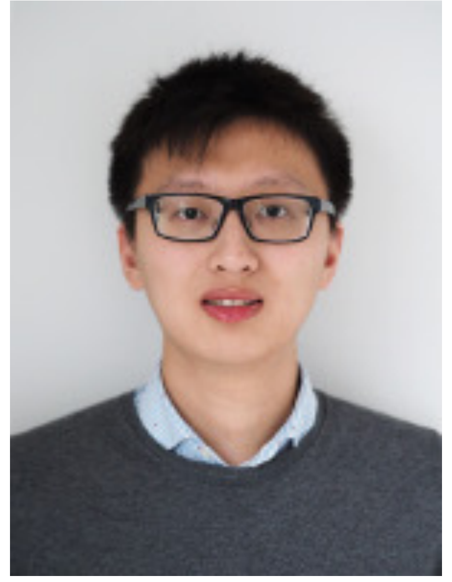
THE UNIVERSITY OF BRITISH COLUMBIA

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Neural Fourier Filter Bank

Zhijie Wu

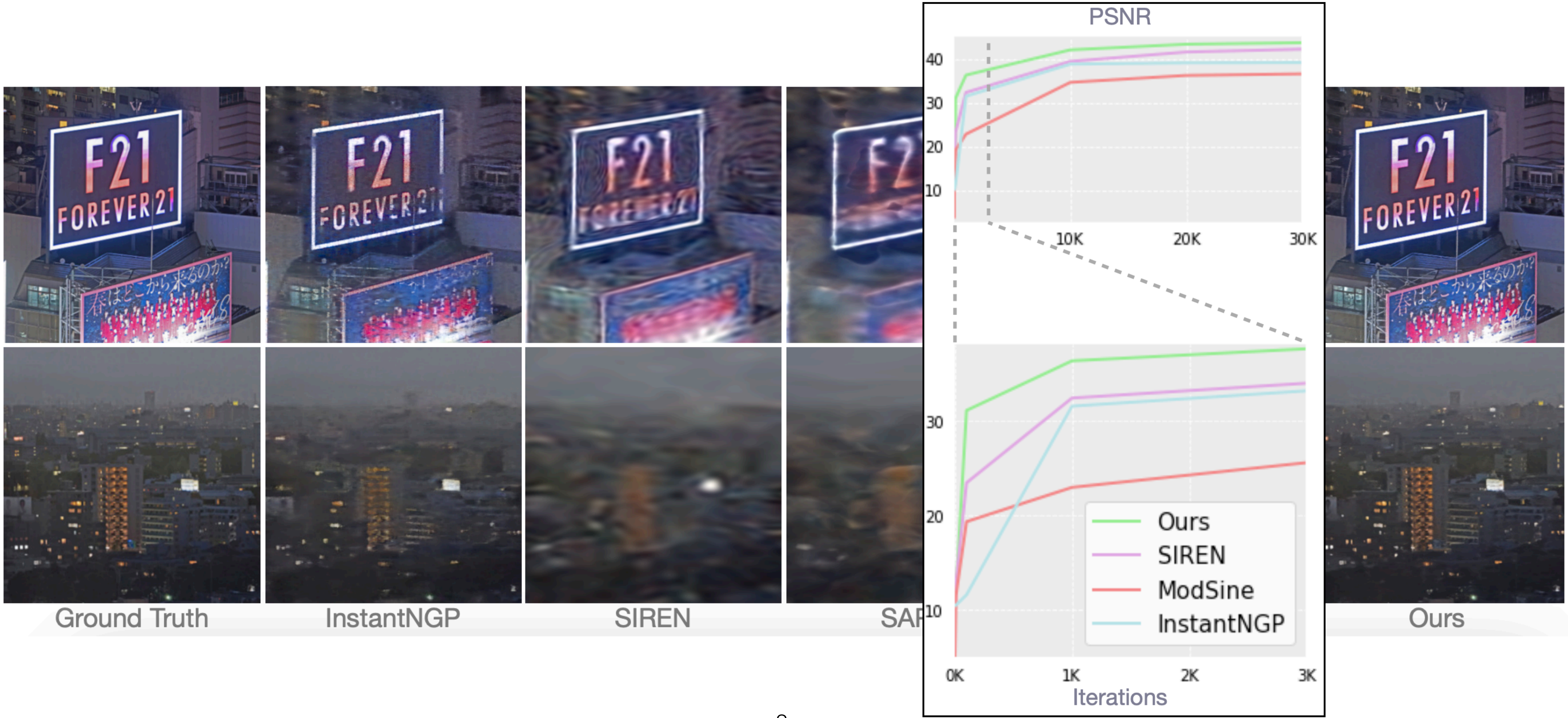


Yuhe Jin

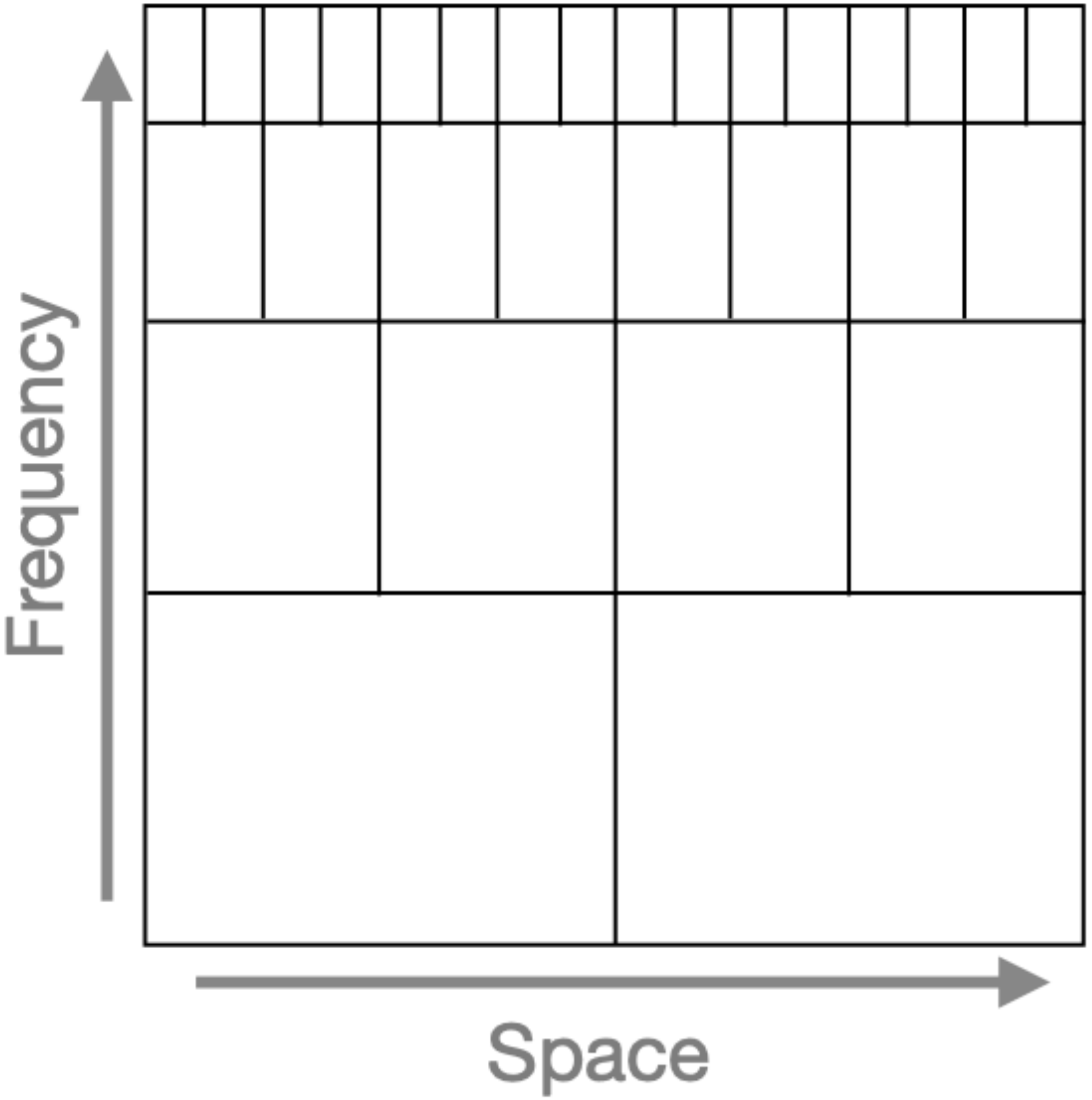
Kwang Moo Yi

WED-PM-172

Reconstruction quality↑ + compute↓



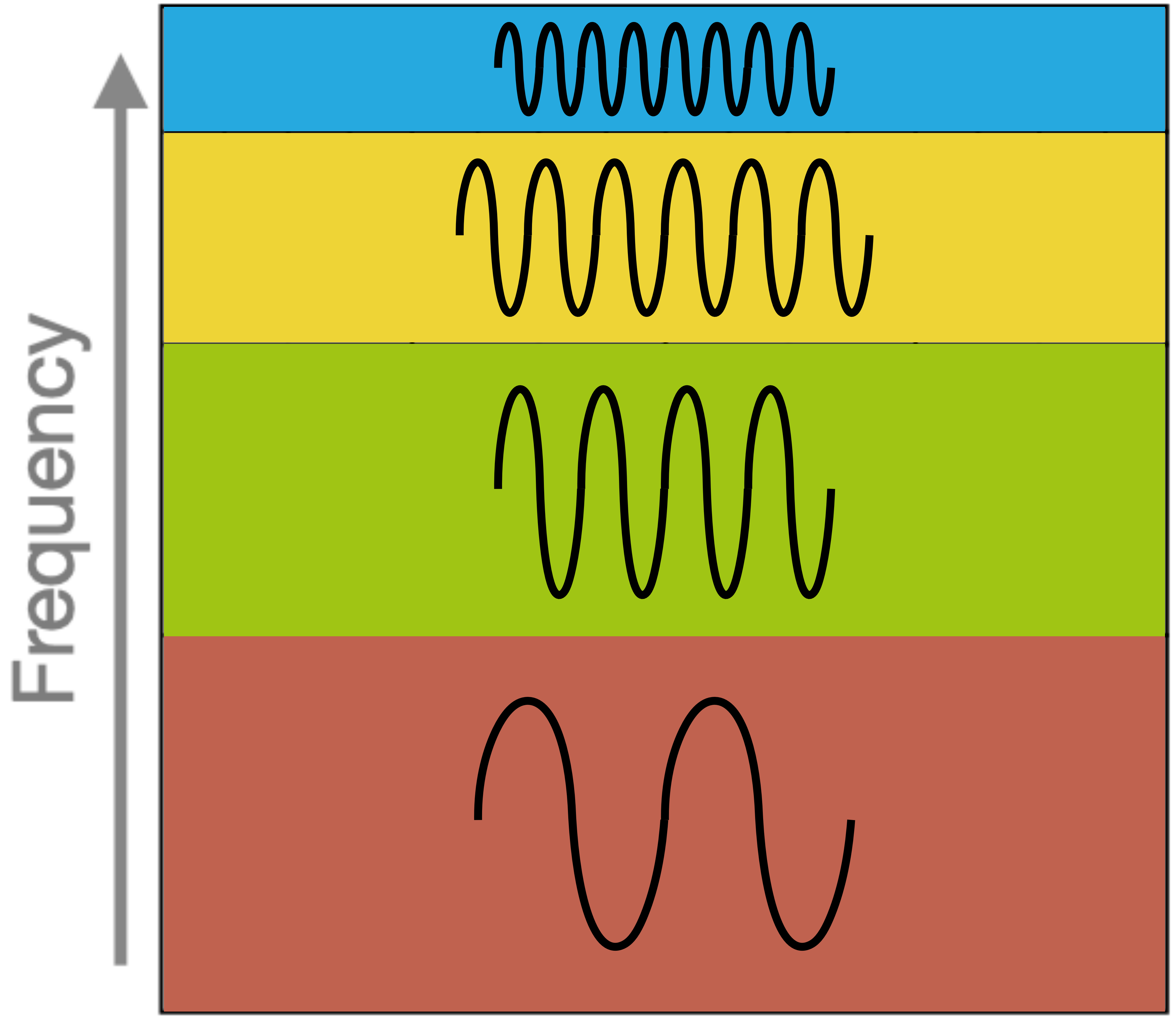
Considering both space and frequency



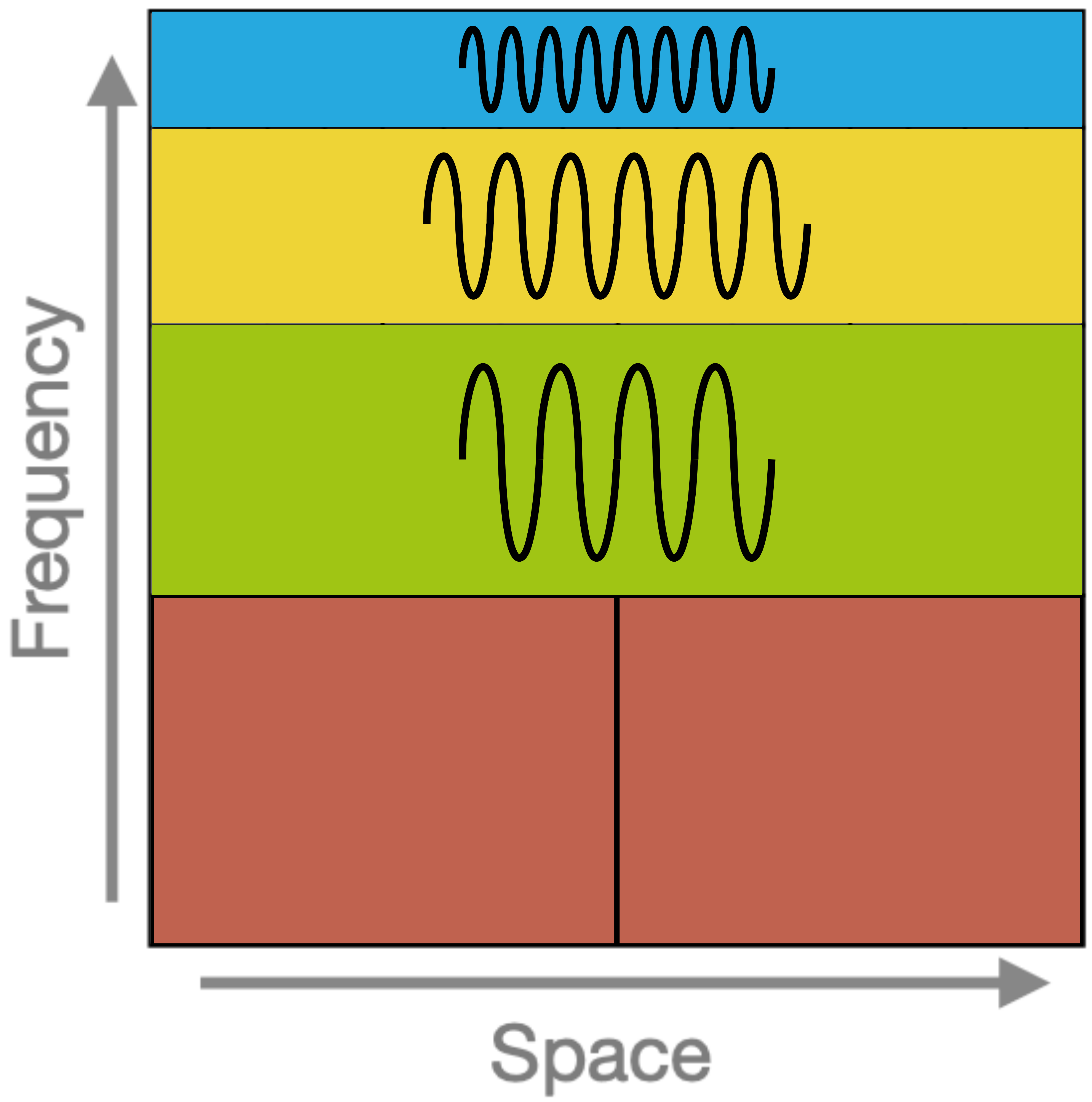
