PHYSICS-INFORMED OPTICAL KERNEL REGRESSION USING COMPLEX-VALUED NEURAL FIELDS

Guojin Chen[1], Zehua Pei[1], Haoyu Yang[2], Yuzhe Ma[3], Bei Yu[1], Martin Wong[1]

[1]The Chinese University of Hong Kong, [2]nVIDIA, [3]HKUST(GZ)

Introduction

- Lithography is fundamental to integrated circuit fabrication, necessitating large computation overhead.
- All previous methods regard the lithography system as an image-to-image black box mapping.
- In this paper, we propose a new ML-based paradigm disassembling the rigorous lithographic model into non-parametric mask operations and learned optical kernels containing determinant source, pupil, and lithography information.

Keywords: PINN, NeRF, Lithography

Background of Lithography

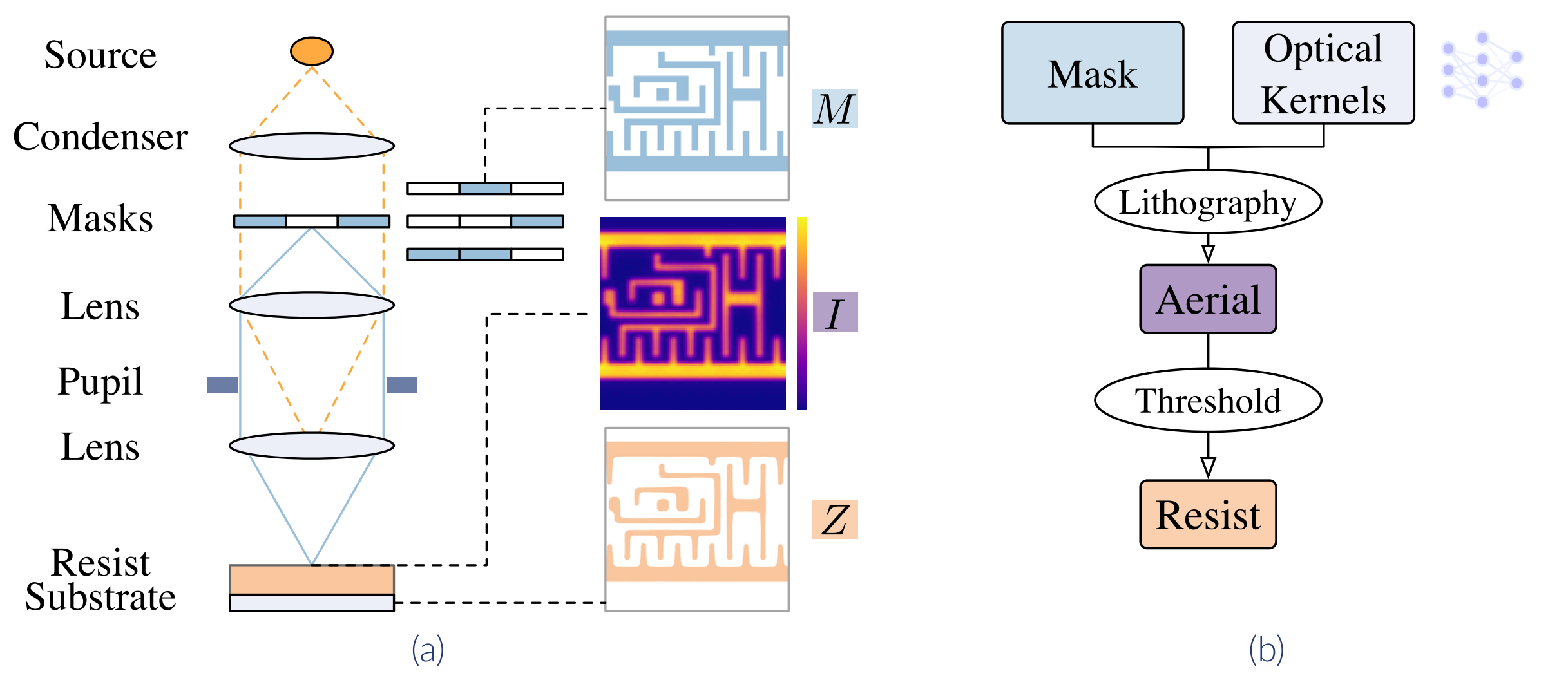


Figure 1. (a) Components of the lithography imaging system: illumination source, lenses, and pupil. (b) Lithography simulation flow using source- and pupil-dependent optical

