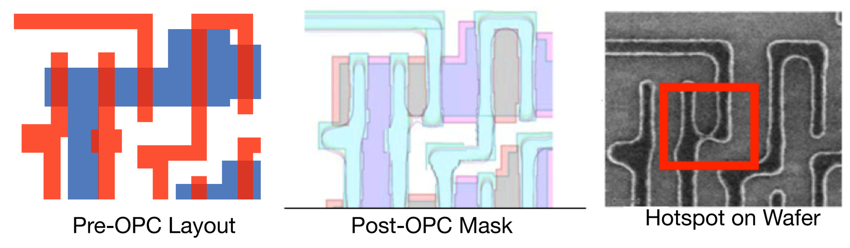
Faster Region-based Hotspot Detection

Ran Chen[1], Wei Zhong[2], Haoyu Yang[1], Hao Geng[1], Xuan Zeng[3], Bei Yu[1]

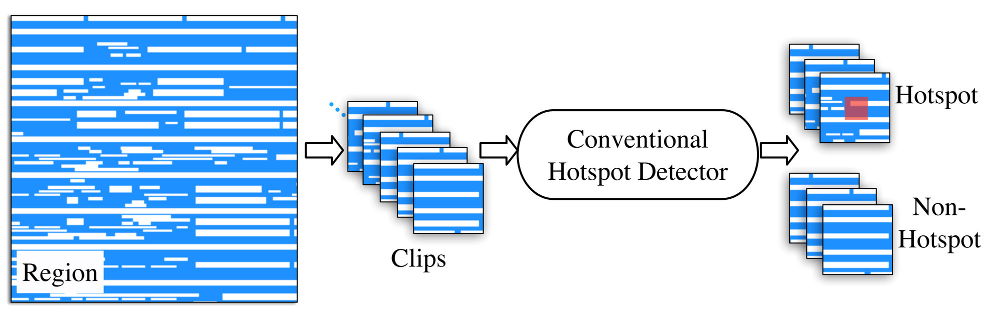
[1]The Chinese University of Hong Kong
[2]Dalian University of Technology
[3]Fudan University

Background



- What you see ≠ what you get
- Diffraction information loss
- RET: OPC, SRAF, MPL
- Worse on designs under 10nm or beyond