THE UNIVERSITY OF BRITISH COLUMBIA

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

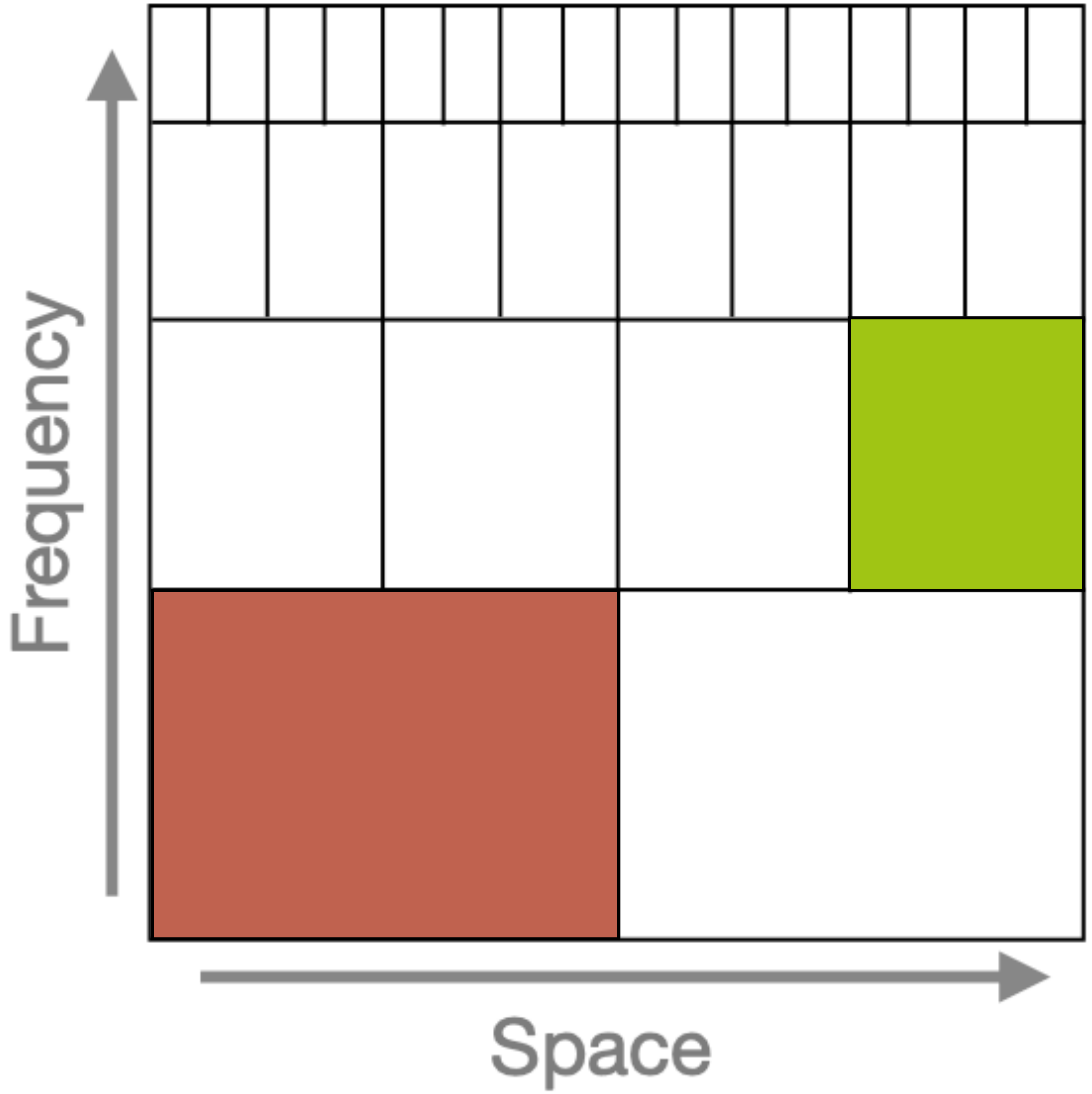
Considering both space and frequency



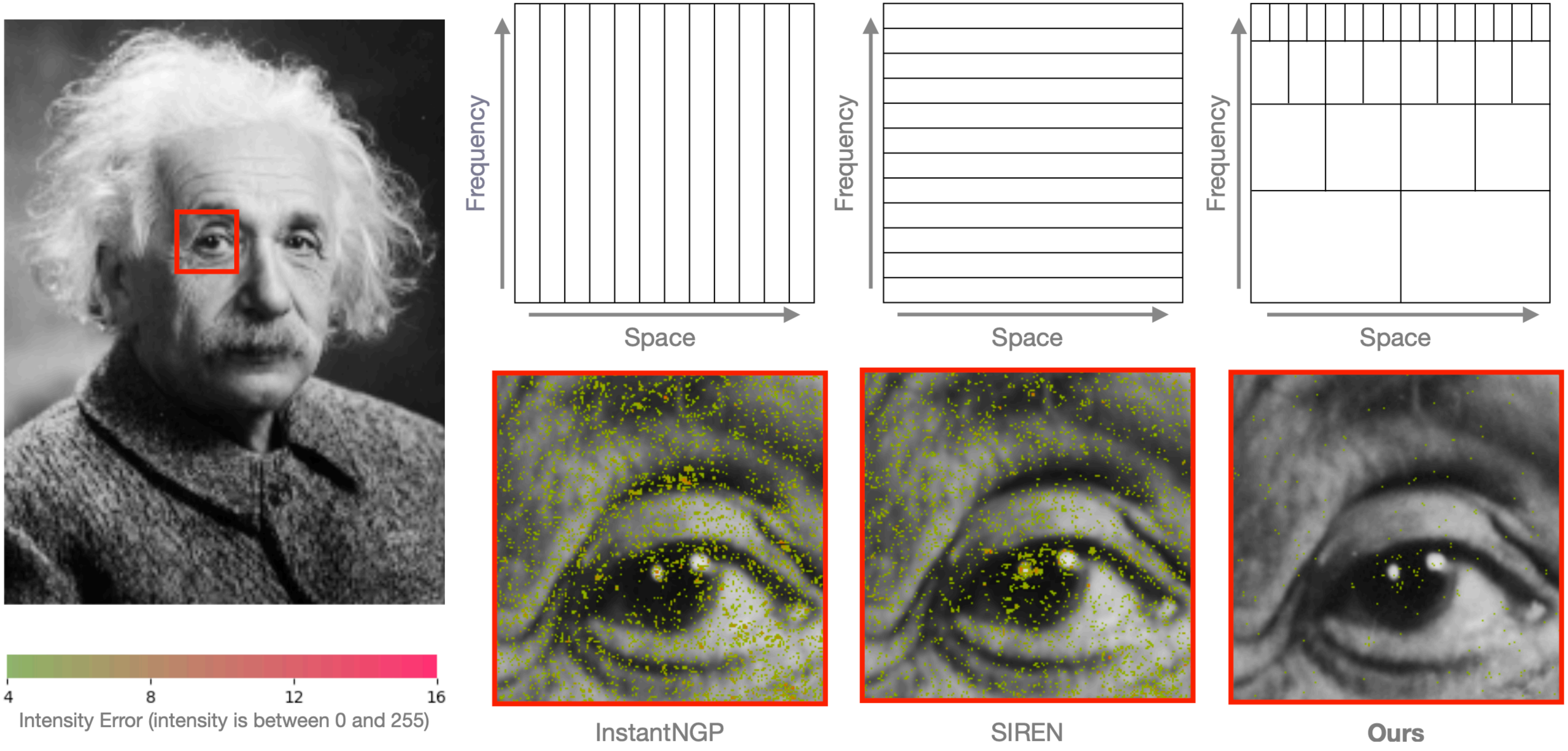
Considering both space and frequency



Considering both space and frequency



Considering both space and frequency



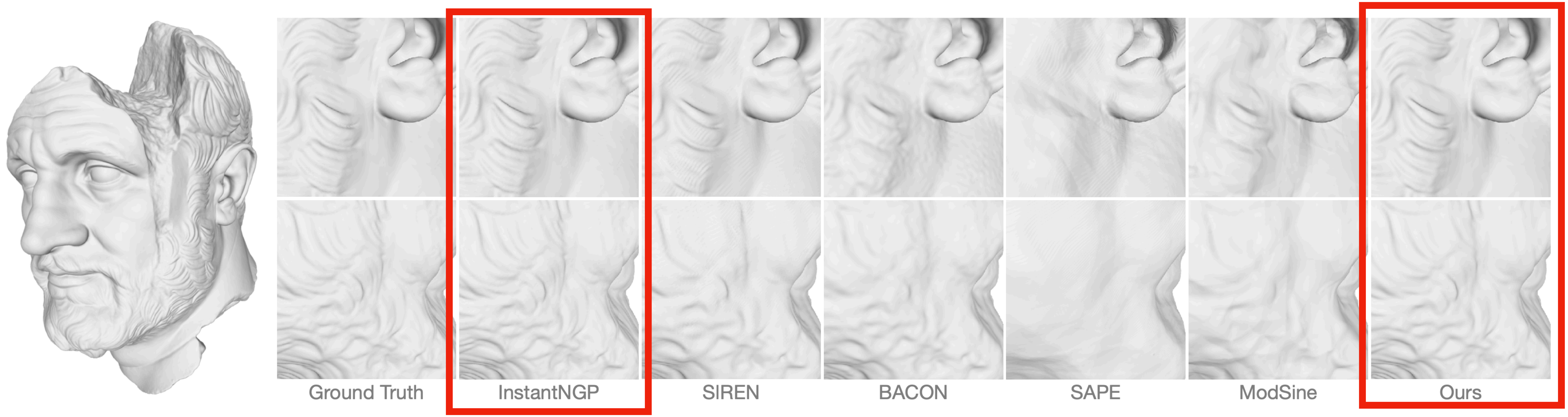
THE UNIVERSITY OF BRITISH COLUMBIA

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