Summary of previous works

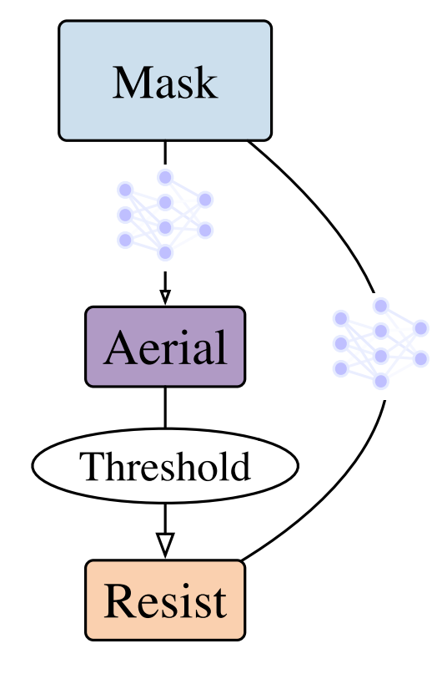


Figure 2. General flow of previous SOTA.

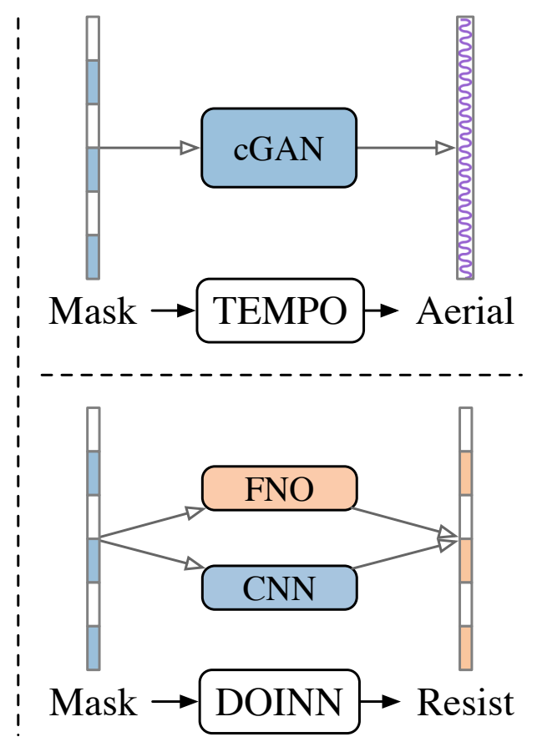


Figure 3. Previous SOTA work on Aerial stage (TEMPO) and Resist stage (DOINN).

- Previous works, two stage: mask-to-resist, mask-to-aerial
- Modeling the lithography process as a black box, utilizing neural networks to fit this black box.

