PowerNet: Transferable Dynamic IR Drop Estimation via Maximum Convolutional Neural Network

Zhiyao Xie*, Haoxing Ren+, Ye Sheng+, Santosh Santosh+, Brucek Khailany+, Jiang Hu± and Yiran Chen*

*Duke University, +Nvidia, ± Texas A&M University

Outline

- Introduction
- Method
- Results
- Discussion
- Conclusion

Introduction: IR Drop Problem

- Dynamic IR drop is voltage drop at power pin of a cell
- Problem
  - IR drop increases cell delay and may cause time violation
  - IR drop becomes more serious as technology node shrinks
- Reason for IR drop
  - V=IR, high current demand, high resistive path
  - High cell and power density (spatial)
  - Timing window overlap (temporal)

Introduction: Previous Works

- Design independency:
  - model **transferable** to new designs that not in training set
- Vector-based IR drop compared with Vectorless IR drop:
  - requires many simulation patterns to cover most regions, can be very slow
  - accurate power simulation patterns not available in early design process
  - much easier to be estimated by ML models (will be discussed)

ML Methods	Model	Design Independent
Yamato et al. (ITC 12)	Linear Regression	No
Ye et al. (VTS 14)	SVM	No
Dhotre et al. (ATS 17)	Clustering	Unsupervised
Lin et al. (VTS 18)	ANN	No
Fang et al. (ICCAD 18)	XGBoost	No
PowerNet	Max-CNN	Yes

PowerNet can be design independent

PowerNet solve both vector-based & vectorless IR predictions, while focusing on vectorless scenario.

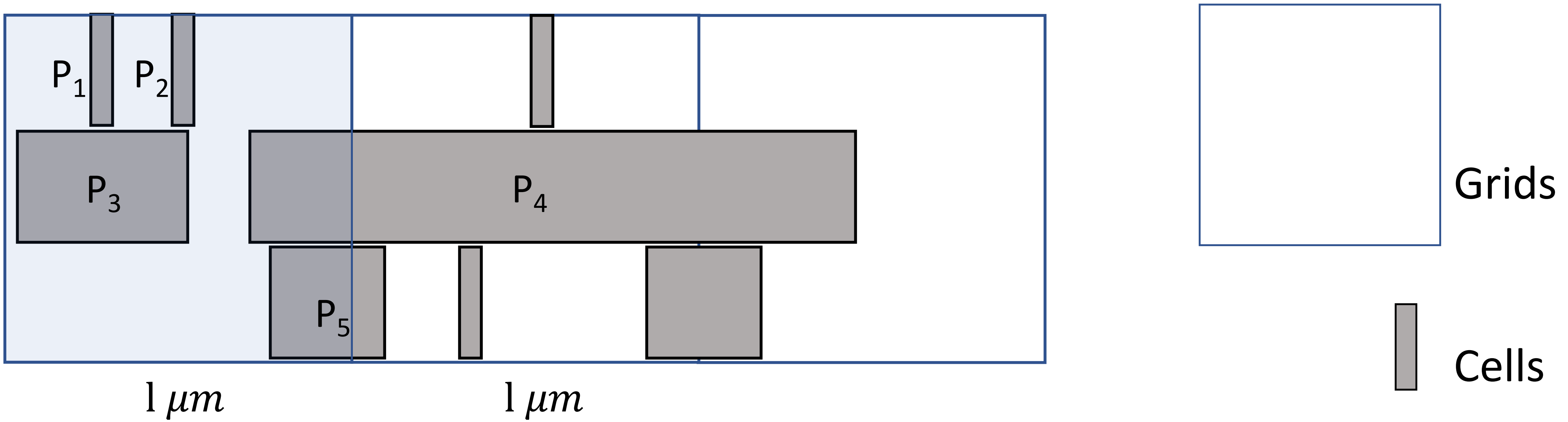
Method: Features Extraction for Each Cell

- Power: Three types of power values are extracted.
  - Internal power (p_i)
  - Switching power (p_s)
  - Leakage power (p_l)

Overall power: $p_{all} = p_i + p_s + p_l$
Scaled overall power: $p_{sca} = r_{tog}*(p_i + p_s) + p_l$
Resistance is assumed uniform (will be discussed)

- Coordinates: The cell location after placement.
  - Min and max x axis (x_{min}, x_{max})
  - Min and max y axis (y_{min}, y_{max})
- Signal arrival time: The min and max signal arrival times in one clock cycle.
  - Min arrival time (t_{min})
  - Max arrival time (t_{max})
- Toggle rate: how often output changes with regard to a given clock input. (r_{tog})
- IR drop value as label. (IR)