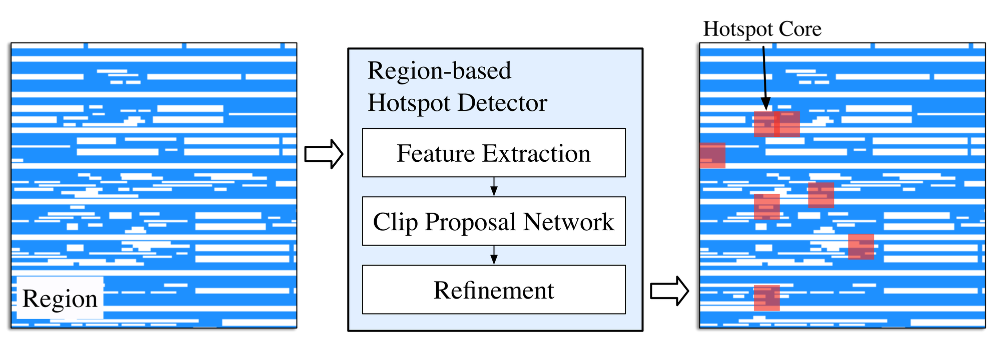
Previous Solutions



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

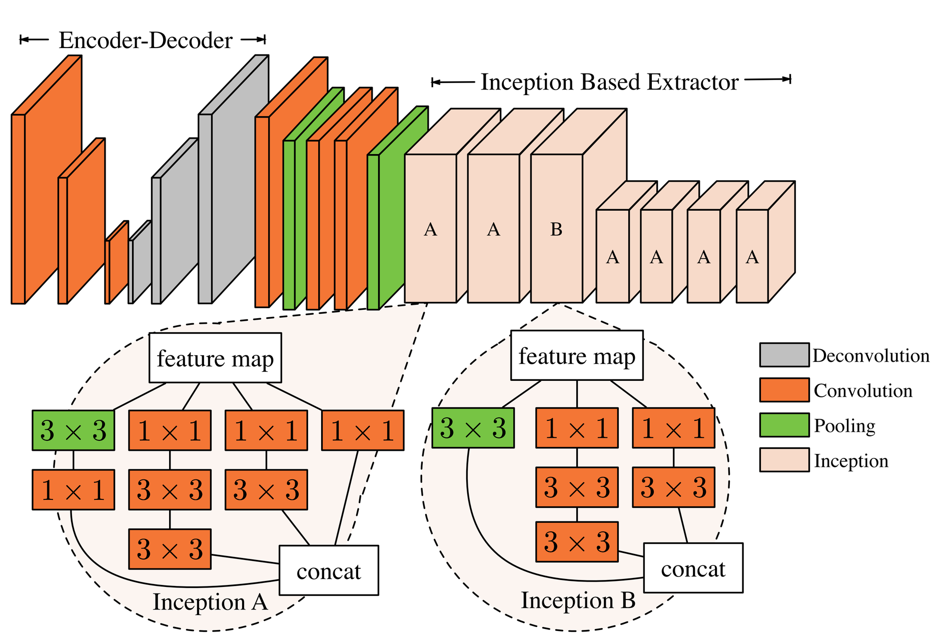
- Scanning is time consuming and single stage is not robust to false alarm.

Region Based Approach



- Learning what and where is hotspot at same time.
  - Multi-task on Classification and Regression.

Feature Extraction



- Encoder-decoder preprocess
  - Symmetric Structure for feature encoding and decoding.
  - Much faster than discrete cosine transformation.

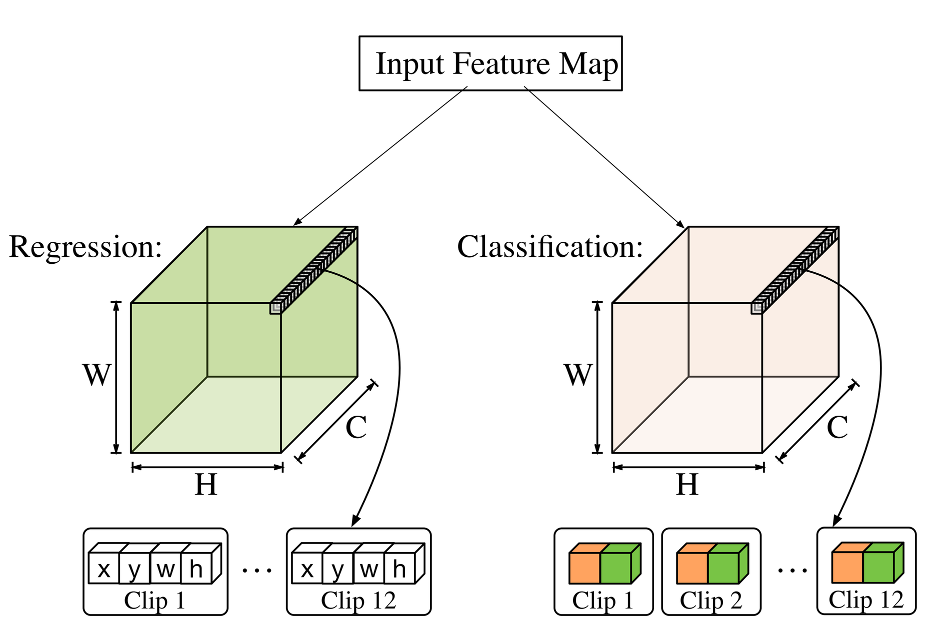
- Inception based structure
  - Multi thread feature extraction.
  - Prune the depth of the output channel for each stage.
  - Downsample the feature map size in height and width direction.

Clip Proposal Network

Definition

- Clip: Predefined box to crop hotspot features in region.
- Proposal: Selected clip which contribute to classification and regression.

- Based on extracted features, Clip Proposal Network is designed to locate and classify hotspots.
- Classification and regression branches share features.



Details on Clip Proposal Network

- To a classifier, we have to balance the positive and negative samples.
- As a regression task on location, we need to select reasonable clips as proposals.
- We also need to consider efficiency and quality of features.