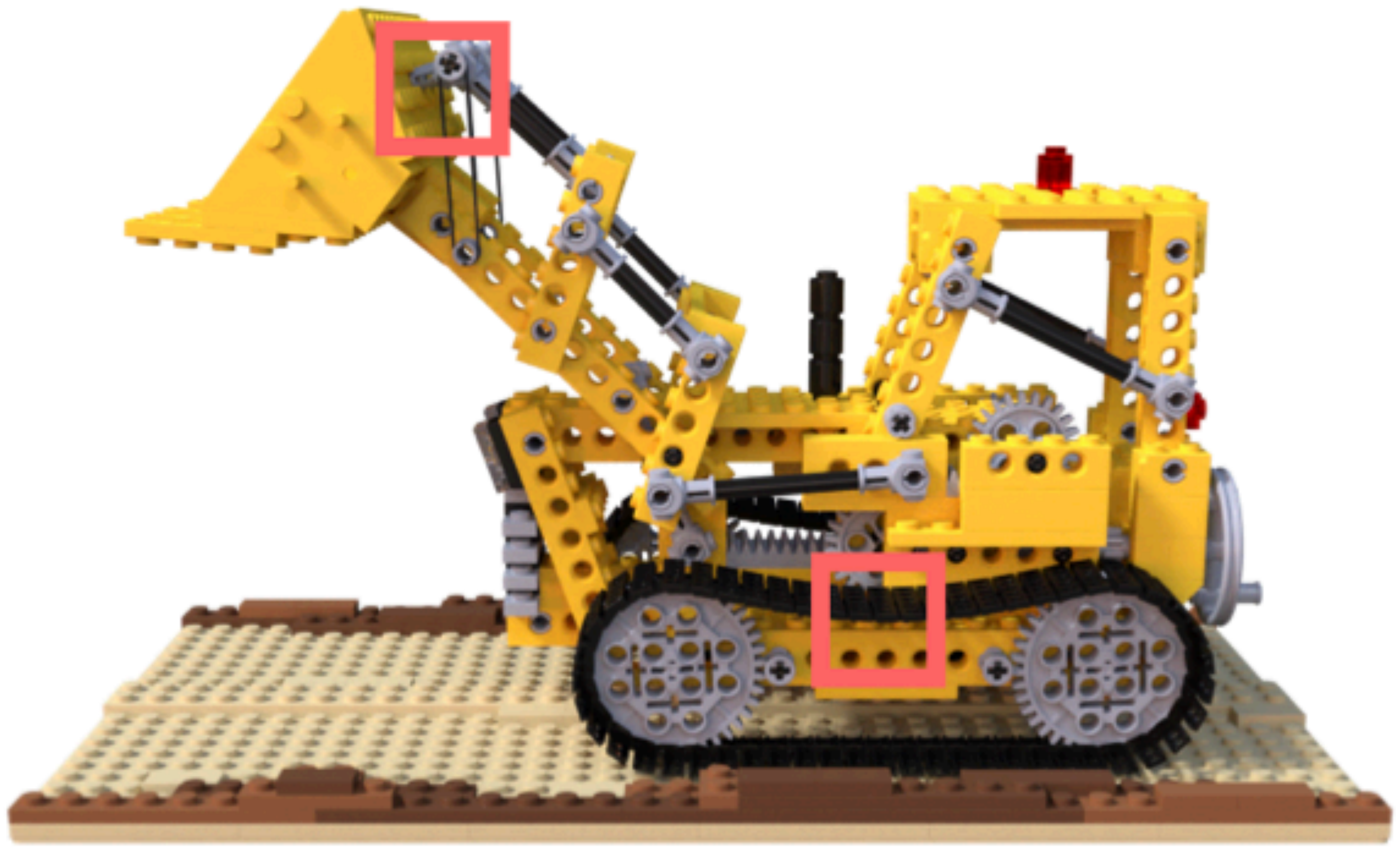
Results — 2D image fitting

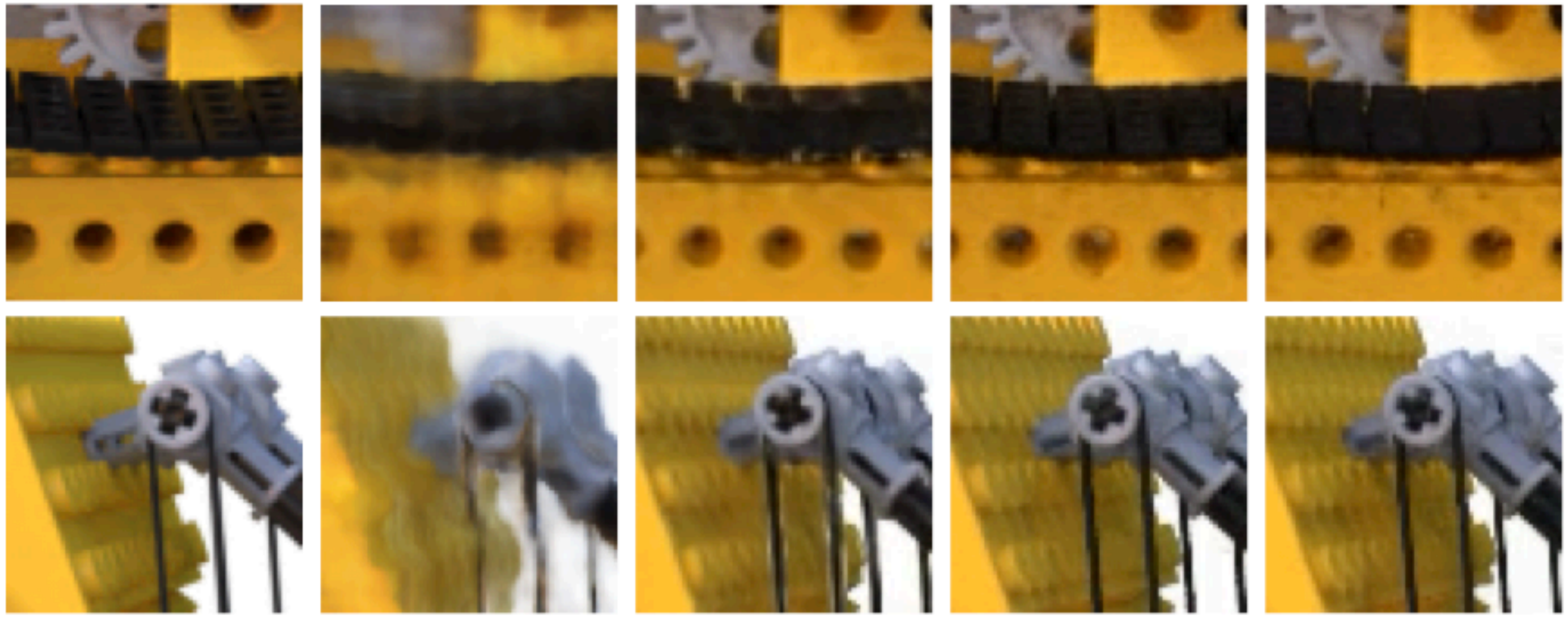
Ground Truth | InstantNGP | SIREN | SAPE | ModSine | Ours

Results — 3D shape fitting



Results — Novel View Synthesis





Neural Fields

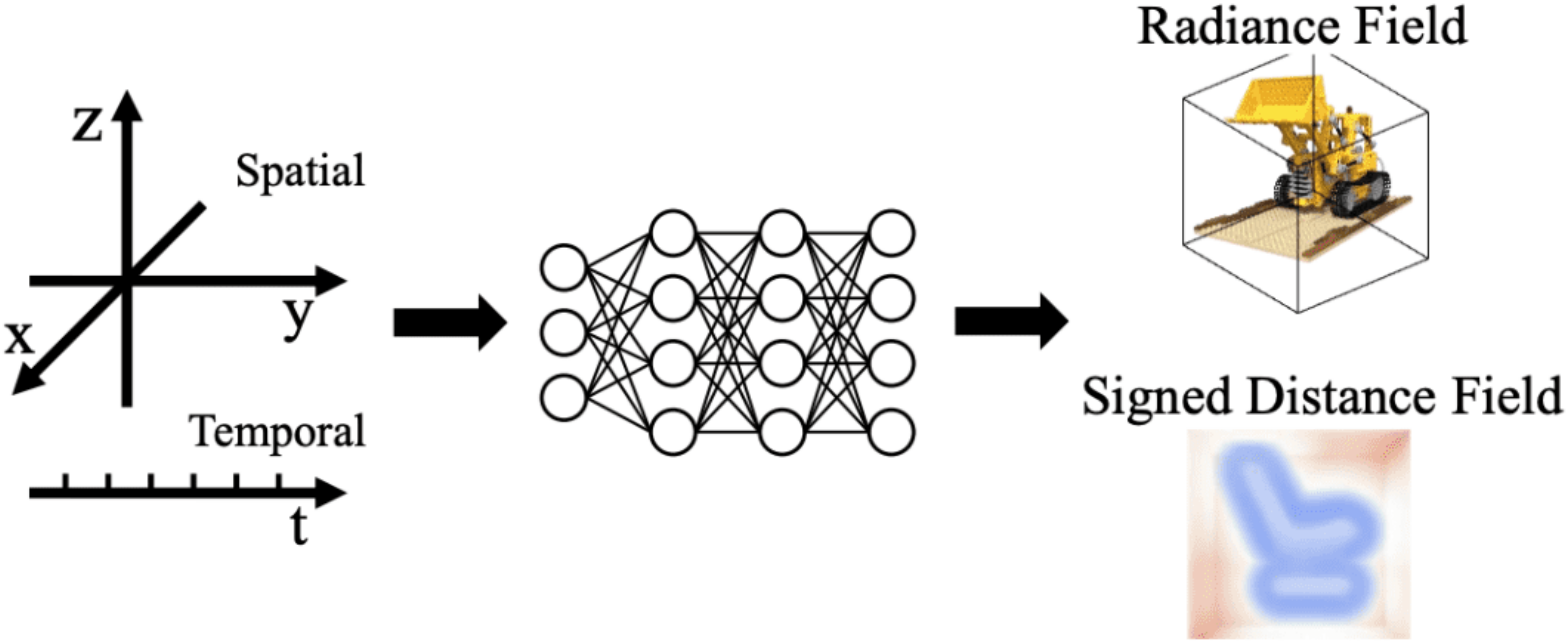
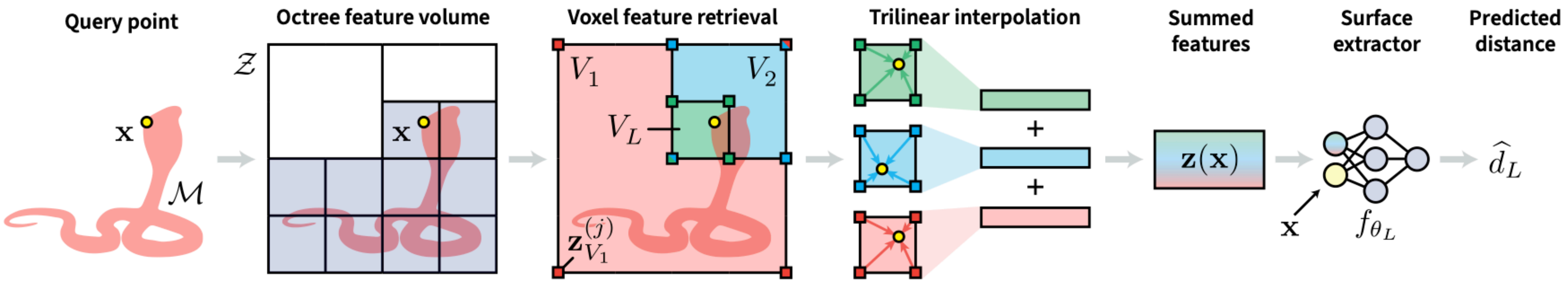
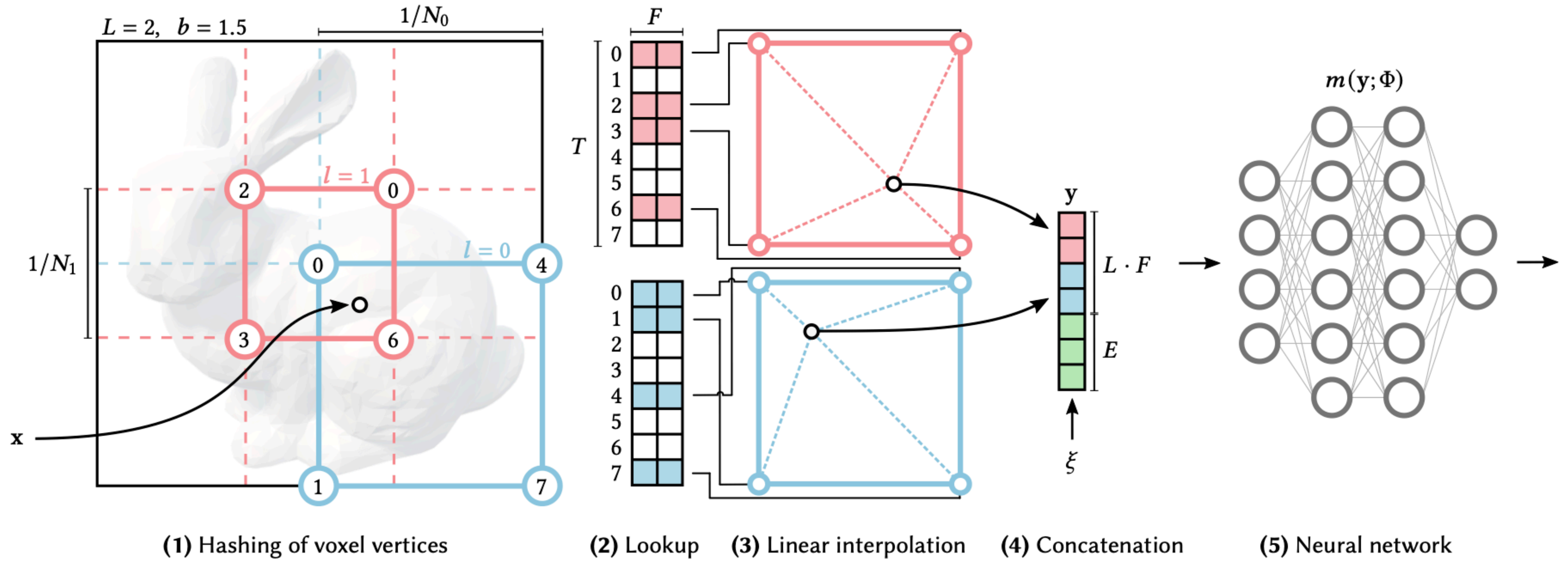


Image from Xie et al., 2022, reproduced for educational use.