Method: Space Decomposition

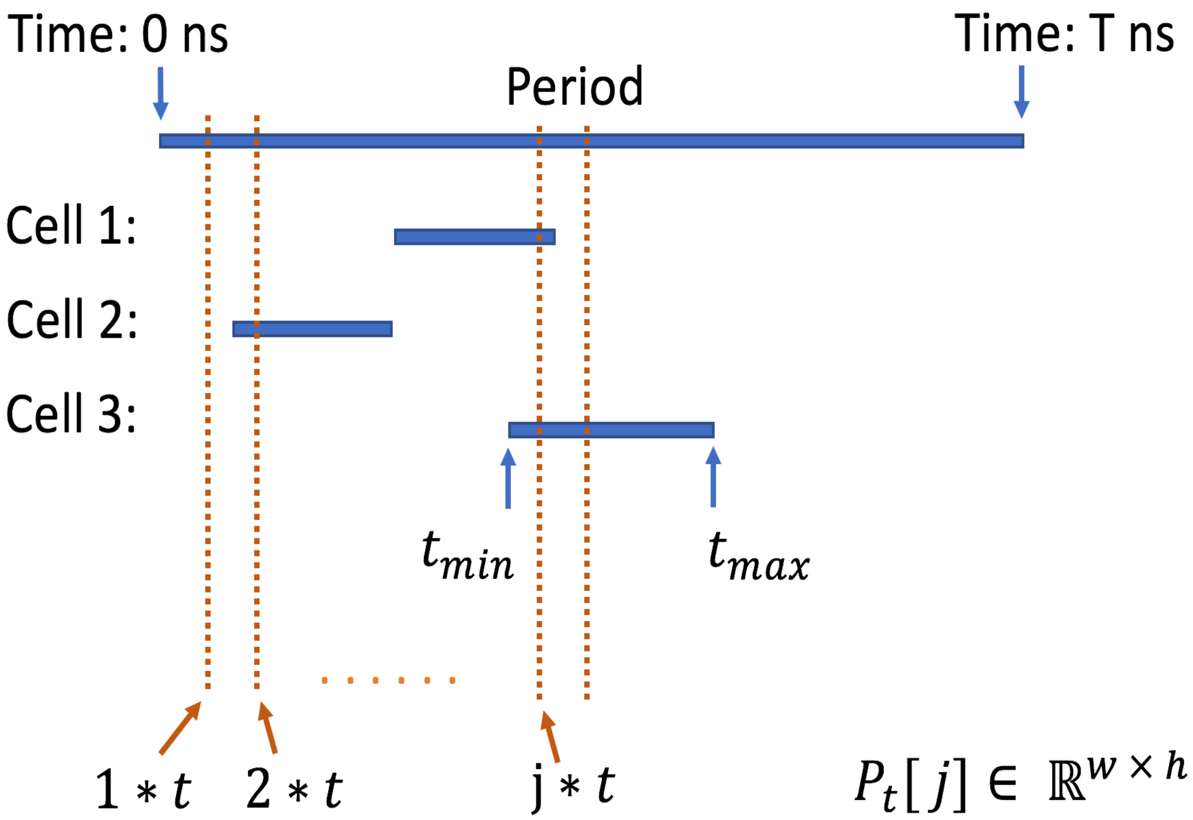


For each type of power, the power for analyzed grid is:

$$\mathbf{P}_{\mathbf{grid}} = P_1 + P_2 + P_3 + P_4 / 3 + P_5 / 2$$

Calculate power density

Method: Time Decomposition



Define N time frames within one cycle.

For each time frame j, generate one time-decomposed power map $P_t[j]$:

For each grid:

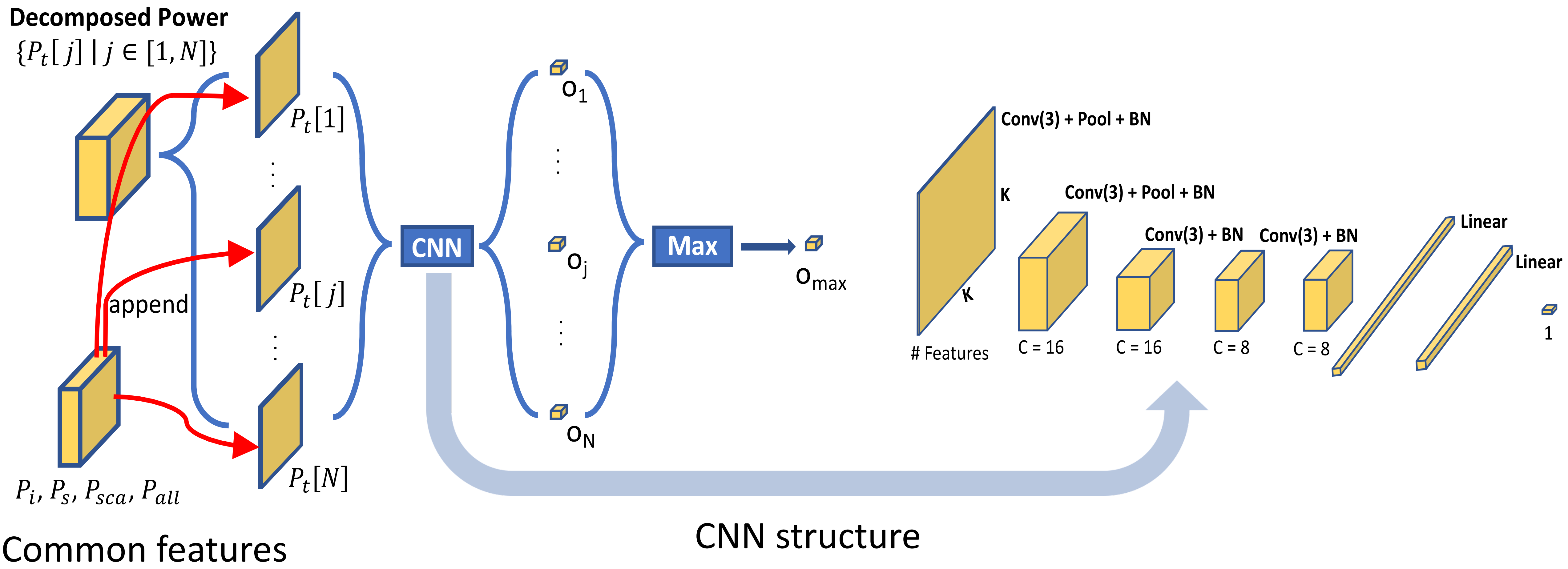
For each cell in it:

Only if $j*t \in [t_{min}, t_{max}]$:

