Wave Interpolation Neural Operator: Interpolated Prediction of Electric Fields Across Untrained Wavelengths

Joonhyuk Seo
Hanyang University
{yhy258}@hanyang.ac.kr

Chanik Kang
Hanyang University
{chanik}@hanyang.ac.kr

Dongjin Seo
Yale University
{dongjin.seo}@yale.edu

Haejun Chung*
Hanyang University
{haejun}@hanyang.ac.kr

Abstract

Existing surrogate solvers are limited to performing inference at fixed simulation conditions, such as wavelengths, and require retraining for different conditions. To address this, we propose Wave Interpolation Neural Operator (WINO), a novel surrogate solver enabling simulation condition interpolation across a continuous spectrum of broadband wavelengths. WINO introduces the Fourier Group Convolution Shuffling operator and a new conditioning method to efficiently predict electric fields from both trained and untrained wavelength data, achieving significant improvements in parameter efficiency and spectral interpolation performance.[2]

1 Introduction

Metalens is at the forefront of next-generation optical technologies. Designing metalenses involves conducting electromagnetic simulations grounded in the partial differential equations (PDEs) called Maxwell's equations. Generally, the simulations numerically solve Maxwell's equations using finite-difference frequency-domain (FDFD) Hughes et al. [2019, 2018] or finite-difference time-domain (FDTD) Kunz and Luebbers [1993], Oskooi et al. [2010] solvers. However, the requirement of high computational resources makes large-scale simulations and designs intractable Kang et al. [2024a]. To address this challenge, there is a need for surrogate solvers that operate faster than traditional simulations and provide precise results.

Most previous studies have employed conventional data-driven neural networks to predict plausible results based on a constrained number of design factors Jiang et al. [2019], Kang et al. [2024b]. However, as demonstrated in a prior study Kovachki et al. [2023], these models frequently fail to guarantee discretization invariance. This poses difficulties in learning PDEs with mesh information that changes according to wavelength or resolution. Furthermore, they typically perform well within the fixed simulation settings in which they are trained. Generating a new dataset and retraining the model to predict outcomes for new simulation parameters are necessary.

In this work, we propose a parameter-efficient surrogate solver named Wave Interpolation Neural Operator (WINO). For the first time, WINO enables the interpolation of simulation parameters by training with discrete wavelength simulation data through a novel conditioning method, allowing it to

*Corresponding author.

[2]The code and dataset are available at https://anonymous.4open.science/r/WINO_interp-4B11.

D3S3: Data-driven and Differentiable Simulations, Surrogates, and Solvers @ NeurIPS 2024.

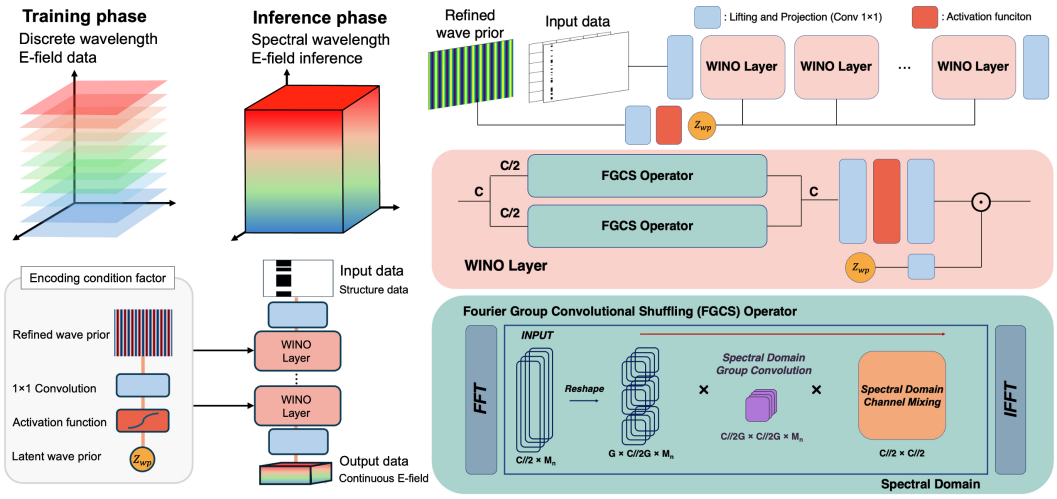

Figure 1: (left) Overall framework of WINO. (right) Schematic illustration of (right top) WINO including (right middle) WINO layer and (right bottom) Fourier Group Convolution Shuffling (FGCS) operator.

infer across a continuous spectrum of broadband wavelengths. The main contributions of our study are as follows:

1. To the best of our knowledge, our work presents the first broadband electromagnetic surrogate solver.
2. Parameter efficiency improves by approximately 74% over the previous state-of-the-art (SOTA) model.
3. Performance improvements: 81% enhancement for untrained wavelengths and 10% enhancement for trained wavelengths compared to the SOTA.