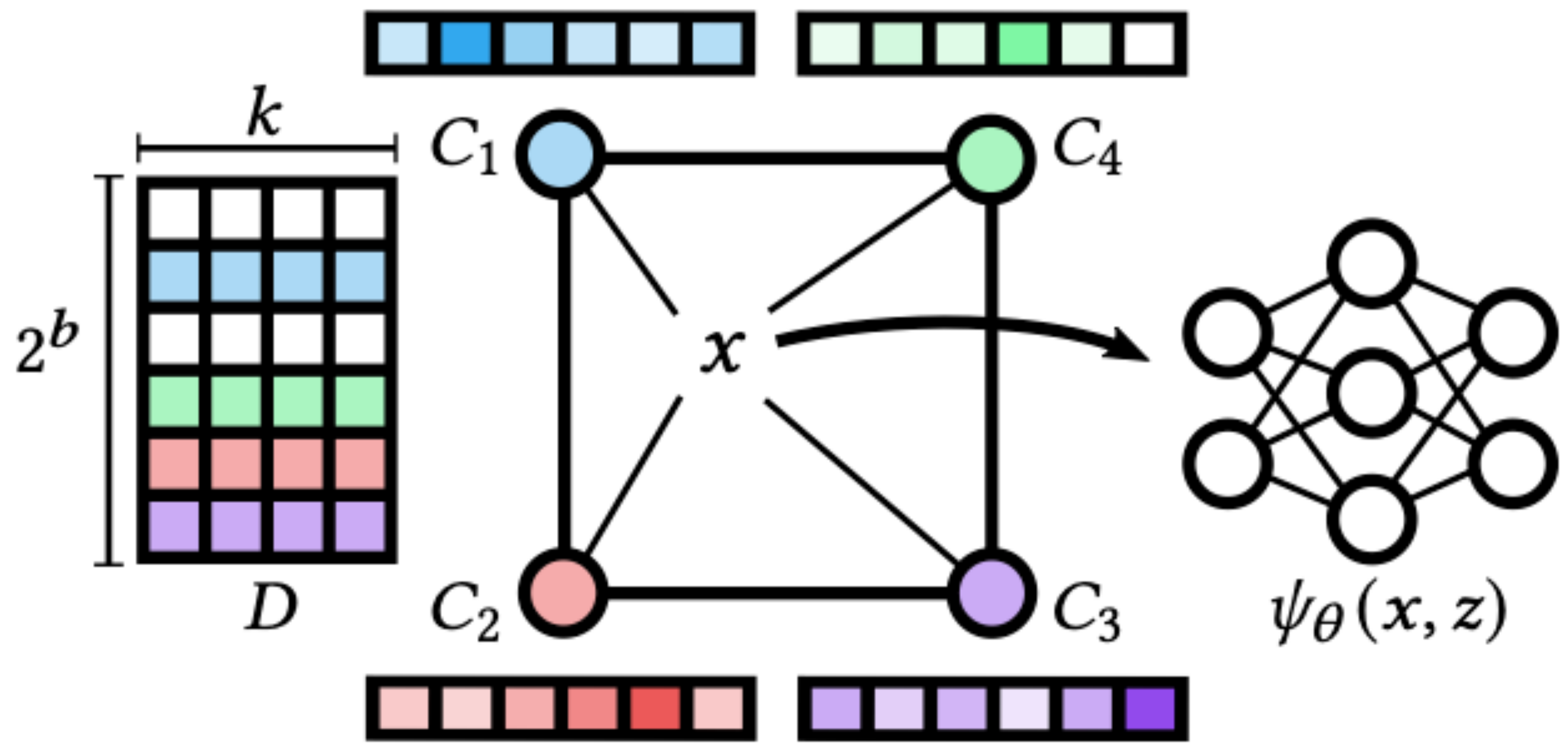
State-Of-The-Art Methods



[Takikawa et al., 2021]



[Müller et al., 2022]

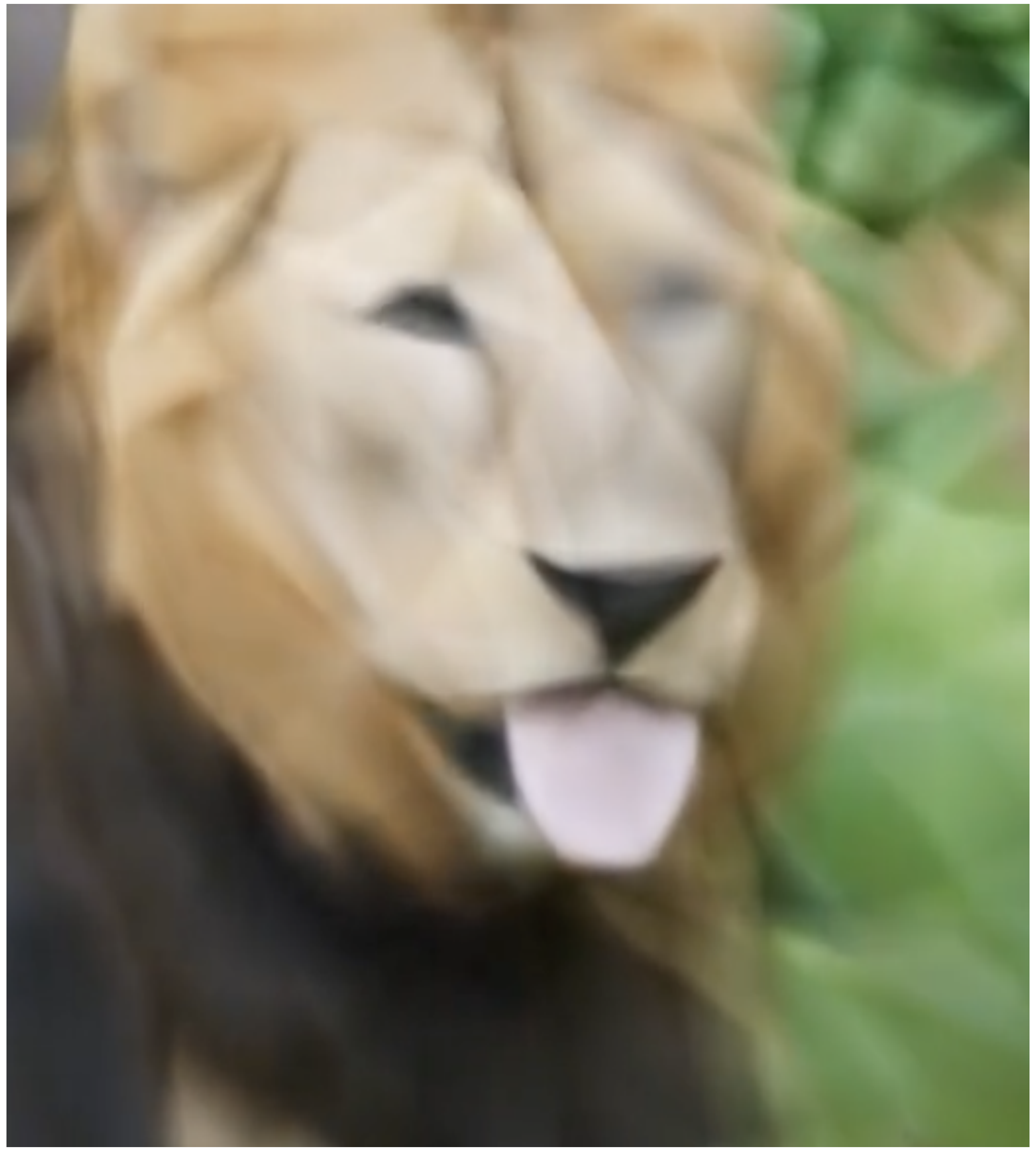


[Takikawa et al., 2022]

A missing component

[Muller et al., 2022]