Count the p_{sca} of this cell

To capture the worst transient local IR drop

Method: PowerNet Architecture



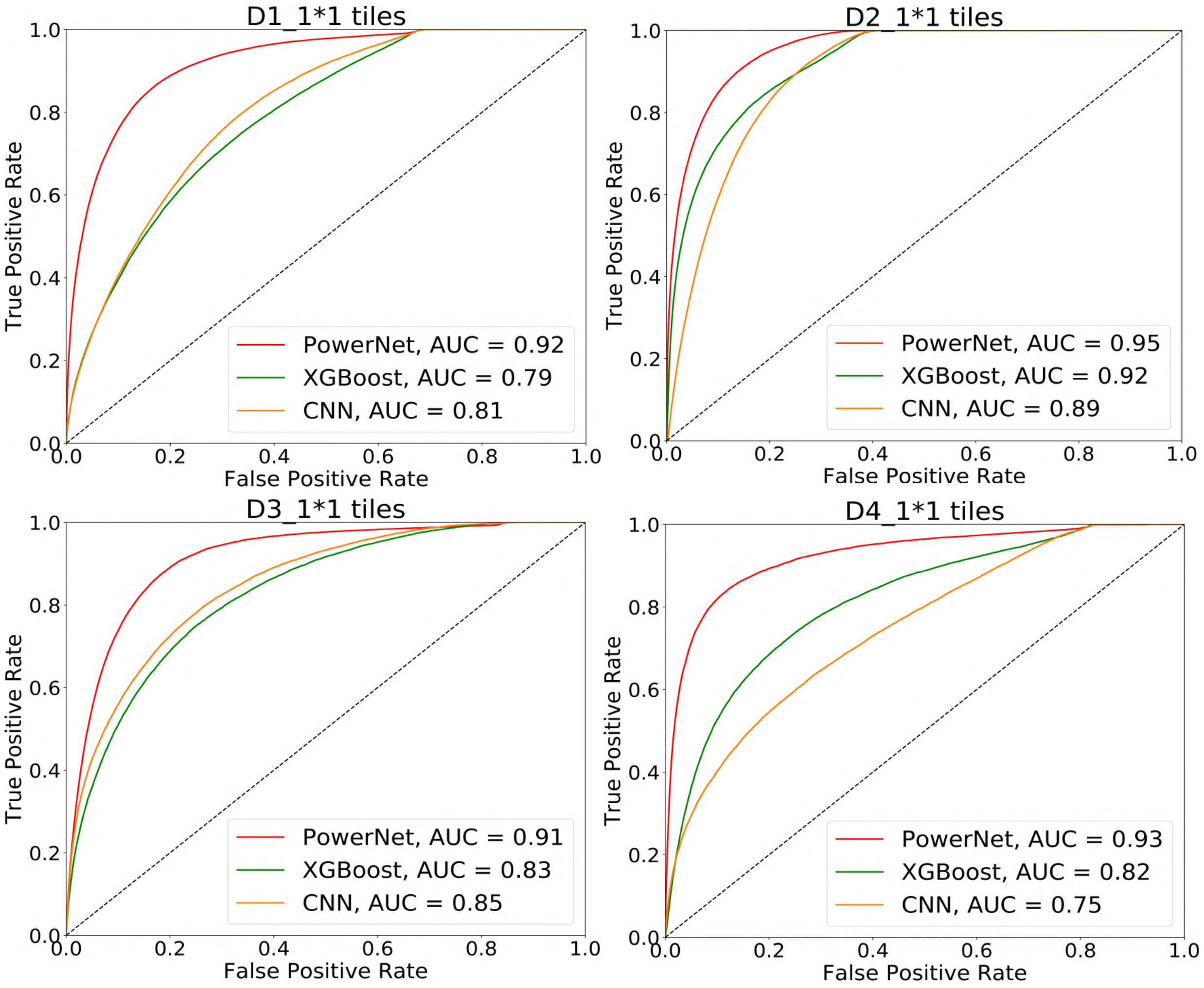
Idea: calculate IR drop for N instants and take the worst one.