Drawbacks of previous works

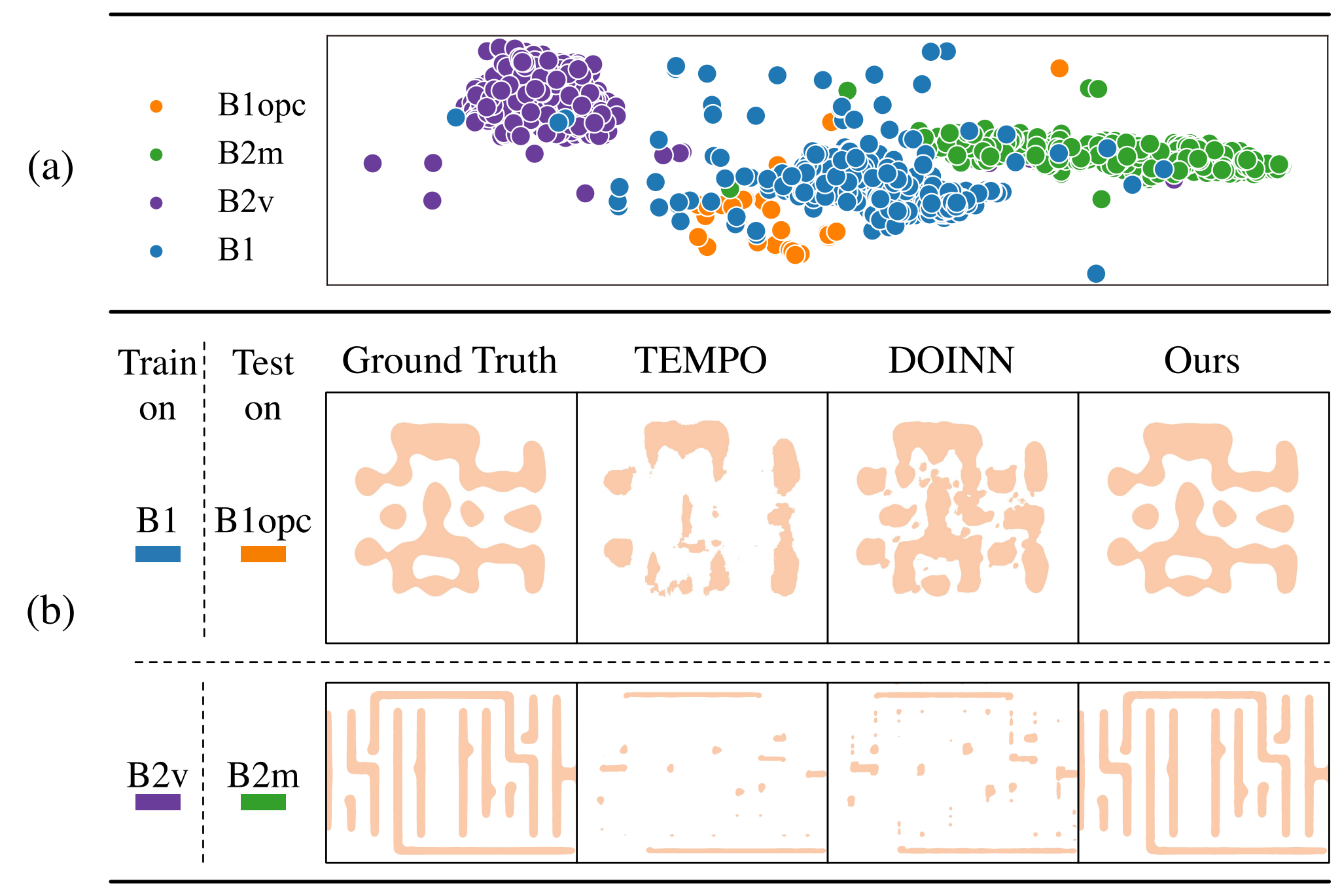


Figure 4. (a) t-SNE distribution of datasets. (b) Comparison of generalization capability on out-of-distribution (OOD) datasets.