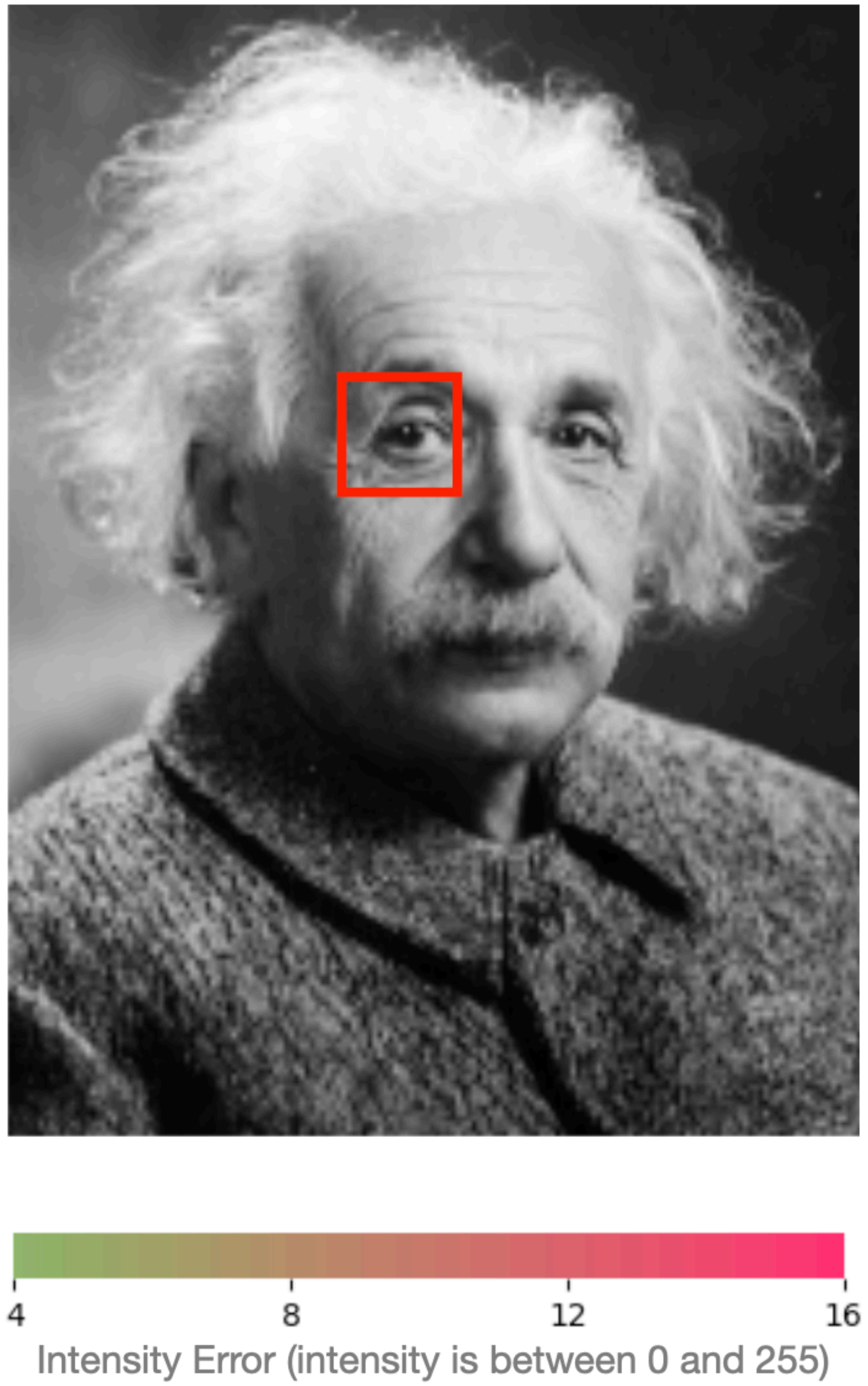
Extremely fast training via hash grids



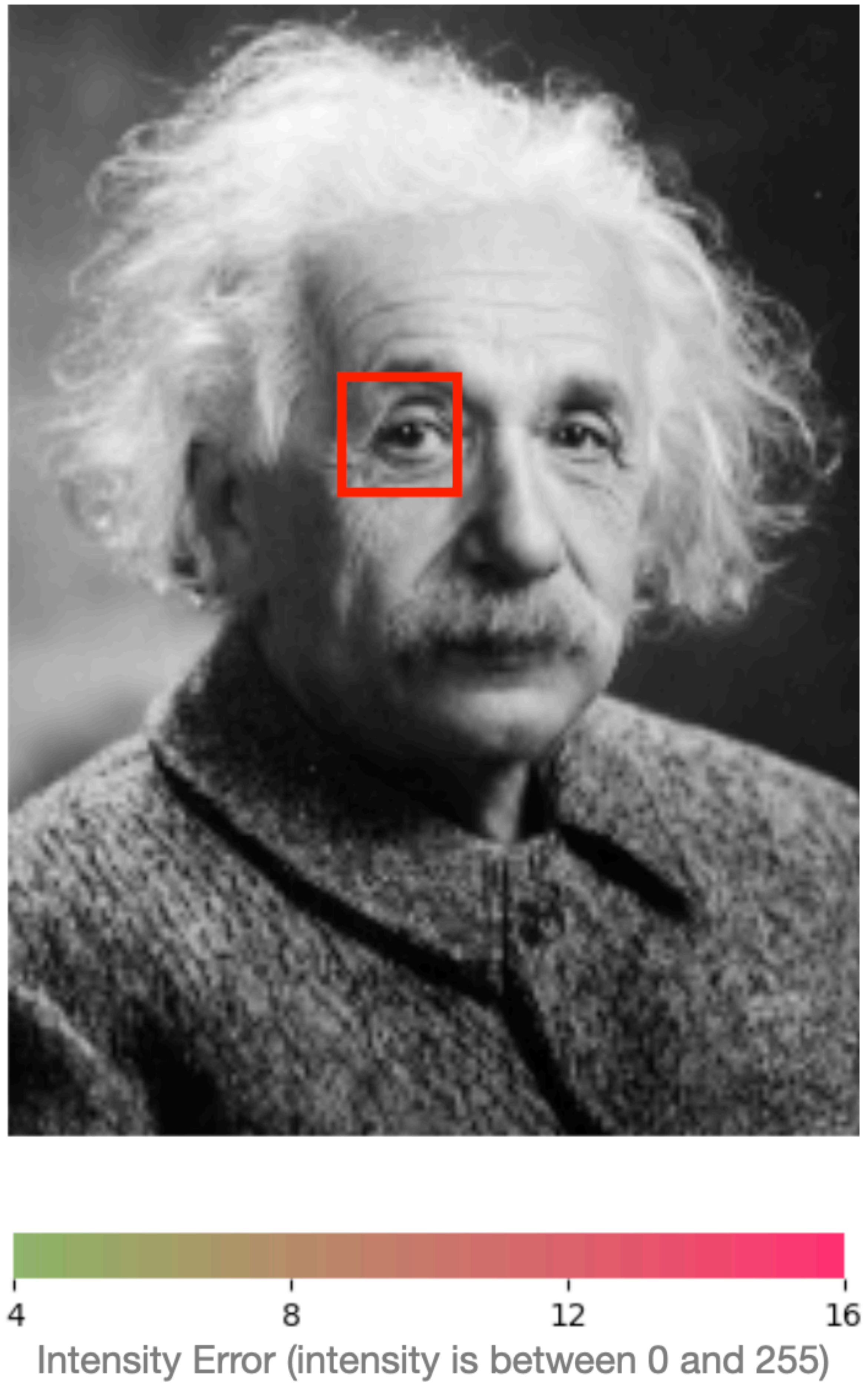
[Mildenhall et al., 2020]