Results: Experiments Setup

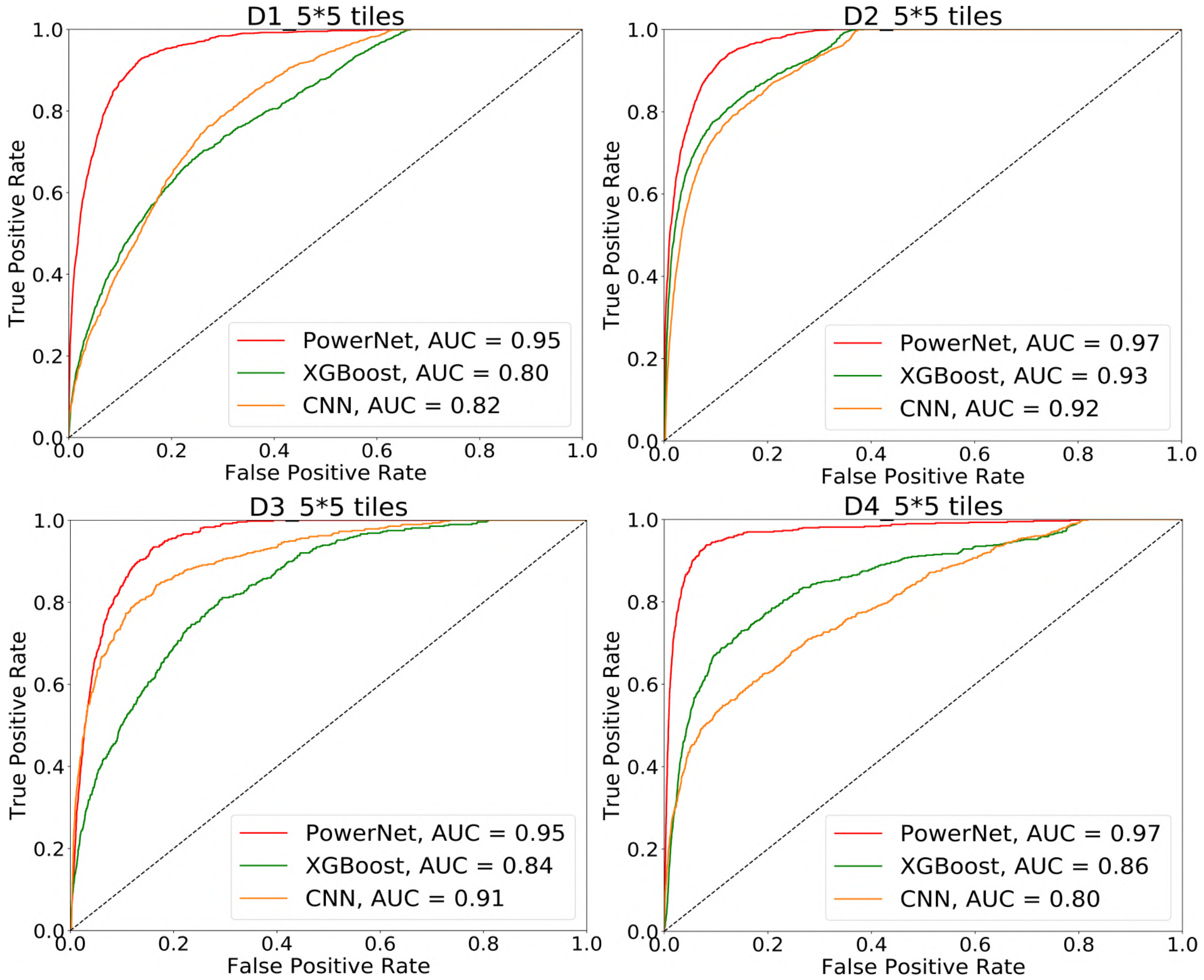
- D1-D4 measure prediction accuracy, MD1 & MD2 measure the performance on IR drop mitigation (hotspot portions much lower)
- All designs in sub-10nm technology
- The label is vectorless IR drop measured by commercial tool
- IR drop hotspot threshold is 56mV , 6% of the supply voltage (0.94V)
- Data extracted after CTS stage (can also be applied to other stages)

Design	D1	D2	D3	D4	MD1	MD2
# cells (million)	1.7	0.81	2.0	1.9	1.7	2.4
Hotspot Portion	5.6%	7.7%	3.1%	3.1%	0.65%	0.50%

Results: Accuracy Measured by ROC Curve



Measured in 1 grid length *1 grid length



Measured in 5 grid length * 5 grid length

Results: Accuracy Measured by MSE & Rank

- MSE (Mean squared error)
  - High MSE may be contributed by a consistent bias for all inferenced tiles

