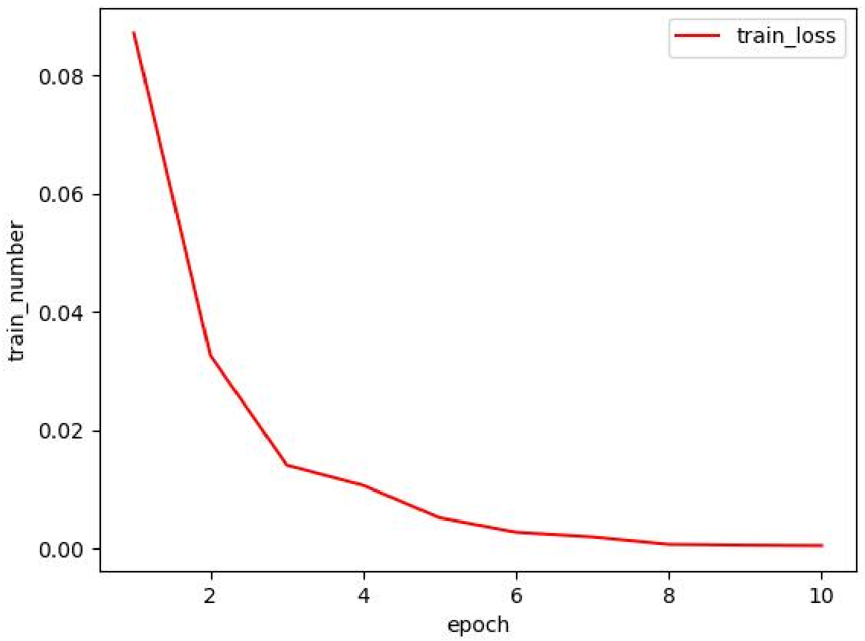


Fig. 6. Training loss on benchamark5.

Experimental result

Table3:The other results are shown in Table 2.

	YU		ZHANG		Yang		Ours	
	FA	RECALL	FA	RECALL	FA	RECALL	FA	RECALL
benchmark1	1493	94.7	788	100	147	99.6	1761	99.6
benchmark2	11834	98.2	544	99.4	561	99.8	117	98.2
benchmark3	13850	91.9	2052	97.5	2660	97.8	3113	98.1
benchmark4	3664	85.9	3354	97.7	1785	96.4	359	93.9
benchmark5	1205	92.9	94	95.1	242	95.1	248	98.6

Conclusion

Conclusion

We propose a photolithography hotspot model based on convolutional neural network, introduce residual network, use a variety of data enhancement forms, and introduce OHEM algorithm for data imbalance. The experimental results show that our method can maintain a high recall rate with low false alarm.

THANKS