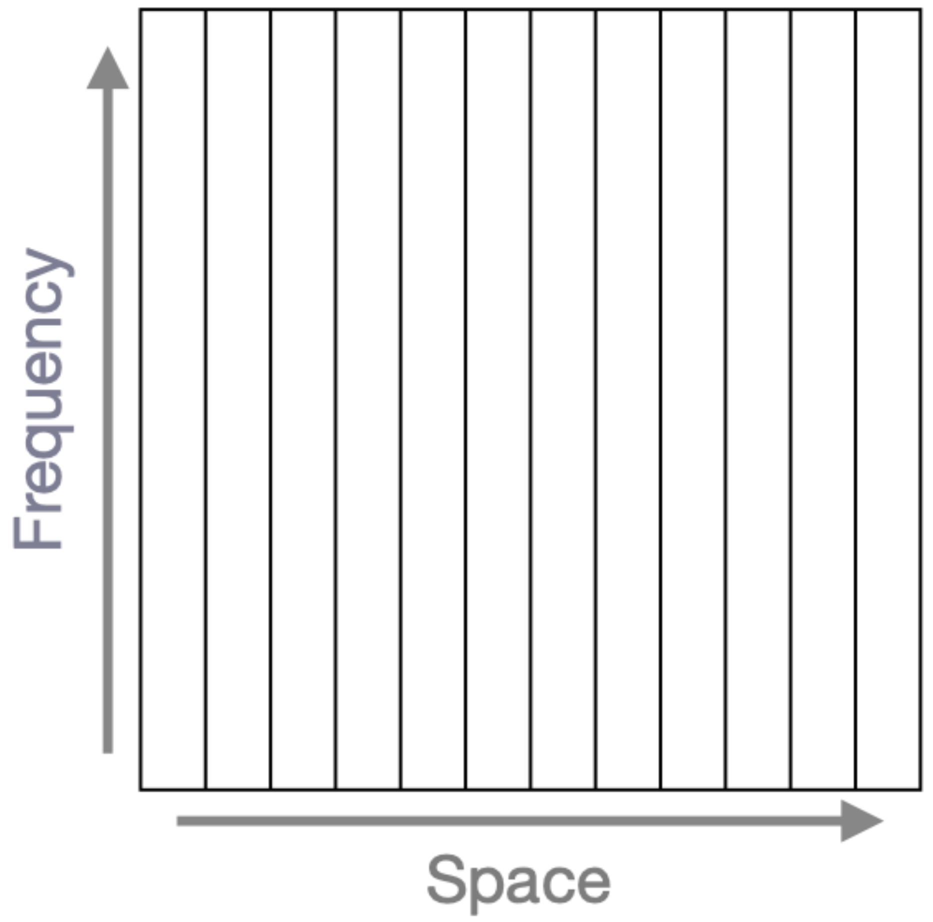
Frequency is important

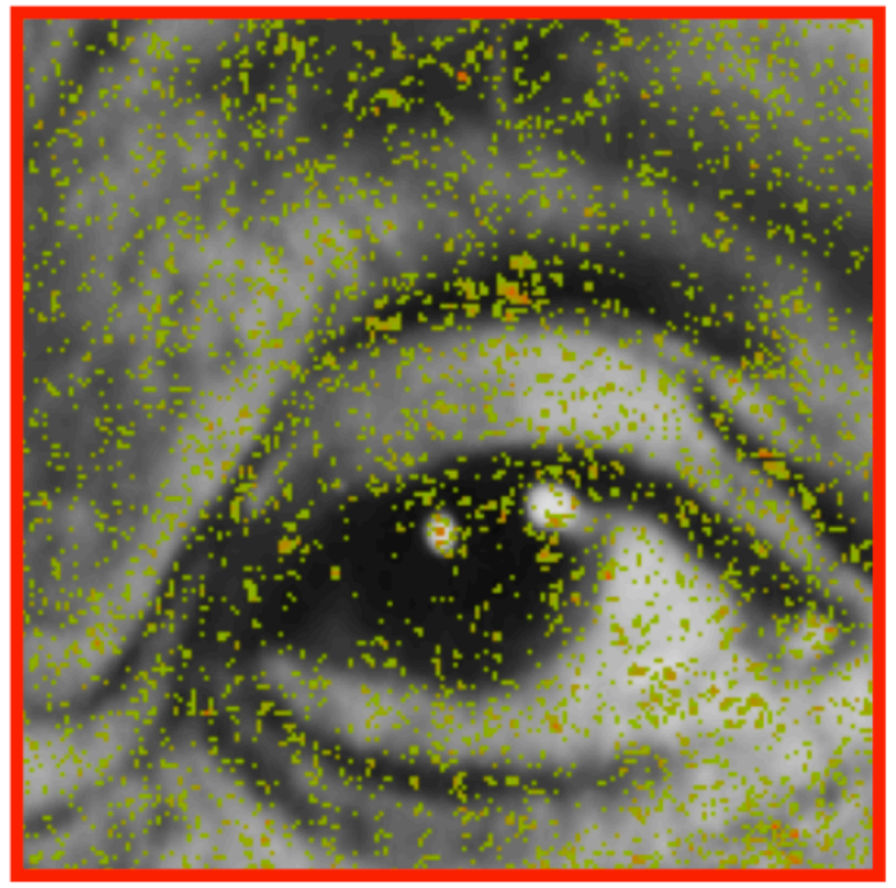
Considering both space and frequency



Considering both space and frequency



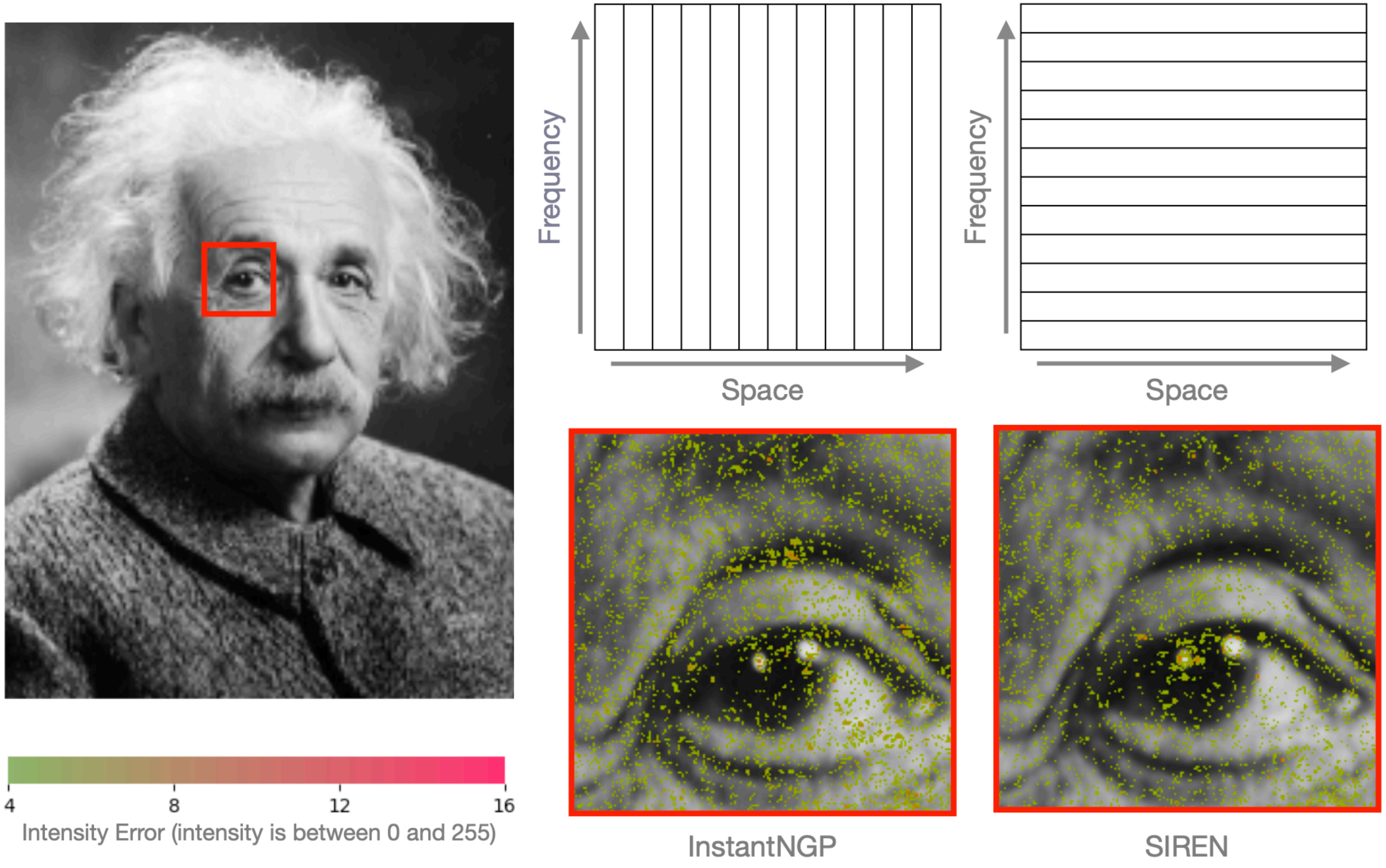




InstantNGP

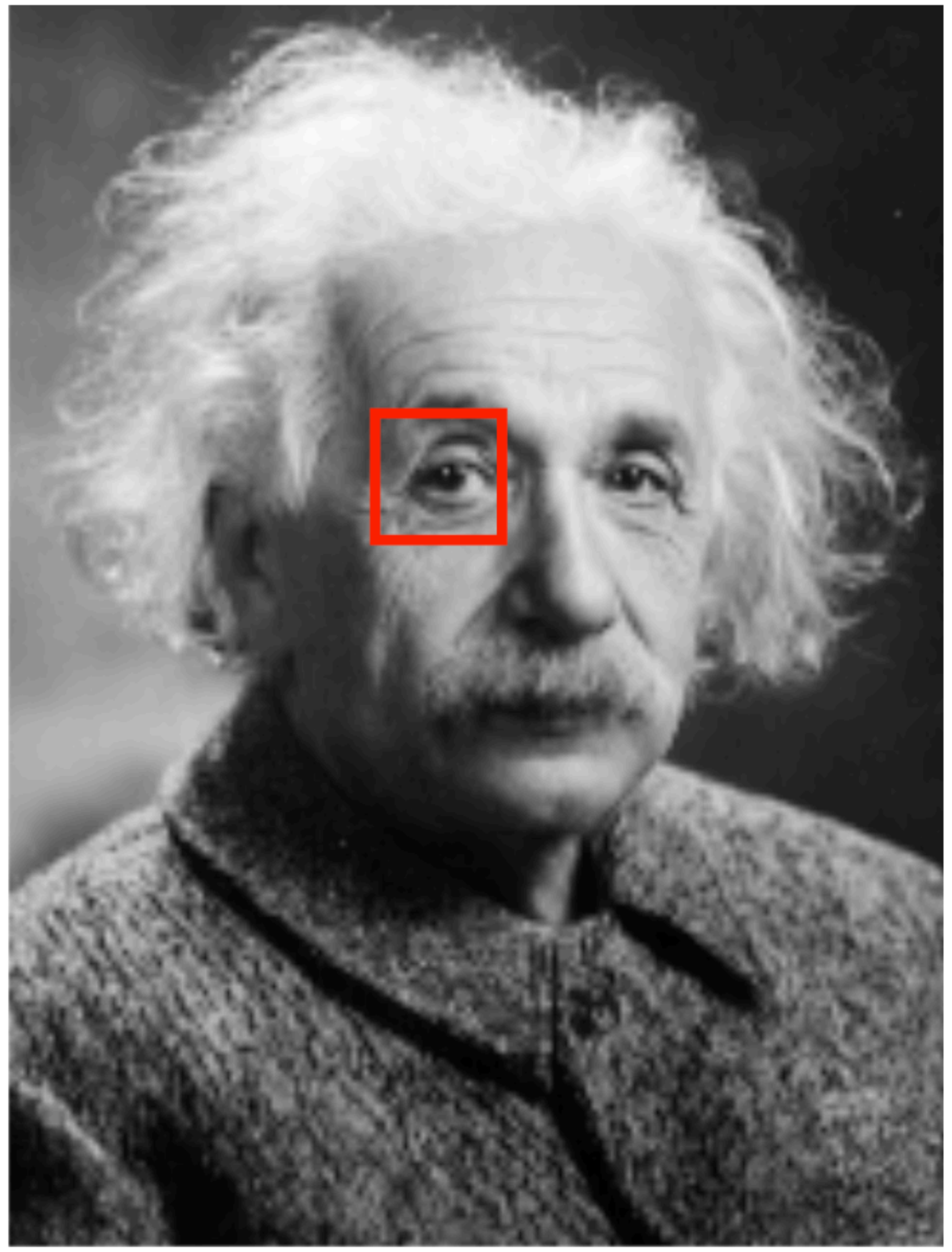
Use <u>hashgrid</u> to divide space into local regions and speed up training

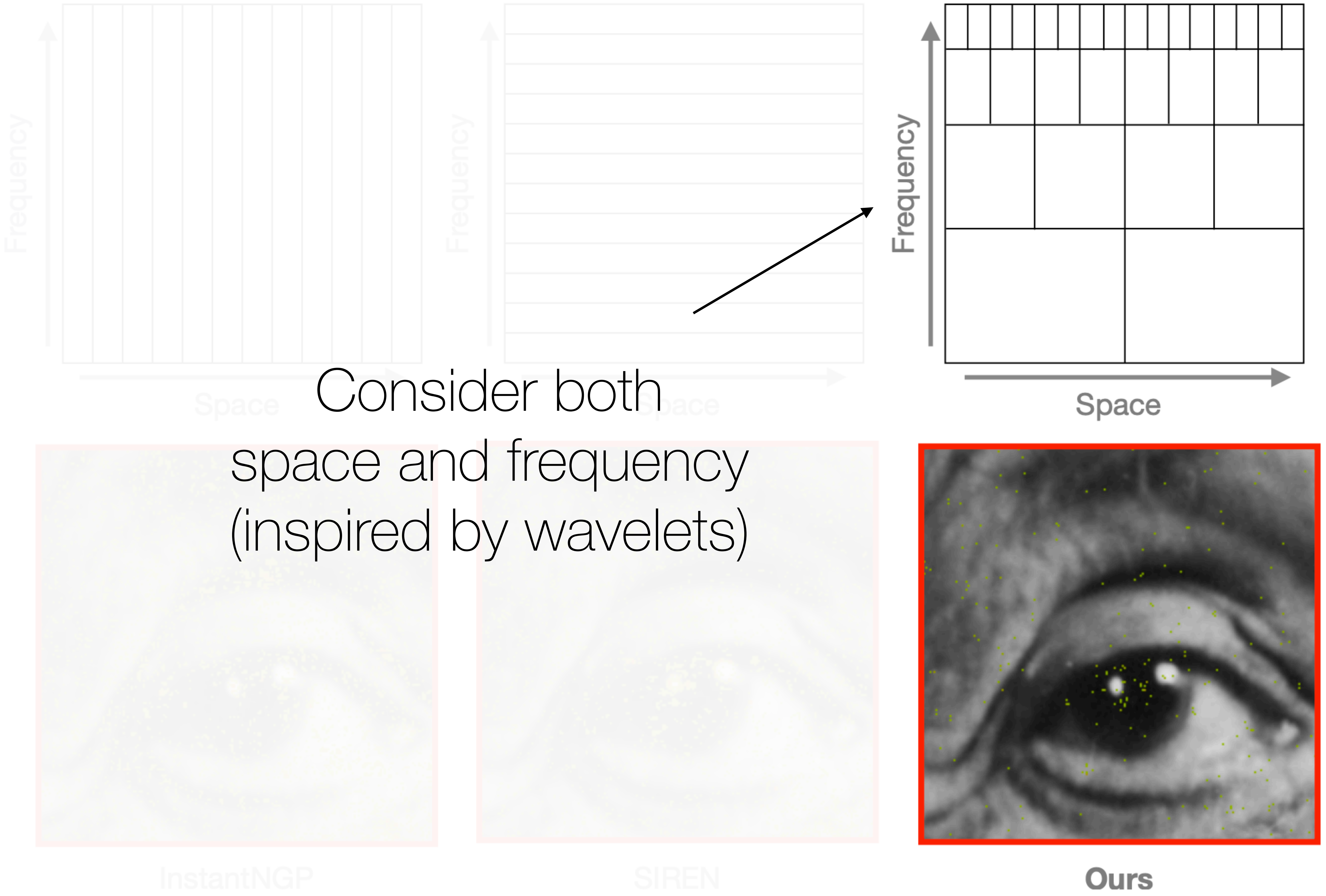
Considering both space and frequency



Analogous to neurons operating in the frequency domain via <u>sine activations</u>

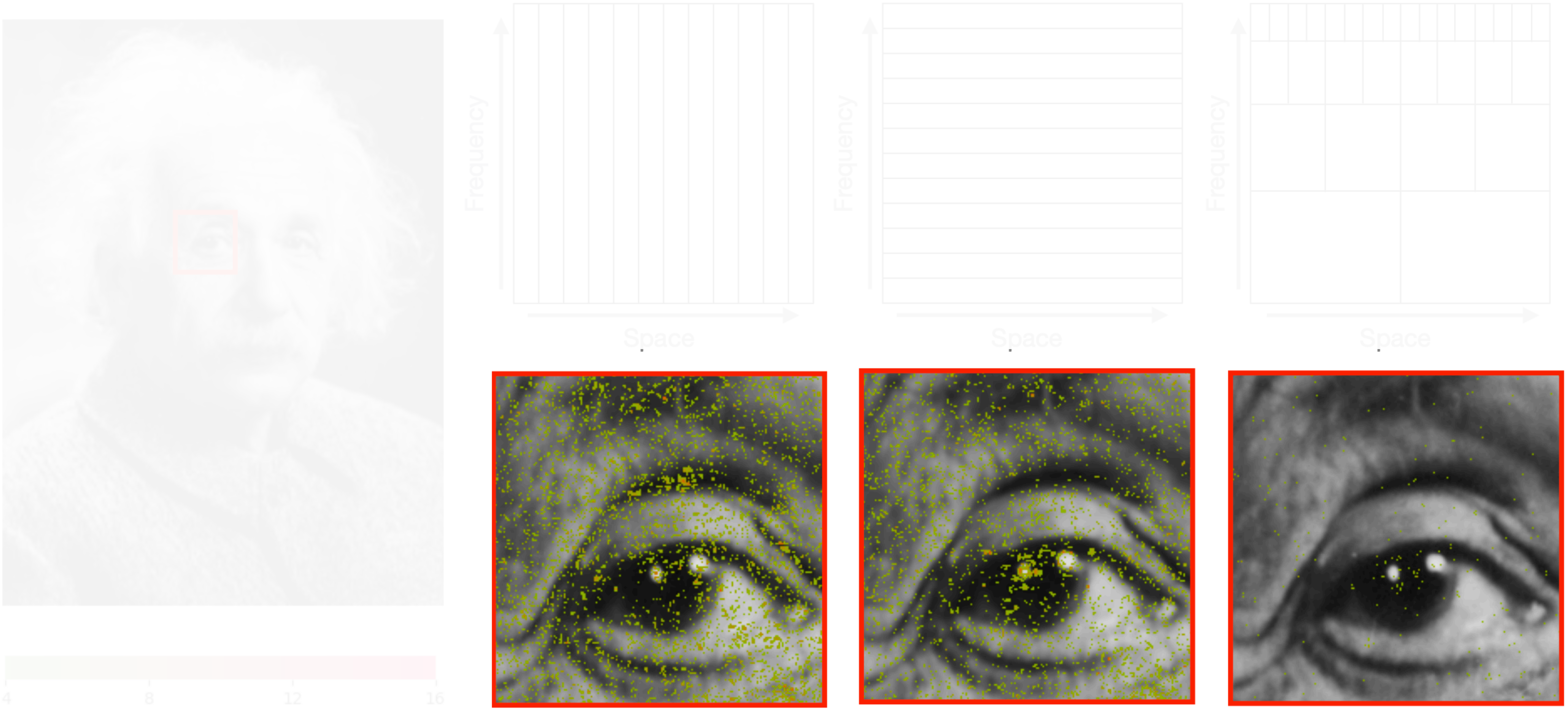
Considering both space and frequency





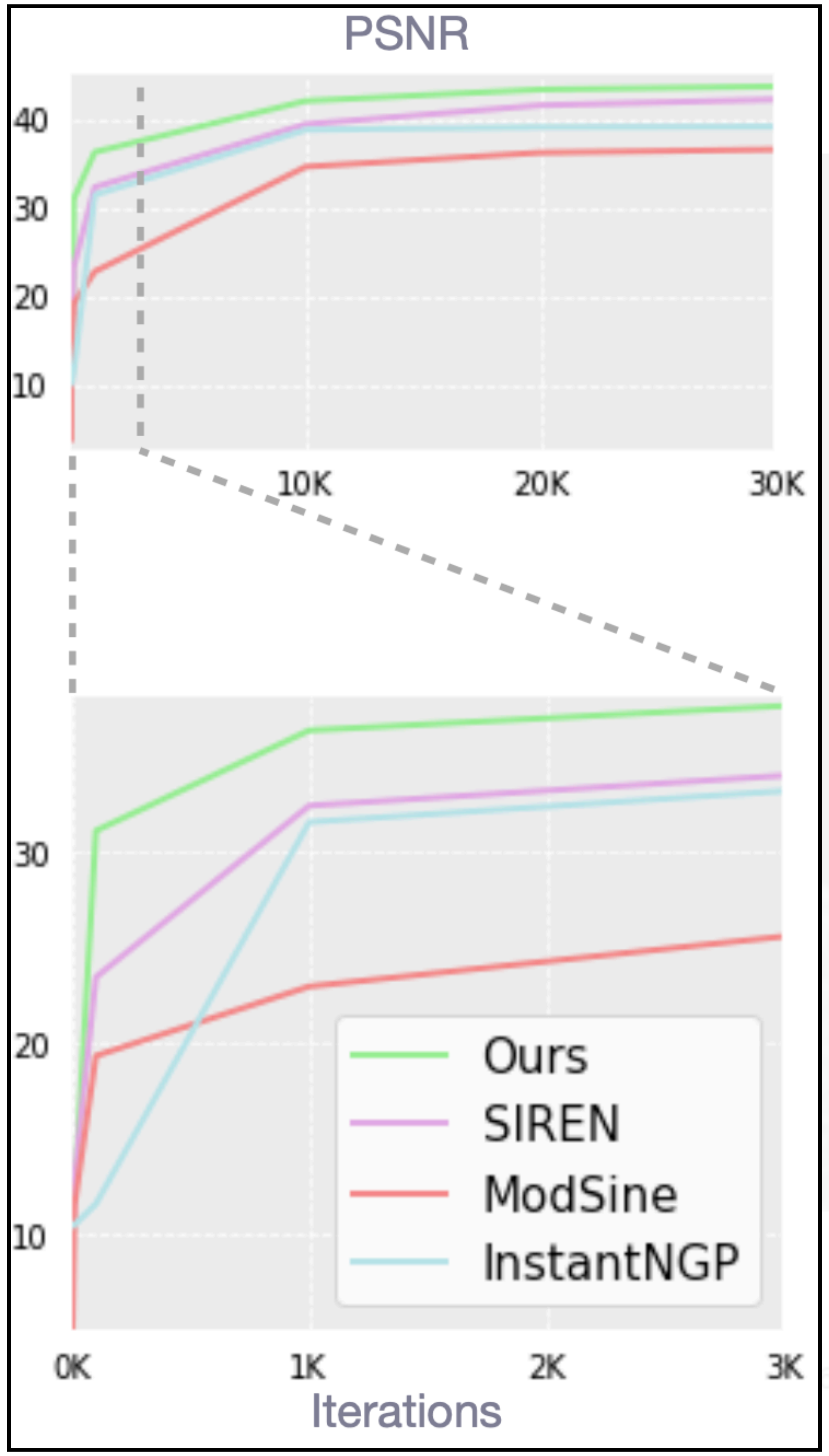
Consider both
space and frequency
(inspired by wavelets)