- Kendall rank coefficient

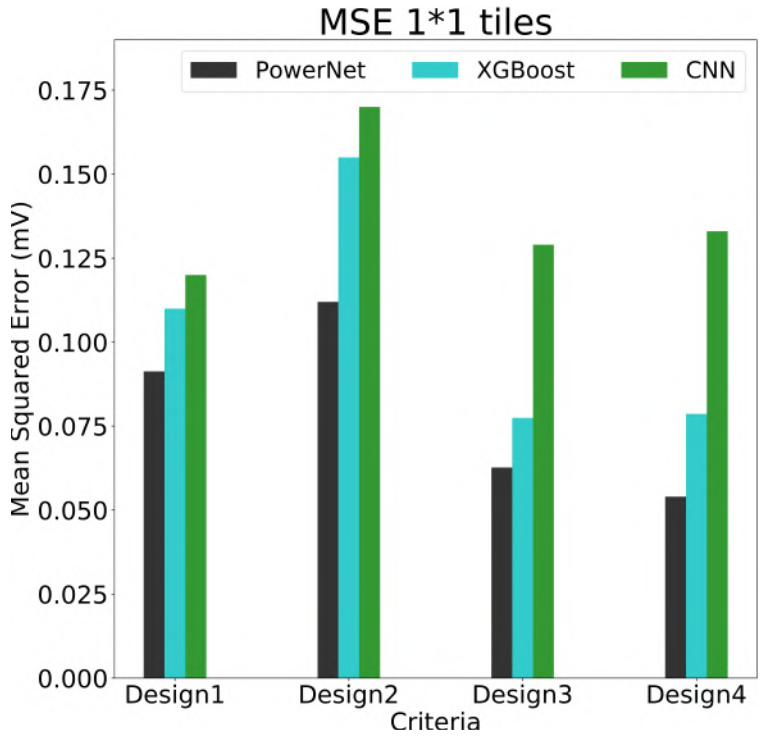
C is number of concordant pairs

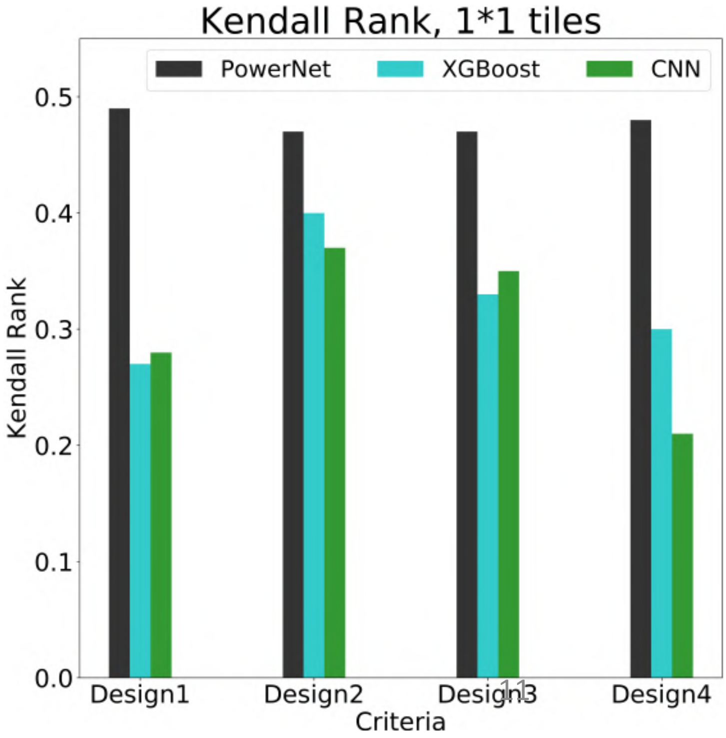
D is number of disconcordant pairs

n is overall number of elements

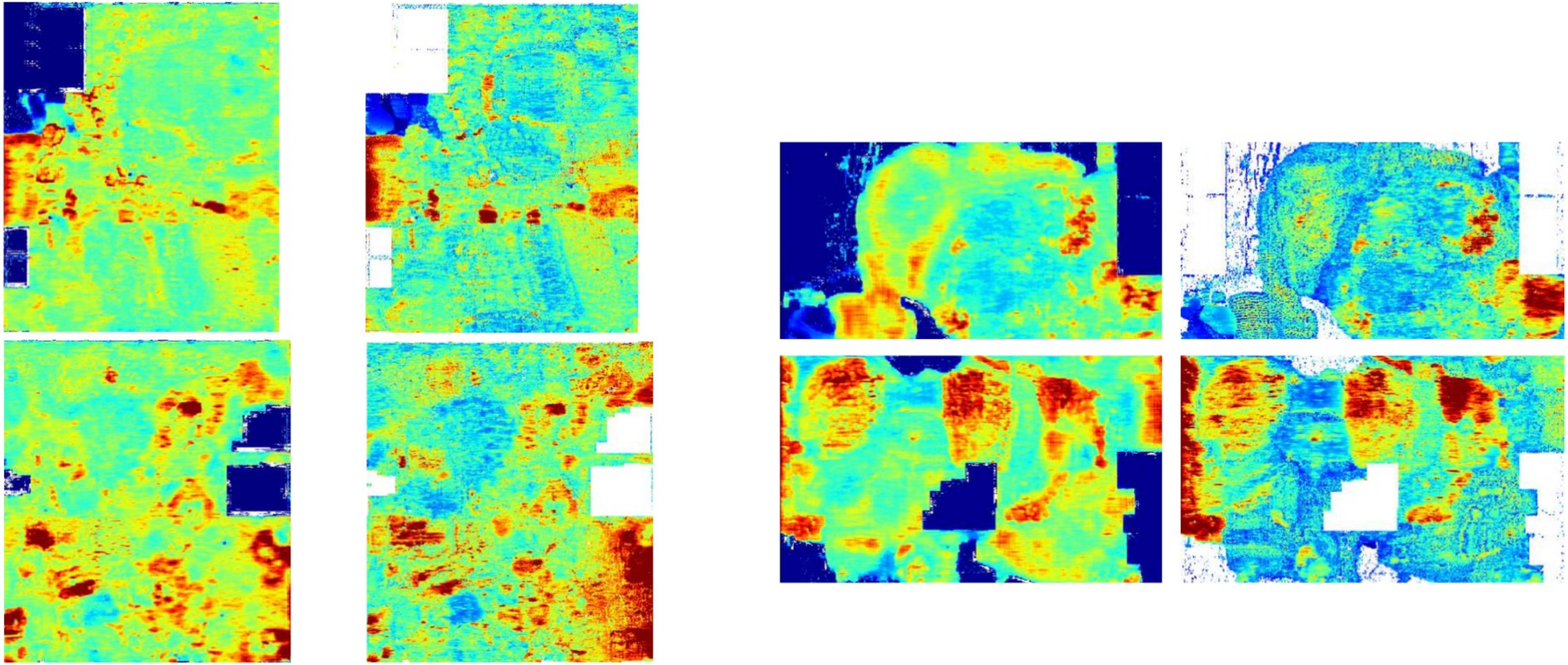
$$\text{Rank } \tau = \frac{C-D}{n(n-1)/2} \in [-1, 1]$$