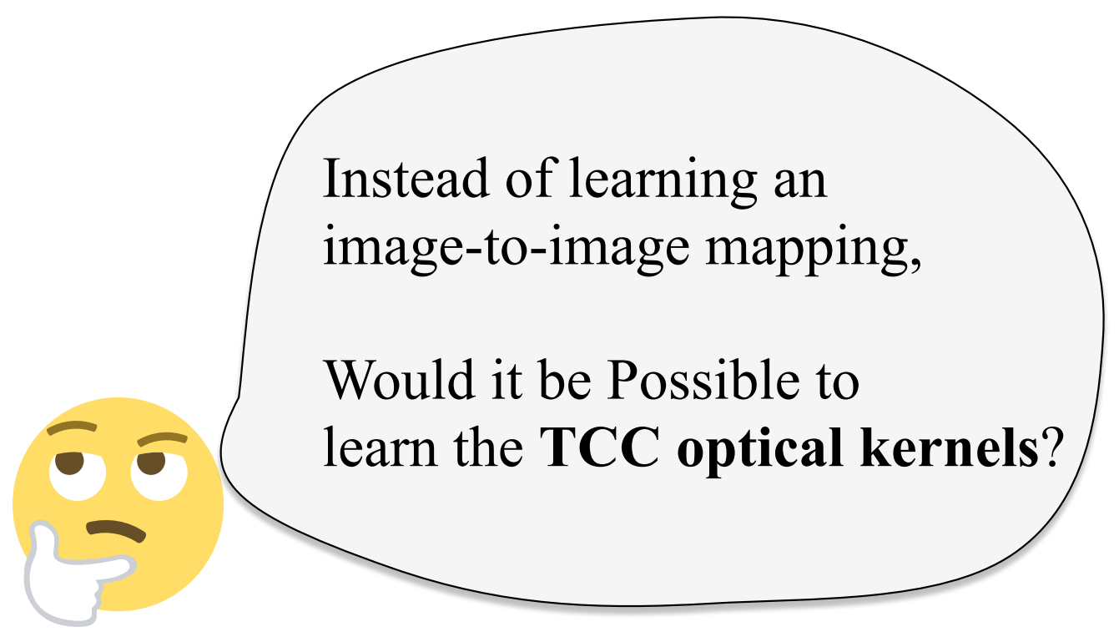
Previous image-learning based

- ✘ ☹ Bias on image distribution.
- ✘ ☹ Large models.
- ✓ 😄 Fast prediction.

Industrial lithography simulator.

- ✓ 😄 Good generalization capability.
- ✘ ☹ Computationally expensive.

Motivation



Homepage

LinkedIn

Overall flow of Nitho

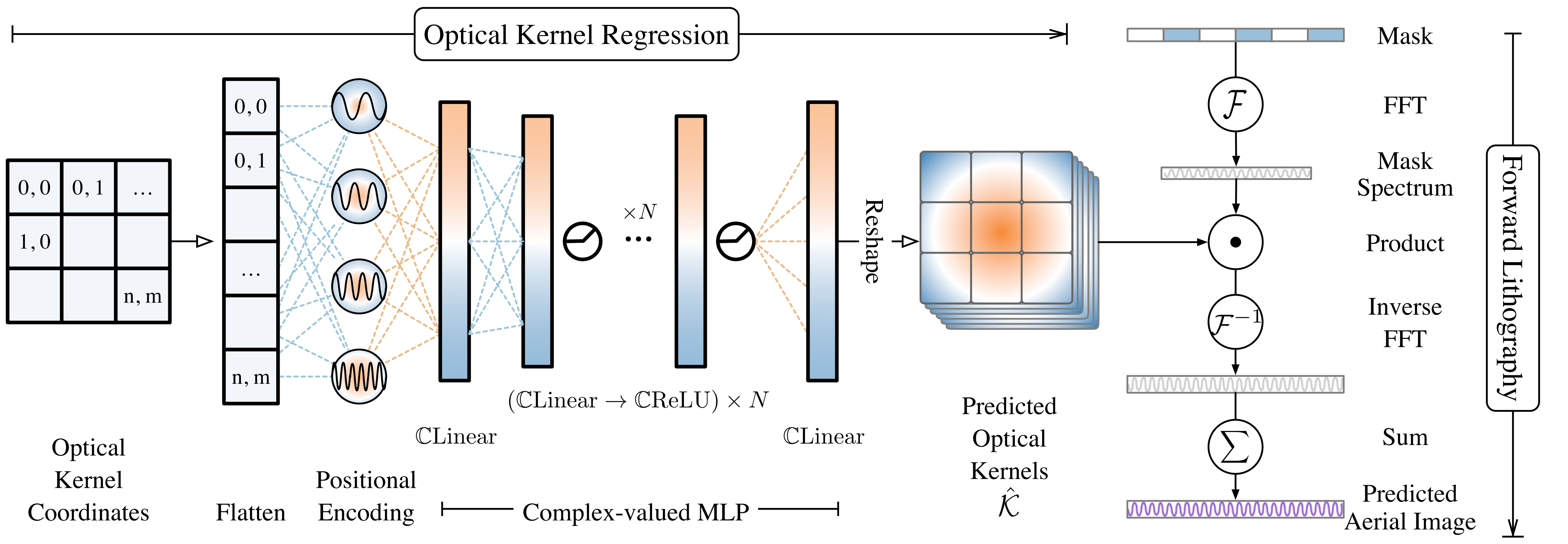


Figure 5. The overall aerial image prediction pipeline of Nitho framework, which separates mask-related linear operations from optical kernel regression using coordinate-based ℂMLP.