2 Methodology

2.1 Wave Interpolation Neural Operator

We propose Wave Interpolation Neural Operator (WINO), which enables the interpolated prediction of unseen wavelengths from simulation observations of discrete wavelengths (Figure 1). WINO comprises two main components: (1) the Fourier Group Convolution Shuffling (FGCS) operator that is designed to be highly parameter-efficient while maintaining performance, and (2) a novel conditioning method that enables the seamless interpolation of data conditions.

We define the WINO layers as $\mathcal{L}^{(l)}(z^{(l)}) = z^{(l)} + W_2^{(l)}\sigma(W_1^{(l)}\tilde{\mathcal{K}}^{(l)}(z^{(l)}) + b_1^{(l)}) + b_2^{(l)}$. We employ a residual connection after two linear transformations with an activation function to enhance the capacity of our model, like Vaswani et al. [2017], Tran et al. [2021]. Detailed information on problem setting is in the Appendix section titled “Problem Setting."

2.2 Kernel Integral Operator $\tilde{\mathcal{K}}$ of WINO

The kernel integral operator in the original FNO is defined as follows:

$$\mathcal{K}^{(l)} = \mathcal{F}^{-1}(R^{(l)} \cdot \mathcal{F}),\ R \in \mathbb{C}^{M_v \times M_h \times C \times C}$$

where $\mathcal{F}$ is a Fourier transformation, M_v, M_h, and C denote the number of frequency modes of vertical and horizontal components and the number of channels, respectively. Thus, the number of parameters in a 2D problem is $C^2 M_v M_h$. This leads to a significantly high parameter complexity. Thus, as illustrated in Figure 1 (middle), we construct the cross-shaped block Gu et al. [2022] to achieve superior parameter efficiency. We split the inputs into two chunks $z^{(l)} = [z_h^{(l)}, z_v^{(l)}]$ along the channel axis and factorize the Fourier transforms over the horizontal and vertical dimensions $\tilde{\mathcal{K}}^{(l)}(z^{(l)}) = [\mathcal{K}_h^{(l)}(z_h^{(l)}), \mathcal{K}_v^{(l)}(z_v^{(l)})]$. Therefore, the weights $R_h^{(l)} \in \mathbb{C}^{M_h \times \frac{1}{2}C \times \frac{1}{2}C}$ and

$R_v^{(l)} \in \mathbb{C}^{M_v \times \frac{1}{2}C \times \frac{1}{2}C}$ of both two blocks have $\frac{C^2}{4}(M_h + M_v)$ parameters, fewer than the $C^2 M_h M_v$ parameters in the kernel integral operator of an FNO.

Furthermore, we propose an efficient Fourier Group Convolution Shuffling (FGCS) operator that is extremely parameter-efficient while maintaining high performance, as shown in Figure 1 (middle and bottom). We group the input in the Fourier domain along the channel dimension using the group parameter G, as inspired by Xie et al. [2017], Guibas et al. [2021], Kim et al. [2024], but we share weights across the divided groups. Let the horizontal and vertical components be represented by $\mathcal{F}(z_h^{(l)}) \in \mathbb{C}^{H \times M_h \times \frac{1}{2}C}$ and $\mathcal{F}(z_v^{(l)}) \in \mathbb{C}^{M_v \times W \times \frac{1}{2}C}$, respectively. By introducing a group parameter G, horizontal component becomes $R_h^{(l)} \in \mathbb{C}^{M_h \times \frac{1}{2G}C \times \frac{1}{2G}C}$, and the vertical component is transformed into $R_v^{(l)} \in \mathbb{C}^{M_v \times \frac{1}{2G}C \times \frac{1}{2G}C}$. Consequently, we construct a kernel integral operator with a parameter complexity of $C^2 M/G^2$. However, the weight-sharing scheme in the group convolution leads to a significant information loss, resulting in a drastic performance decline. Thus, we further perform channel shuffling in the Fourier domain in order to compensate for the lost information, which can be functioned through simple matrix multiplication along the channel dimension with $W_{\text{ch}}^{(l)} \in \mathbb{C}^{C \times C}$.