Clip Pruning

Intersection over Union (IoU)

$$\mathrm{IoU} = \frac{clip_{groundtruth} \bigcap clip_{generated}}{clip_{groundtruth} \bigcup clip_{generated}}.$$

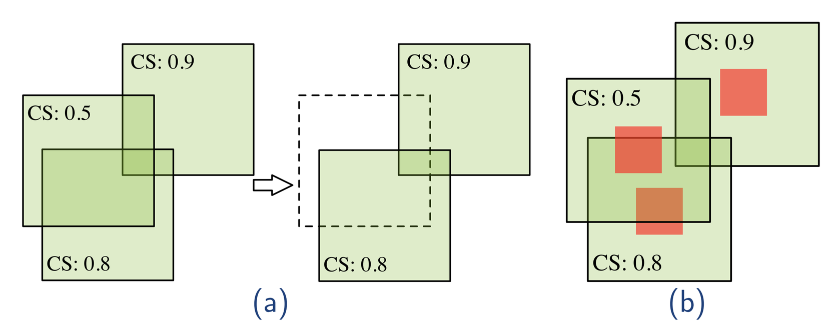
- Clip generation: generate group of clips with different aspect ratios and scales in dense.
  - Number of clips: $w * h * clips\ per\ location$

- Clip Pruning before Classification and Regression.
  - IoU > 0.7, reserved as positive sample;
  - IoU with any ground truth highest score should be reserved as positive sample;
  - IoU < 0.3, reserved as negative sample;
  - Rest of clips do no contribution to the network training.

hotspot non-maximum suppression

```
sorted_ws ← sorted clip set;
k ← size of clip set;
for i ← 1, 2, ..., k do
    current_w ← sorted_ws[i];
    for j ← i, i + 1, ..., k do
        compared_w ← sorted_ws[j];
        Overlap ← Centre_IoU(current_w, compared_w);
        if Overlap > threshold then
            Remove compared_w; k ← k − 1;
        end if
    end for
end for
return sorted_ws;
```

- Comparison with conventional non-maximum suppression



- Examples of (a) conventional non-maximum suppression, and (b) the proposed hotspot non-maximum suppression.
- Hotspot non-maximum suppression takes advantage of the structural relation between core region and clips which avoid the error dropout during the training.

Loss Function Design

Parameterizations of coordinates

- Origin parameters may affect the training stability.

$$\begin{aligned} l_x &= (x - x_g)/w_g, & l_y &= (y - y_g)/h_g, \\ l'_x &= (x' - x_g)/w_g, & l'_y &= (y' - y_g)/h_g, \\ l_w &= \log(w/w_g), & l_h &= \log(h/h_g), \\ l'_w &= \log(w'/w_g), & l'_h &= \log(h'/h_g), \end{aligned} \quad (1)$$

Classification and Regression Loss

- L2 regularization penalizes peaky weight vectors and prefers diffuse weight vectors, which has appealing property of encouraging the network to use all of its inputs rather than skewed on partial of its inputs.

$$\begin{aligned} L_{C\&R}(h_i, l_i) = &\alpha_{loc} \sum_i h'_i l_{loc}(l_i, l'_i) + \sum_i l_{hotspot}(h_i, h'_i) \\ &+ \frac{1}{2}\beta(\|\boldsymbol{T}_{loc}\|_2^2 + \|\boldsymbol{T}_{hotspot}\|_2^2), \end{aligned} \quad (2)$$

- Smooth L1 loss for robust regression, which makes gradient smooth when offset is small:

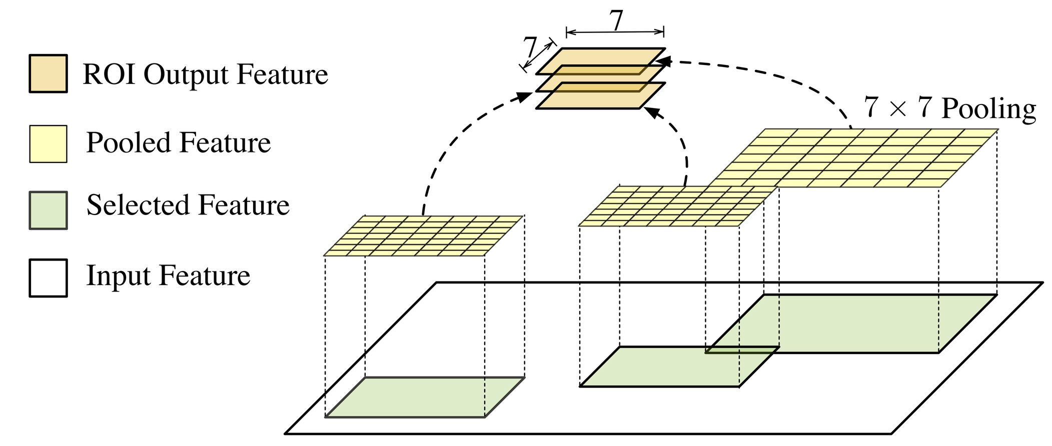
$$l_{loc}(l_i[j], l'_i[j]) = \begin{cases} \frac{1}{2}(l_i[j] - l'_i[j])^2, & \text{if } |l_i[j] - l'_i[j]| < 1, \\ |l_i[j] - l'_i[j]| - 0.5, & \text{otherwise,} \end{cases} \quad (3)$$