Hopkins Model & Transmission Cross-Coefficient (TCC)

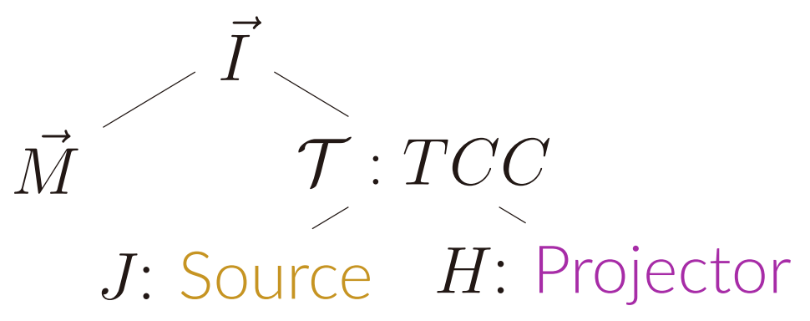
The imaging equation:

$$\mathcal{F}(\vec{I})(f,g)=\iint_{-\infty}^{\infty}\underline{\mathcal{T}\left(\left(f'+f,g'+g\right),\left(f',g'\right)\right)}\,\mathcal{F}(\vec{M})\left(f'+f,g'+g\right)\mathcal{F}(\vec{M})^{*}\left(f',g'\right)\mathrm{d}f'\mathrm{d}g', \tag{1}$$

where $\vec{M}$ is the mask, (f,g) is its frequencies. $\mathcal{T}$ is TCC given by:

$$\mathcal{T}\left(\left(f',g'\right),\left(f'',g''\right)\right):=\iint_{-\infty}^{\infty}\underline{\mathcal{F}(J)(f,g)}\,\mathcal{F}(H)\left(f+f',g+g'\right)\mathcal{F}(H)^{*}\left(f+f'',g+g''\right)\mathrm{d}f\,\mathrm{d}g, \tag{2}$$

where the weight factor J solely depends on effective source, H is projector transfer function.



When the projector and source are fixed,

$\boxed{J\text{: constant matrix}}$ $\boxed{H\text{: constant matrix}}$

$$\Downarrow$$

$\boxed{\mathcal{T}\text{: TCC is a constant matrix}}$

The benefits of learning optical kernels

- Get rid of negative influence of layer types & dataset distribution.
- Less training data required & smaller model size.

The solutions of learning optical kernels

1. Design the kernel dimension based on physical “*resolution limit*”.
2. Implement a set of differentiable complex-valued neuron layers.
3. A new training paradigm separates the influence of masks and optical kernels

SOCS w. optical kernels / Kernel Dimensions

$$\vec{I}=\sum_{i=1}^{r}\alpha_i\left|\mathcal{F}^{-1}\left(\mathcal{F}\left(\boldsymbol{h}_i\right)\odot\mathcal{F}(\boldsymbol{M})\right)\right|^2.$$

$$\Downarrow$$

$$\vec{I}=\sum_{i}^{r}\left|\mathcal{F}^{-1}\left(\mathcal{K}_i\odot\mathcal{F}(\boldsymbol{M})\right)\right|^2,$$

$$m=(W\times\frac{2\mathrm{NA}}{\lambda})\times 2+1,$$

$$n=(H\times\frac{2\mathrm{NA}}{\lambda})\times 2+1,$$

$$\Rightarrow\mathcal{K}\in\mathbb{C}^{r\times n\times m}$$

Computing using complex-valued neural network

Complex Liner Layer: $\vec{W}=\vec{A}+i\vec{B}$ by a complex vector $\vec{h}=\vec{x}+i\vec{y}$, where $\vec{A}$ and $\vec{B}$ are real matrices and $\vec{x}$ and $\vec{y}$ are real vectors. We obtain:

$$\vec{W}\vec{h}=(\vec{A}\vec{x}-\vec{B}\vec{y})+i(\vec{B}\vec{x}+\vec{A}\vec{y}),$$

Complex ReLU: Complex rectified linear unit ($\mathbb{C}\mathrm{ReLU}$) is applied as:

$$\mathbb{C}\mathrm{ReLU}(z)=\mathrm{ReLU}(\Re(z))+i\,\mathrm{ReLU}(\Im(z)).$$

$\mathbb{C}\mathrm{MLP}$: The $\mathbb{C}\mathrm{MLP}$ is further constructed as,

$$\mathbb{C}\mathrm{MLP}:\mathbb{C}\mathrm{Linear}\rightarrow(\mathbb{C}\mathrm{Linear}\rightarrow\mathbb{C}\mathrm{ReLU})\times N\ldots\rightarrow\mathbb{C}\mathrm{Linear}, \tag{3}$$

where $\times N$ means there are N hidden blocks ($\mathbb{C}\mathrm{Linear}\rightarrow\mathbb{C}\mathrm{ReLU}$).