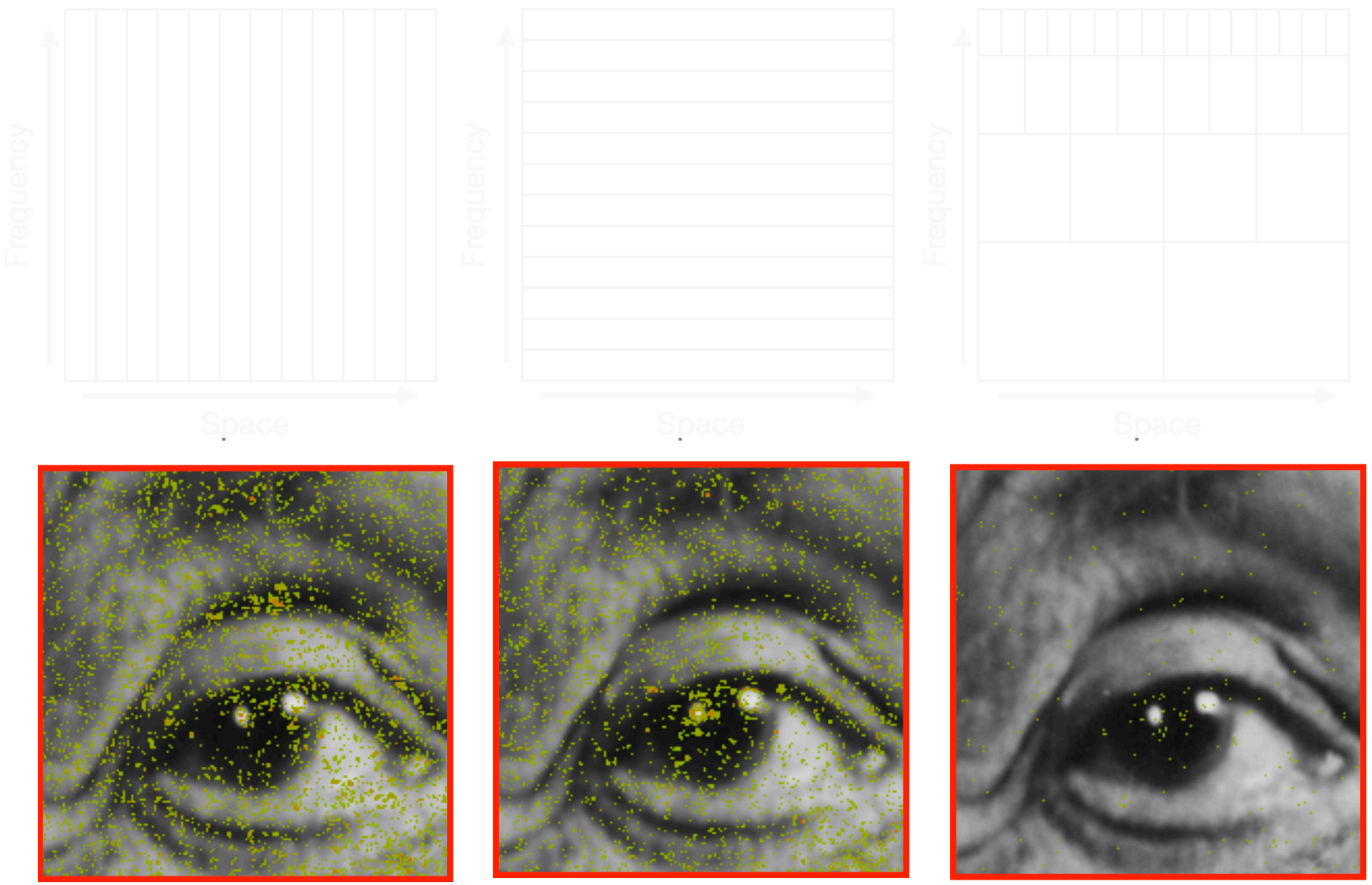
Considering both space and frequency



InstantNGP SIREN **Ours**

Considering both space and frequency

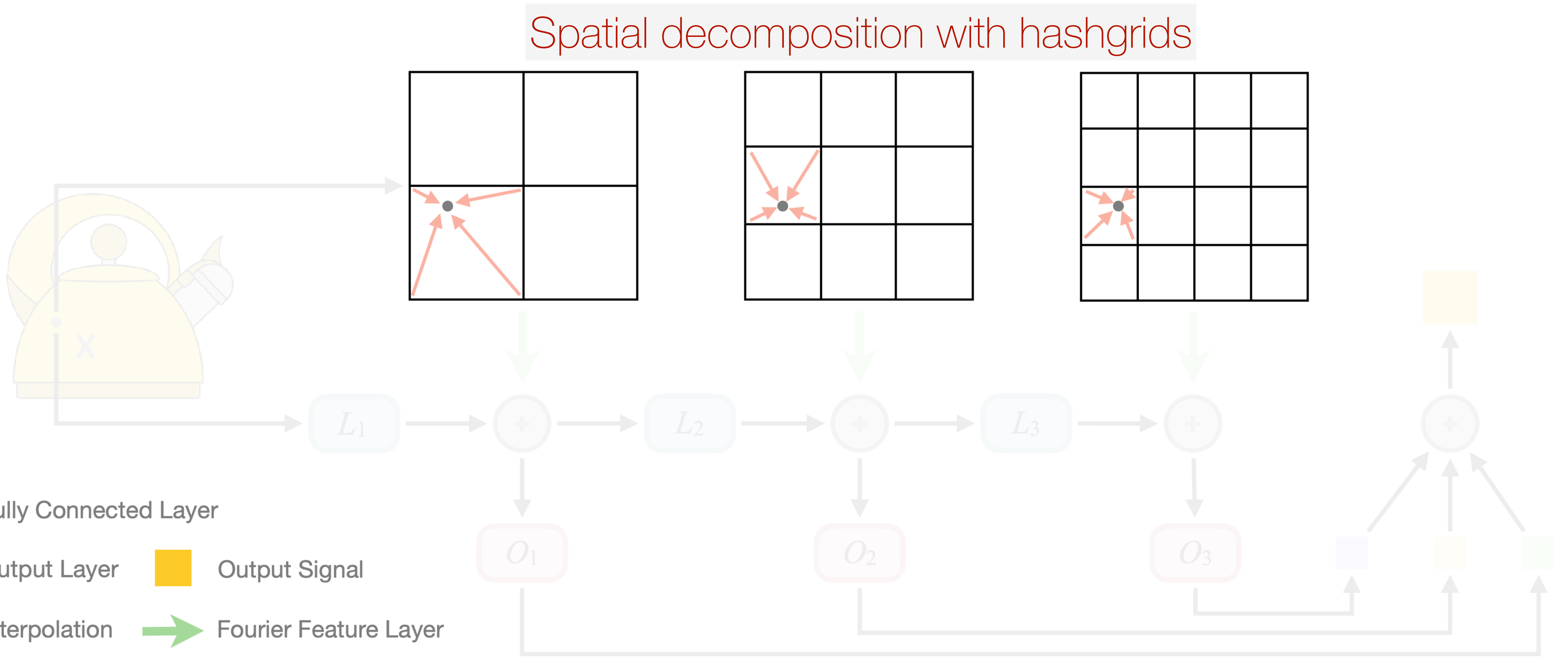




InstantNGP SIREN **Ours**

Implementing the idea as an architecture

Implementing the idea as an architecture



L Fully Connected Layer

O Output Layer

Output Signal

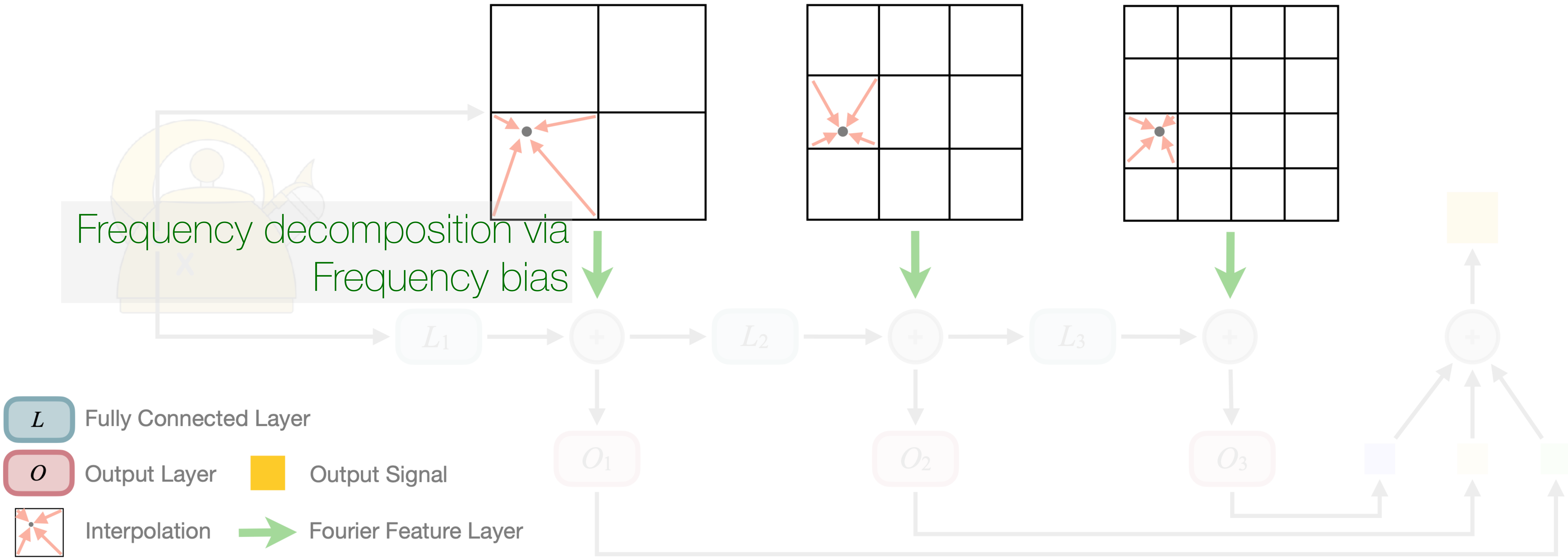
Interpolation

Fourier Feature Layer

Implementing the idea as an architecture

Spatial decomposition with hashgrids

Frequency decomposition via
Frequency bias



Implementing the idea as an architecture

Results — 2D image fitting

Ground Truth | InstantNGP | SIREN | SAPE | ModSine | Ours

Results — 2D image fitting

	Tokyo				Albert			
	Size (MB)↓	PSNR↑	SSIM↑	LPIPS↓	Size (MB)↓	PSNR↑	SSIM↑	LPIPS↓
InstantNGP [33]	36.0	33.38	0.9452	0.201	3.7	41.61	0.9623	0.152
SIREN [41]	5.2	28.52	0.8921	0.474	5.0	42.51	0.9661	0.478
SAPE [15]	**3.2**	21.64	0.5357	0.745	**3.2**	34.26	0.9219	0.399
ModSine [30]	3.5	23.23	0.7587	0.607	4.2	36.74	0.9184	0.438
Ours	4.1	31.57	0.9403	0.187	-	-	-	-
Ours	10.0	**33.62**	**0.9555**	**0.141**	3.7	**43.83**	**0.9763**	**0.142**

Results — 2D image fitting

	Tokyo				Albert			
	Size (MB)↓	PSNR↑	SSIM↑	LPIPS↓	Size (MB)↓	PSNR↑	SSIM↑	LPIPS↓
InstantNGP [33]	36.0	33.38	0.9452	0.201	3.7	41.61	0.9623	0.152
SIREN [41]	5.2	28.52	0.8921	0.474	5.0	42.51	0.9661	0.478
SAPE [15]	**3.2**	21.64	0.5357	0.745	**3.2**	34.26	0.9219	0.399
ModSine [30]	3.5	23.23	0.7587	0.607	4.2	36.74	0.9184	0.438
Ours	4.1	31.57	0.9403	0.187				
Ours	10.0	**33.62**	**0.9555**	**0.141**	3.7	**43.83**	**0.9763**	**0.142**