- Cross Entropy loss:

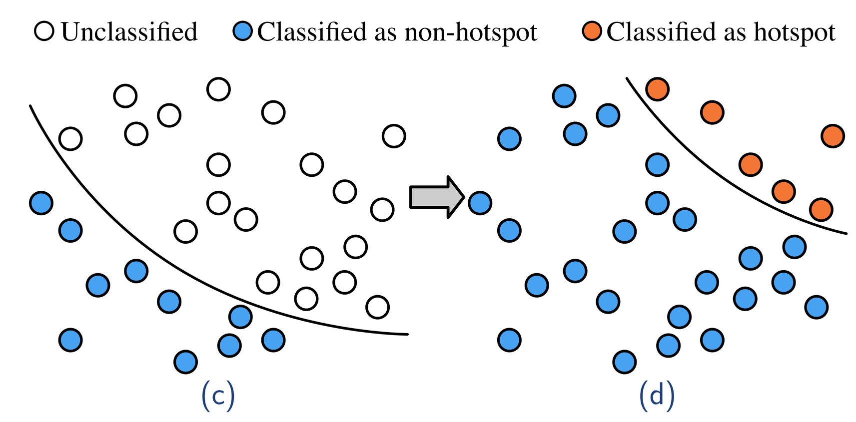
$$l_{hotspot}(h_i, h'_i) = -(h_i \log h'_i + h'_i \log h_i). \quad (4)$$

RoI Pooling

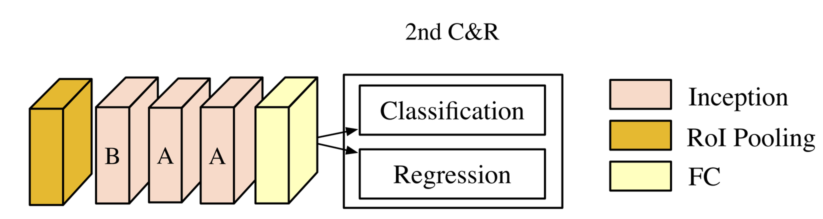
- Classified clips have different sizes.
- We need resize them to same size for second stage refinement.



Refinement



- (a) 1st hotspot classification in clip proposal network;
- (b) The labelled hotspots are fed into 2nd hotspot classification in refinement stage to reduce false alarm.



- We got a rough prediction after clip proposal network regression and classification.
- Refinement stage is applied to further decrease the false alarm and improve accuracy.

Experimental Results

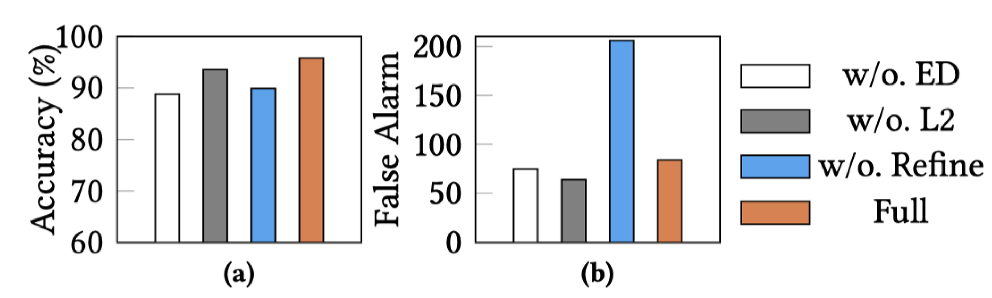
Comparison with State-of-the-art

- ICCAD CAD Contest 2016 Benchmarks
- Three different design styles
- 45 times faster and 6.14 % accuracy improvement compare to [Yang,TCAD'18].
- Much better than two well known object detecton based frameworks.