How well predictions rank grids by IR drop compared with ground-truth ranking



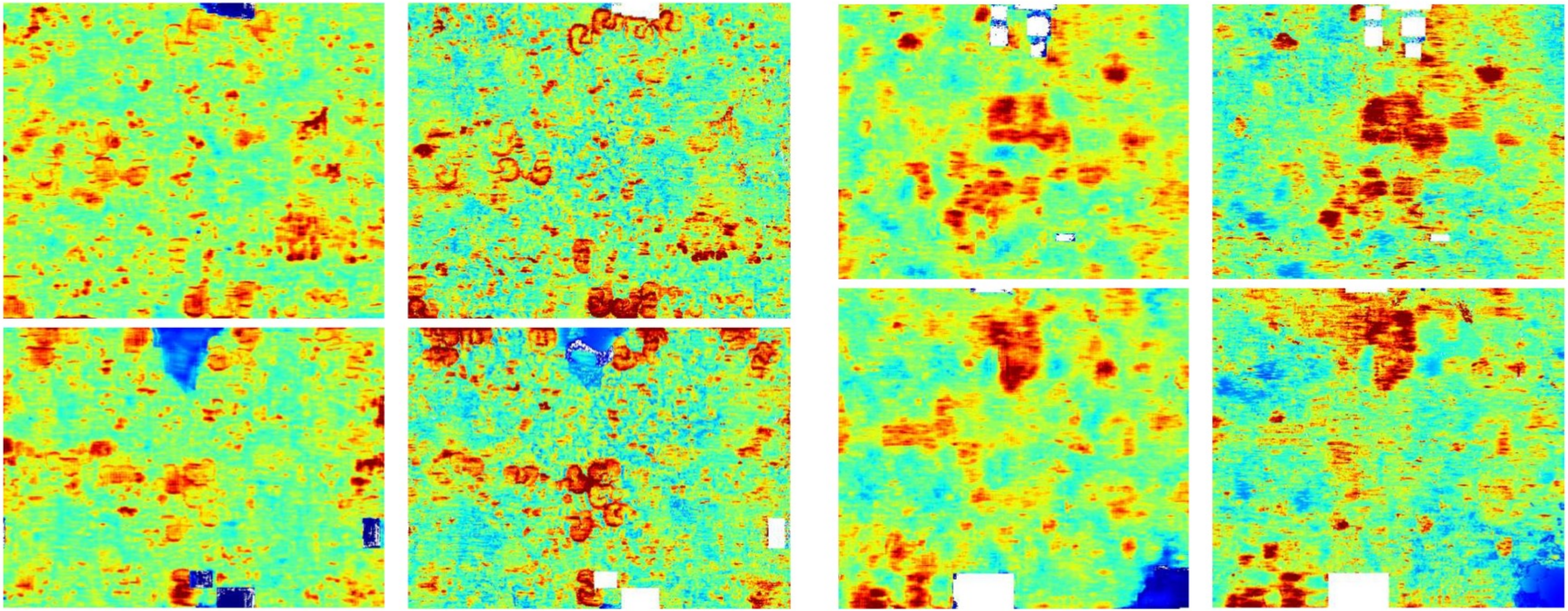


Results: Visualization of Predictions



D1 Label *D1* Prediction *D2* Label *D2* Prediction

Results: Visualization of Predictions



D3 Label *D3* Prediction *D4* Label *D4* Prediction

Results: IR Drop Mitigation with PowerNet

- Add very thin PG straps (0.04 µm) at the PowerNet-estimated hotspots
- Use this conservative method to prevent using much routing resources
- Averaged IR drop for all tiles improves only 0.4 mV
- Averaged IR drop for actual hotspots improves 4.3 mV and 2.6 mV

- PG enhancement is effective at the right places
- 26% and 31% reduction on #hotspots

Design MD1	Violated Cell	# Hotspots	All IR (mV)	Hotspot IR (mV)
Before Mitigate	22185	5092	26.4	66.6
After Mitigate	17052	3778	26.0	62.3
Improvement	23%	26%	0.4	4.3

Design MD2	Violated Cell	# Hotspots	All IR (mV)	Hotspot IR (mV)
Before Mitigate	31097	3627	31.4	62.2
After Mitigate	23941	2489	31.0	59.6
Improvement	23%	31%	0.4	2.6

Results: Inference Time

- Tested on a design with around two million cells
- The 2.5 hour for the commercial tool only includes analysis time, overall it takes 4 hours
- PowerNet achieves a 30× speedup over the commercial tool
- PowerNet is slower than the baseline ML methods because its CNN f generates N outputs o_j for each grid

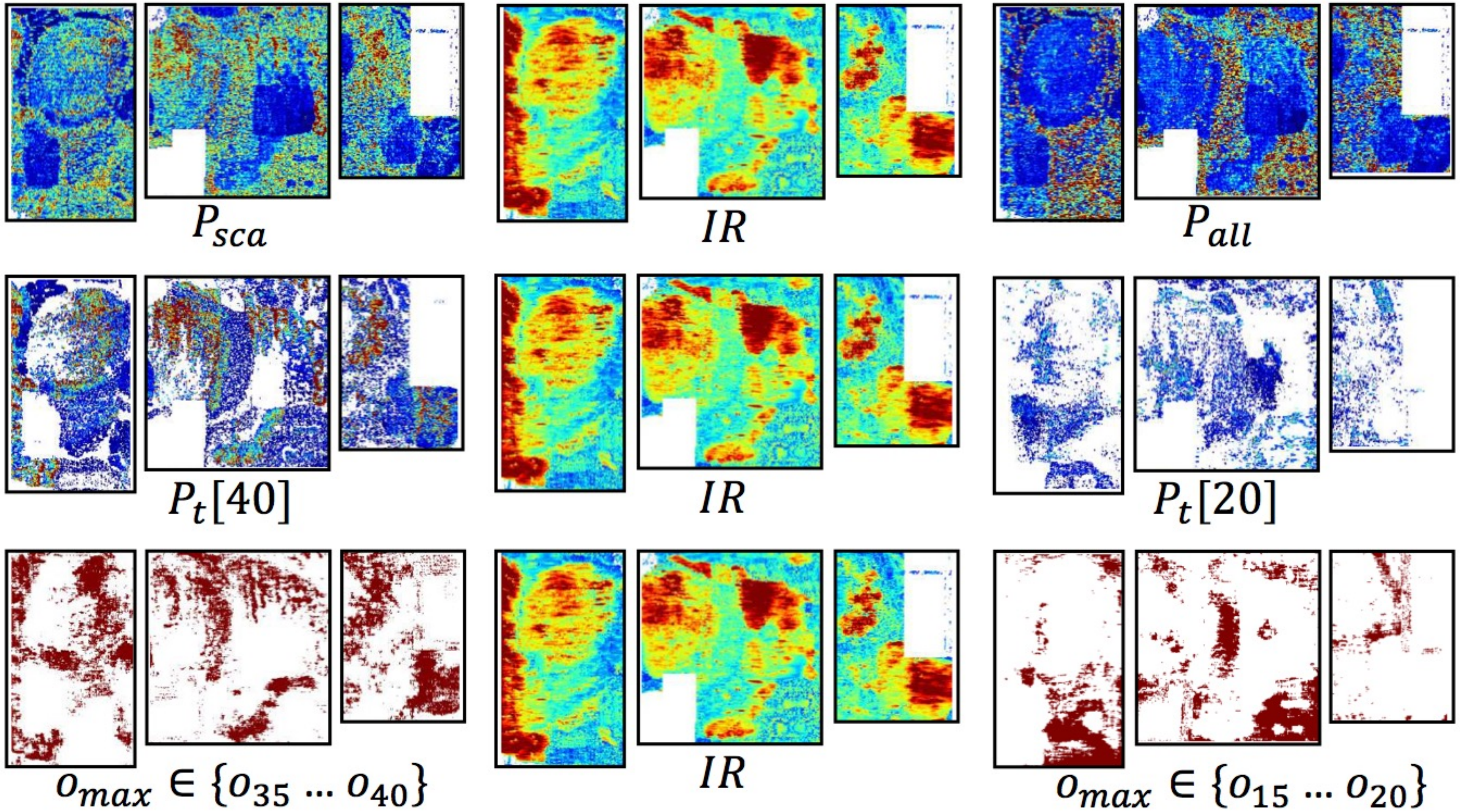
Method	Commercial Tool	PowerNet	CNN	XGBoost
Time	2.5 hour	5 min	1.5 min	1.5 min