Best quality among similar #params

THE UNIVERSITY OF BRITISH COLUMBIA

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Results — 2D image fitting

	Tokyo				Albert			
	Size (MB)↓	PSNR↑	SSIM↑	LPIPS↓	Size (MB)↓	PSNR↑	SSIM↑	LPIPS↓
InstantNGP [33]	36.0	33.38	0.9452	0.201	3.7	41.61	0.9623	0.152
SIREN [41]	5.2	28.52	0.8921	0.474	5.0	42.51	0.9661	0.478
SAPE [15]	**3.2**	21.64	0.5357	0.745	**3.2**	34.26	0.9219	0.399
ModSine [30]	3.5	23.23	0.7587	0.607	4.2	36.74	0.9184	0.438
Ours	4.1	31.57	0.9403	0.187	-	-	-	-
Ours	10.0	**33.62**	**0.9555**	**0.141**	3.7	**43.83**	**0.9763**	**0.142**

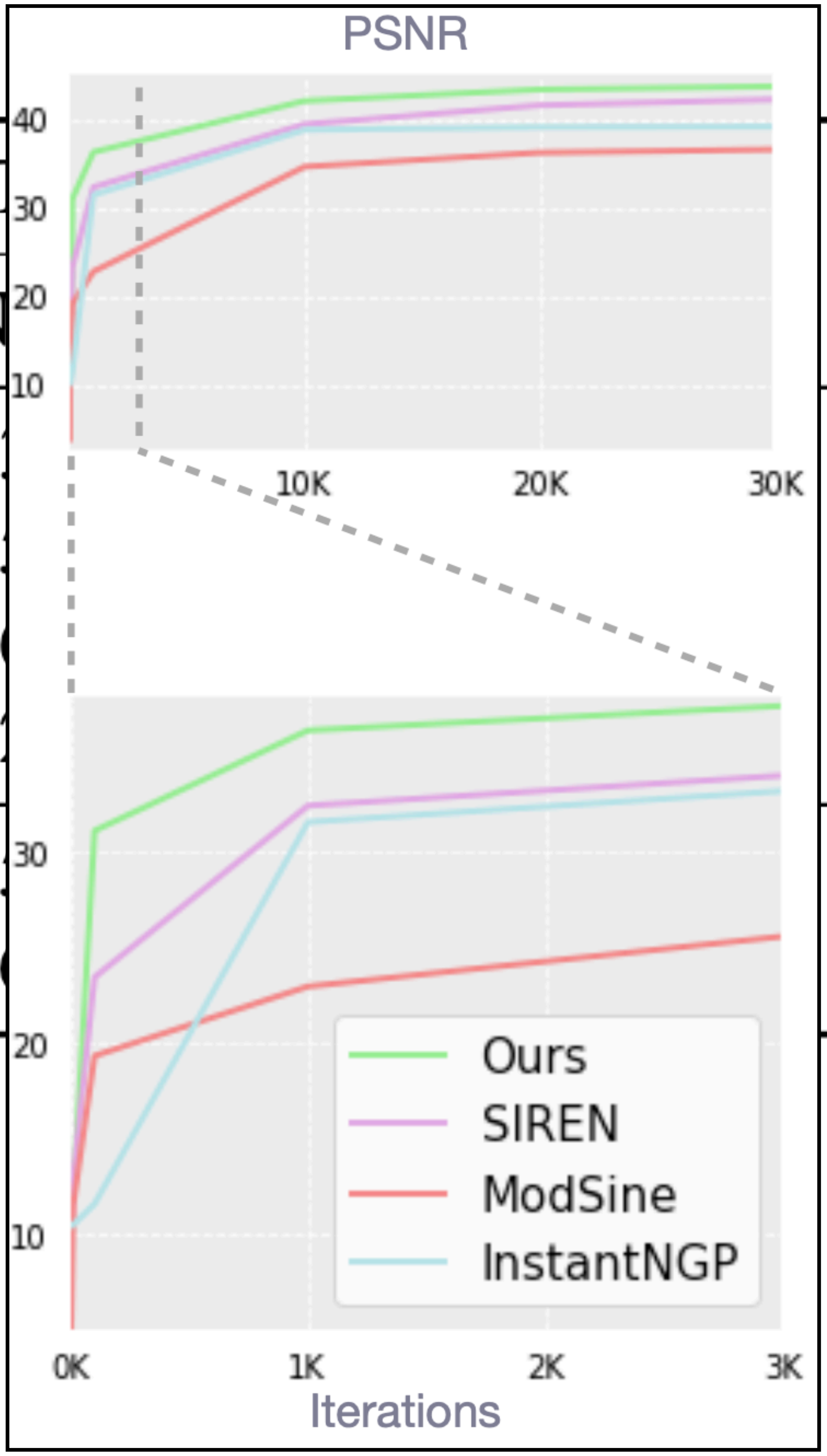
Improved quality with <1/3 #param

THE UNIVERSITY OF BRITISH COLUMBIA

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Results — 2D image fitting

	Size (MB)↓	PSN	Size (MB)↓	PSNR↑	SSIM↑	LPIPS↓
			Albert			
InstantNGP [33]	36.0	33.	3.7	41.61	0.9623	0.152
SIREN [41]	5.2	28.	5.0	42.51	0.9661	0.478
SAPE [15]	**3.2**	21.	**3.2**	34.26	0.9219	0.399
ModSine [30]	3.5	23.	4.2	36.74	0.9184	0.438
Ours	4.1	31.	-	-	-	-
Ours	10.0	**33.**	3.7	**43.83**	**0.9763**	**0.142**



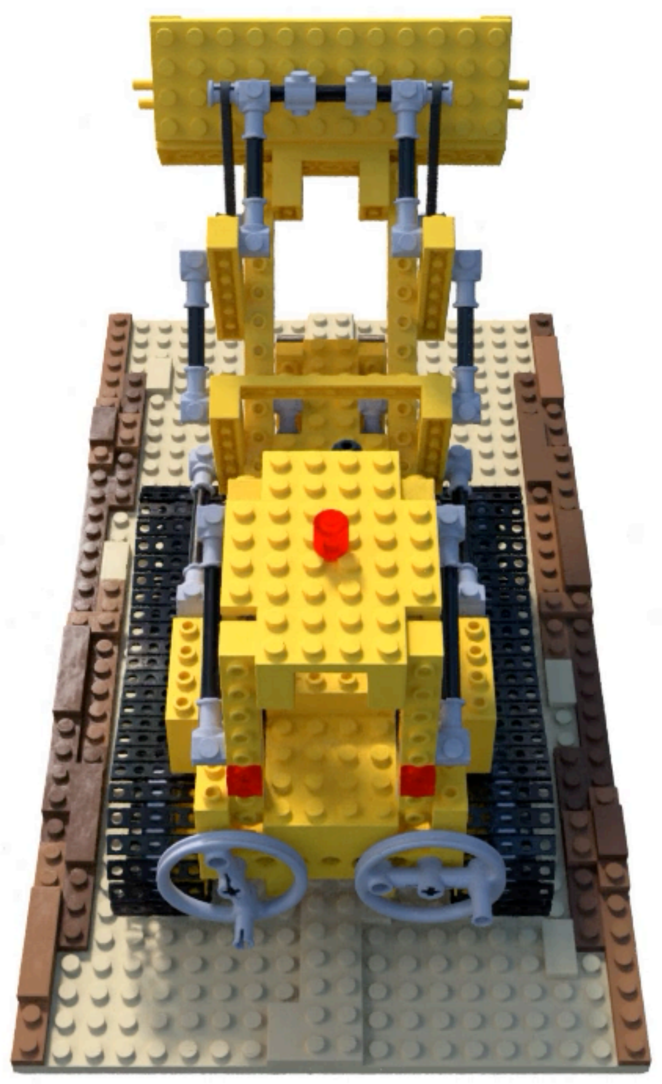
Best result with faster convergence

Results — 3D shape fitting

On par, but with 1/40 #params

	Size (MB)↓	Asian Dragon			Beard Man		
		F-score↑	IoU↑	Cham dist↓	F-score↑	IoU↑	Cham dist↓
InstantNGP [33]	46.5	0.8714	**1.0**	0.00191	**0.999**	0.9970	**0.00272**
Ours	**1.4**	0.8717	**1.0**	**0.00189**	**0.999**	**0.9985**	**0.00272**

Results — NeRF

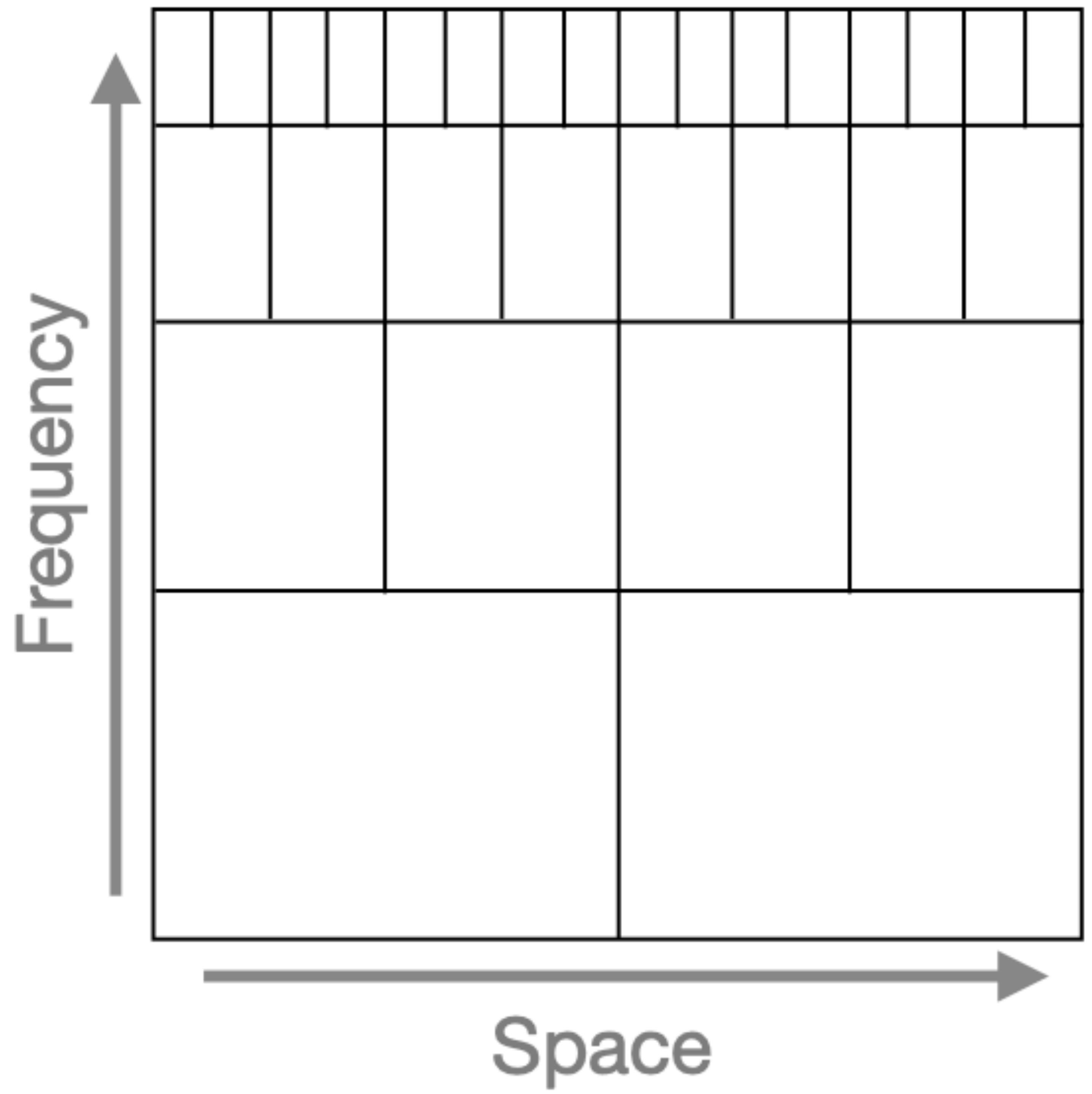


	Steps	Size (MB) ↓	Time ↓	PSNR ↑	SSIM ↑
NeRF [36]	300k	5.0	> 30h	31.01	0.947
Plenoxels [45]	128k	778.1	11.4m	31.71	0.958
DVGO [50]	30k	612.1	15m	31.95	0.957
InstantNGP [37]	30k	46.6	3.4m	32.08	0.955
Ours	30k	14.7	13.1m	32.04	0.955

Similar performance but with less #param and fewer steps

Summary

- We propose a novel framework that decomposes the modeled signal both spatially and frequency-wise
- Our method achieves better trade-off between result quality and network complexity on three tasks



Summary

- We propose a novel framework that decomposes the modeled signal both spatially and frequency-wise
- Our method achieves better trade-off between result quality and network complexity on three tasks

- Released codes & Implementation details: https://github.com/ubc-vision/NFFB
- Project webpage: https://zhijiew94.github.io/NFFB