Table 1: Comparison with State-of-the-art

Bench	TCAD'18 [Yang,TCAD'18]			Faster R-CNN Ren,NIPS'15			SSD[Liu,ECCV'16]			Ours		
	Accu (%)	FA	Time (s)	Accu (%)	FA	Time (s)	Accu (%)	FA	Time (s)	Accu (%)	FA	Time (s)
Case2	77.78	48	60.0	1.8	3	1.0	71.9	519	1.0	93.02	17	2.0
Case3	91.20	263	265.0	57.1	74	11.0	57.4	1730	3.0	94.5	34	10.0
Case4	100.00	511	428.0	6.9	69	8.0	77.8	275	2.0	100.00	201	6.0
Average	89.66	274.0	251.0	21.9	48.7	6.67	69.0	841.3	2.0	95.8	84	6.0
Ratio	1.00	1.00	1.00	0.24	0.18	0.03	0.87	3.07	**0.01**	**1.07**	**0.31**	0.02

Ablation Study



- Comparison among different settings
- (a) average accuracy and (b) average false alarm.

Visualized Result

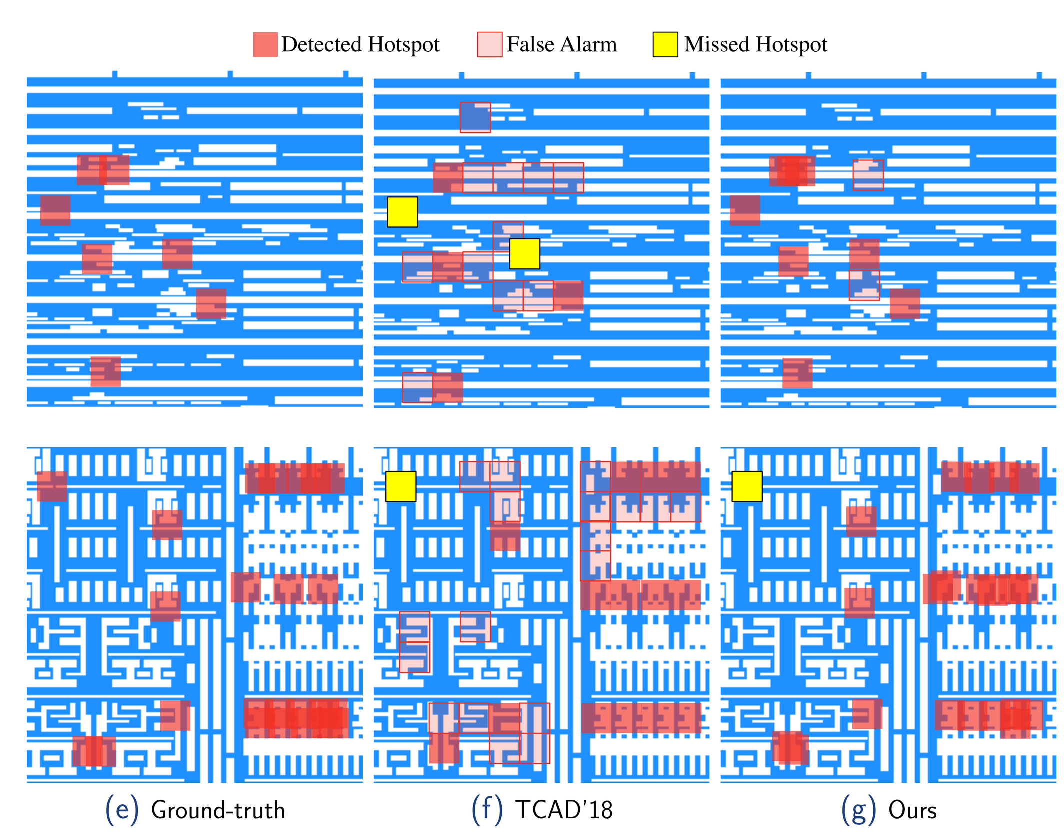


Figure 1: Visualization of different hotspot detection results.