Discussion: Time Decomposition Mechanism

- IR vs P_{sca}
- IR vs P_{all}

- IR vs $P_t[20]$
- IR vs $P_t[40]$

- IR vs $\sum_{j=15}^{20}[o_j]$
- IR vs $\sum_{j=35}^{40}[o_j]$



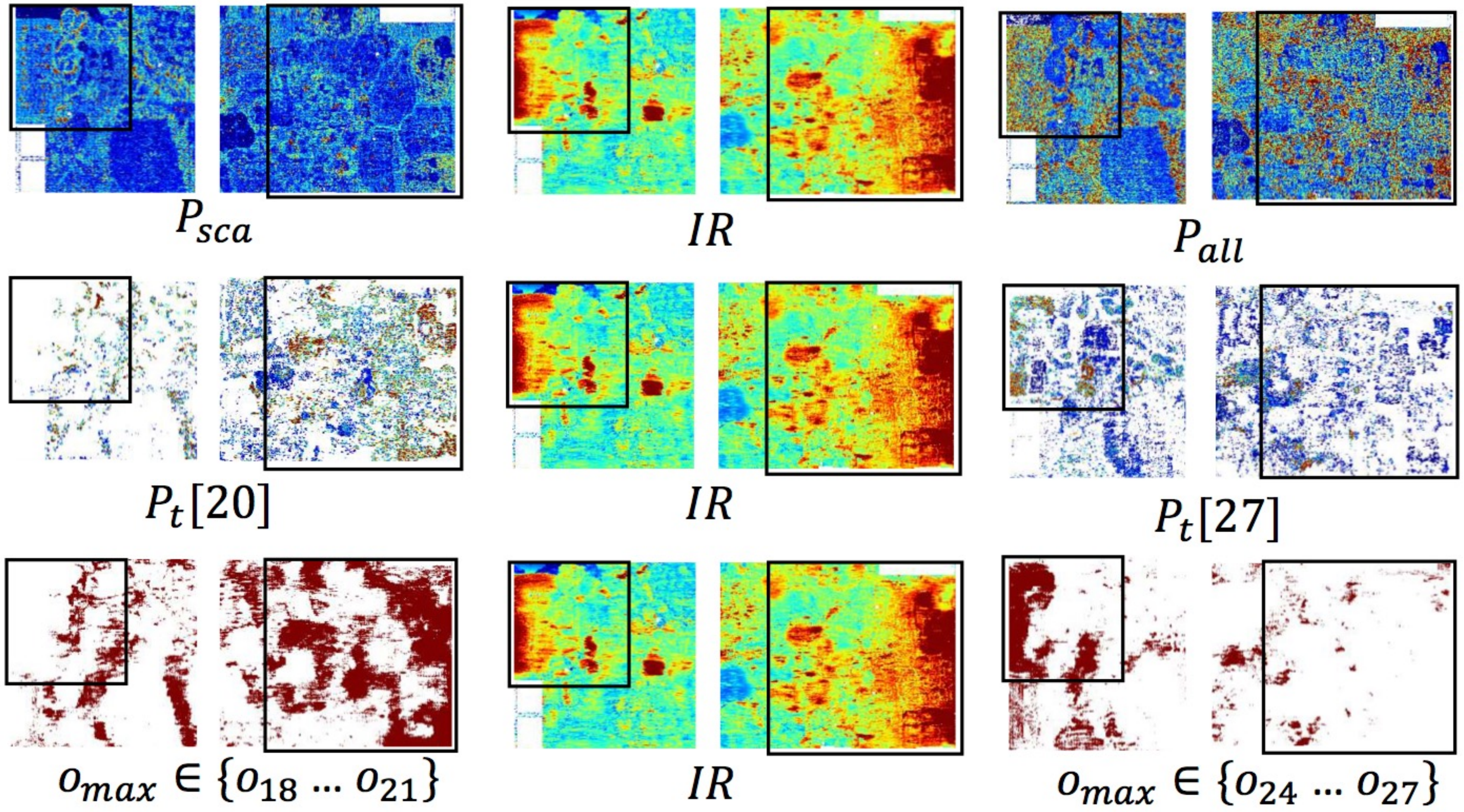
- Measure correlation between features & IR

Discussion: Time Decomposition Mechanism

- IR vs P_{sca}
- IR vs P_{all}

- IR vs $P_t[20]$
- IR vs $P_t[27]$

- IR vs $\sum_{j=18}^{21}[o_j]$
- IR vs $\sum_{j=24}^{27}[o_j]$



- Measure correlation between features & IR

Discussion: Influence of Resistance

- The standard deviation in resistance across the whole design is only 2.8Ω, 0.6% of average resistance
- Thus we choose not to spend extra time calculating per-cell resistance
- We provide another variation of PowerNet where each cell's power is scaled with resistance, named PRNet
- PRNet can be further applied to designs with non-uniform PDNs.