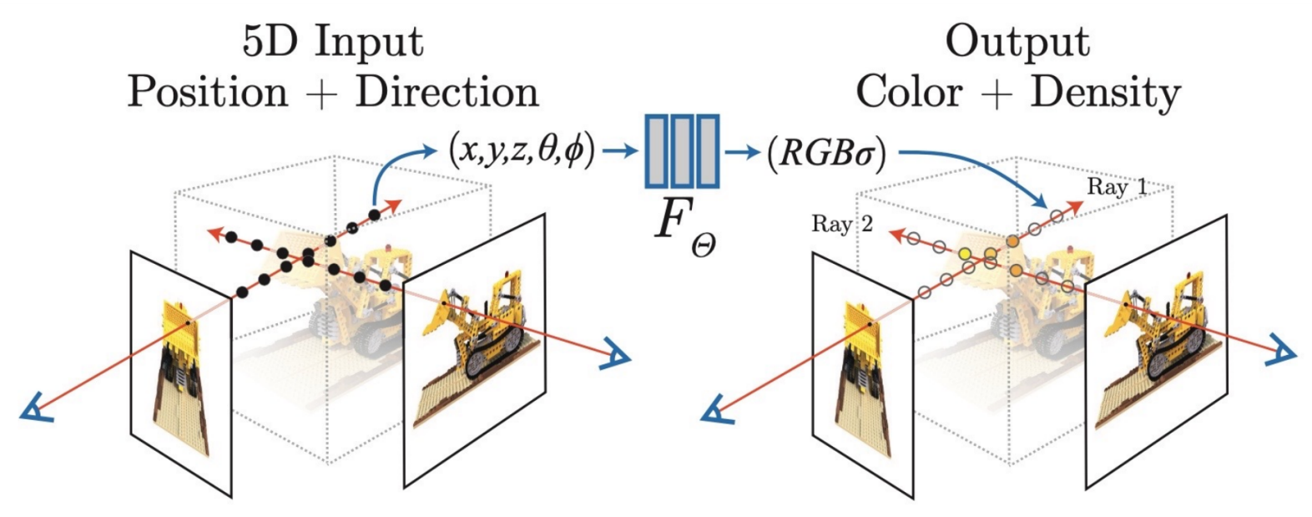
NeRF

Inputs:

- x, y, z: Target position.
- θ, ϕ: orientation.

Outputs:

- $c=(r,g,b)$: Color.
- σ: Volume density.



Nitho: NeRF inspired lithography simulator.

The lithography conditions are location dependent. TCC is given by:

$$\mathcal{T}\left(\left(f',g'\right),\left(f'',g''\right)\right):=\iint_{-\infty}^{\infty}\underline{\mathcal{F}(J)(f,g)}\,\mathcal{F}(H)\left(f+f',g+g'\right)\mathcal{F}(H)^{*}\left(f+f'',g+g''\right)\mathrm{d}f\,\mathrm{d}g, \tag{4}$$

One more thing: positional encoding

NeRF’s positional encoding:

$$\gamma(\vec{v})=\left[\sin(2^0\pi\vec{v}),\cos(2^0\pi\vec{v}),\ldots,\sin\left(2^{L-1}\pi\vec{v}\right),\cos\left(2^{L-1}\pi\vec{v}\right)\right]^{\mathrm{T}},$$

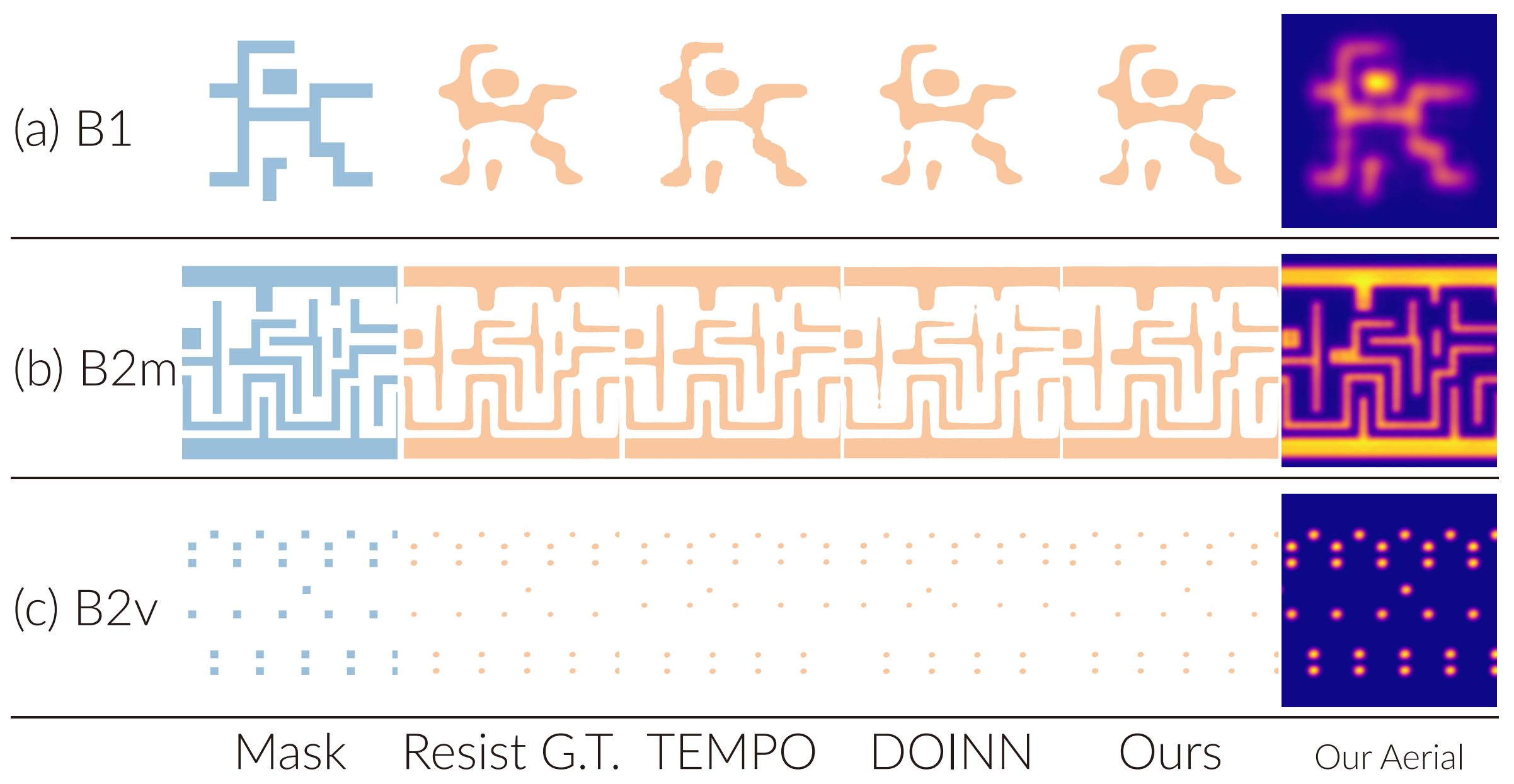
Ours:

$$\gamma(\vec{v})=\left[\cos(2\pi\vec{B}v)*(1+j),\ \sin(2\pi\vec{B}v)*(1+j)\right]^{\mathrm{T}},$$

Table 1. Result Comparison with State-of-the-Art.

Bench	Aerial Image									Resist Image					
	TEMPO			DOINN			Nitho			TEMPO*		DOINN*		Nitho	
	MSE $\times10^{-5}$	ME $\times10^{-2}$	PSNR dB	MSE $\times10^{-5}$	ME $\times10^{-2}$	PSNR dB	MSE $\times10^{-5}$	ME $\times10^{-2}$	PSNR dB	mPA (%)	mIOU (%)	mPA (%)	mIOU (%)	mPA (%)	mIOU (%)
B1	108.29	10.49	32.01	5.55	1.94	47.10	**1.32**	**0.51**	**50.75**	94.60	88.70	99.19	98.32	**99.45**	**99.21**
B2m	1899.04	13.96	30.77	1202.39	6.11	31.64	**25.48**	**0.82**	**49.06**	98.24	96.55	98.79	97.10	**99.15**	**99.02**
B2v	6.54	3.86	42.76	2.26	2.75	46.37	**2.01**	**0.68**	**48.06**	99.06	93.28	99.21	98.41	**99.59**	**99.34**
B2m + B2v	4352.25	15.21	27.10	3114.24	12.35	29.92	**33.13**	**0.78**	**47.88**	98.63	95.84	98.71	96.68	**99.61**	**99.36**
Average	1591.53	10.88	33.16	1081.11	5.79	39.26	**15.49**	**0.70**	**48.94**	97.63	93.59	98.98	97.63	**99.45**	**99.23**
Ratio	102.77	15.55	0.68	69.81	8.27	0.80	**1.00**	**1.00**	**1.00**	0.98	0.94	0.99	0.98	**1.00**	**1.00**