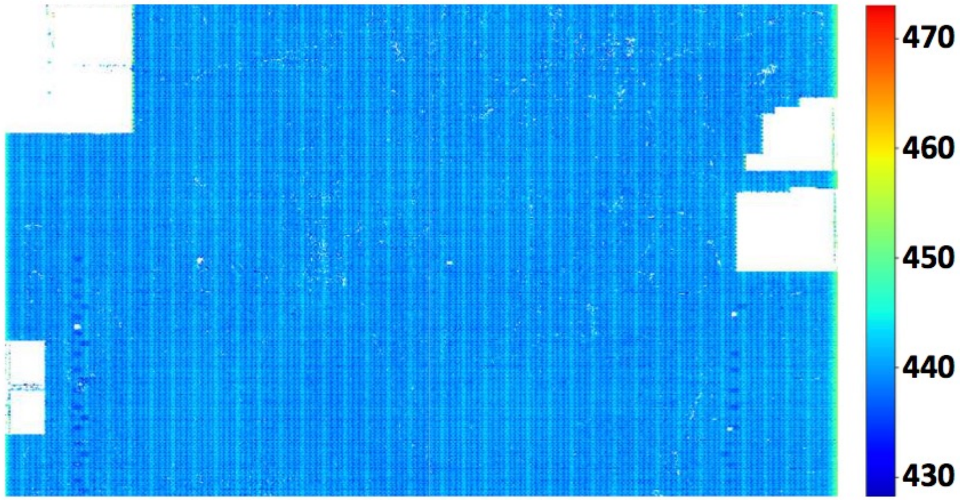


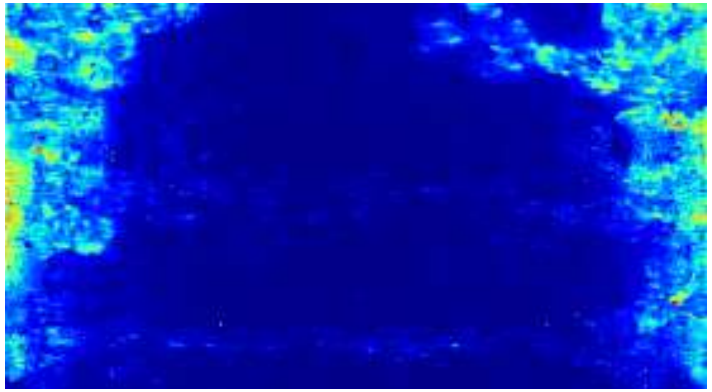
TABLE VI: Inference Accuracy in ROC AUC (0.01*)

ML Methods	D1	D2	D3	D4	Ave
PowerNet (1 × 1 tiles)	92.1	95.4	91.4	92.6	92.9
PRNet (1 × 1 tiles)	92.4	95.5	90.5	93.6	93.0
PowerNet (5 × 5 tiles)	95.4	96.7	94.8	97.0	96.0
PRNet (5 × 5 tiles)	95.7	96.8	93.2	97.5	95.8

Discussion: Vector-based IR Drop Estimation

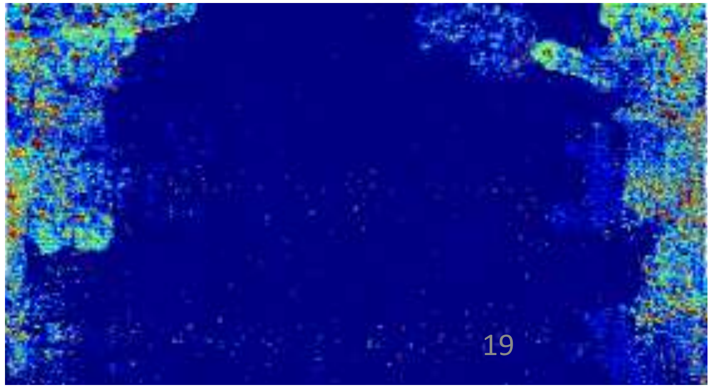
- The correlation between power and IR drop value turns out to be very strong
- We also perform vector-based estimation on four other industrial designs VD1 to VD4
- All methods provide better estimation than the vectorless scenario
- PowerNet still gives the best accuracy

IR



ML Methods	1×1 tiles				5×5 tiles			
	VD1	VD2	VD3	VD4	VD1	VD2	VD3	VD4
XGBoost	97	98	98	96	99	97	98	97
CNN	96	93	95	95	98	92	97	96
PowerNet	98	98	99	97	100	98	100	98

P_{all}



Conclusion

- We propose a CNN-based dynamic IR drop estimator named PowerNet
- The model is general and transferable to new designs
- PowerNet takes an order of magnitude less estimation time than commercial tools
- PowerNet outperforms other ML methods for both vector-based and vectorless IR drop in accuracy
- IR drop mitigation tool guided by PowerNet reduces IR drop by >20% with very limited PG modification

Thanks!