Figure 6. Visualization of the results of Nitho in aerial and resist stage.



Ablation study on smaller training sets and kernels sizes

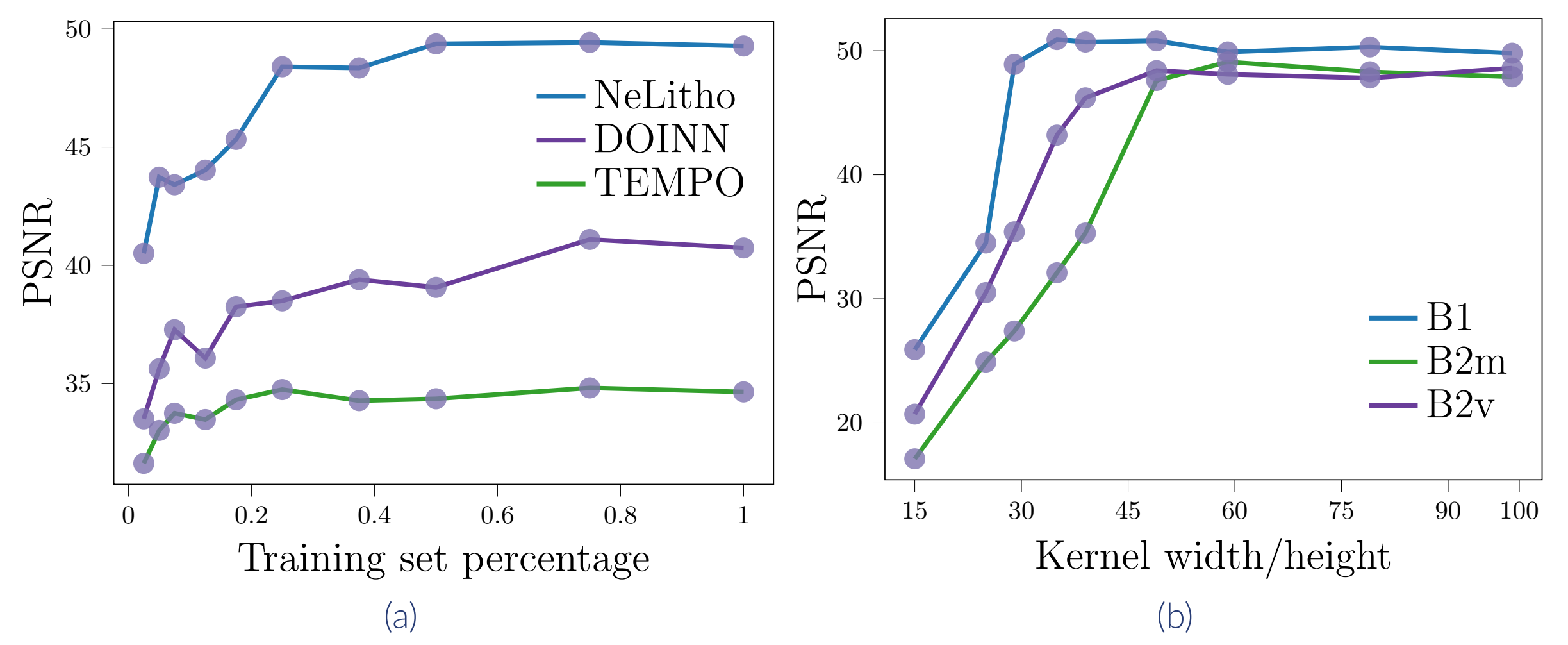


Figure 7. (a) Comparison with SOTA on smaller training sets. (b) Ablation study on kernel size on different datasets.

https://gjchen.me

DAC 2023

cgjcuhk@gmail.com