2.3 Novel Conditioning Method for Interpolating Untrained Wavelengths' Simulations

The changes in the field wave patterns with varying wavelengths, along with resonance and diffraction phenomena during wave-material interactions, are highly nonlinear. This nonlinearity presents challenges in predicting simulation fields for unseen wavelengths during training. To deal with the challenges, first, we propose a conditioning method for interpolation in the parameter space of the broadband spectrum, as shown in Figure 1 (top and middle). We assume that the condition data have high spatial structure similarity with fields and accurately reflect the features of the wave patterns. Under these assumptions, preserving the spatial structure of the condition information when constructing high-dimensional embeddings is necessary to accurately reflect the field changes according to continuous wavelengths. Thus, we create a common embedding, which is then injected into all layers, using a 1×1 convolutional layer with a nonlinear activation function. Then, each layer adaptively transforms the embedding using an additional 1×1 convolutional layer. We highlight that employing only channel-wise operations guarantees the preservation of the spatial structure of conditional information, which is substantially similar to the field. The final conditional embeddings created in each layer are element-wisely multiplied with the output of the feed-forward network before the residual connection in each layer.

Next, we construct condition information that accurately reflects the wave characteristics by employing a wave prior Gu et al. [2022]. The wave prior is an artificial wave pattern derived from the solution of the wave equation and is expressed as $\mathcal{W}_x = e^{j\frac{2\pi\sqrt{\boldsymbol{\epsilon}_r}}{\lambda}x\mathbf{1}^T\Delta l_x}$ and $\mathcal{W}_z = e^{j\frac{2\pi\sqrt{\boldsymbol{\epsilon}_r}}{\lambda}\mathbf{1}z^T\Delta l_z}$. Here, $\boldsymbol{\epsilon}_r \in \mathbb{C}^{H \times W}$ represents the relative permittivity of the structures at each coordinate, λ is the wavelength, Δl_x and Δl_z are the simulation's step length, and x and z are the coordinates. Because the wave prior emulates the solution of the wave equation, it is substantially similar to the wave patterns of the field generated by the simulator, which result from light propagation according to wavelengths, achieving the precluded assumption. However, the wave prior considers only the structures' permittivity values without accounting for the diffraction phenomena, because it is defined by simultaneously reflecting coordinates. The wave prior exhibits sharp and physically implausible features in the regions where the permittivity values change (Figure 8 (top left and right). In addition, since the simulation's permittivity information is already provided as input data, including it in the wave prior is redundant. Hence, we propose a refined wave prior that excludes the ϵ_r term, defined as $\mathcal{W}_x = e^{j\frac{2\pi}{\lambda}x\mathbf{1}^T\Delta l_x}$ and $\mathcal{W}_z = e^{j\frac{2\pi}{\lambda}\mathbf{1}z^T\Delta l_z}$ (Figure 8 (bottom left)). The refined condition information aims to provide a more physically plausible representation of light behavior in free space.

3 Results and Discussion

In the following, we compare the interpolation performance for unobserved wavelengths of Unet Ho et al. [2020], Gupta and Brandstetter [2022], FNO Li et al. [2020], and F-FNO Tran et al. [2021], commonly used in a surrogate solver study. Moreover, we evaluate WINO against NeurOLight Gu et al. [2022], the state-of-the-art model for single wavelength field prediction in electromagnetic

Model	Param (M)	Test loss		Design region test loss	
		Untrained	Trained	Untrained	Trained
Unet Ho et al. [2020]	11.5999	0.003332	**0.001732**	0.00619	**0.003775**
FNO Li et al. [2020]	3.2868	0.0133	0.007719	0.028339	0.025067
F-FNO Tran et al. [2021]	1.887	0.03309	0.02204	0.043035	0.037091
NeurOLight Gu et al. [2022]	1.65	0.01073	0.001973	0.019503	0.004526
WINO	**0.426**	**0.002035**	0.001774	**0.003812**	0.003929

Table 1: Comparisons of field prediction and interpolation. Untrained represents the loss when predicting fields using unobserved wavelength data, which is the loss for the wavelength interpolation.

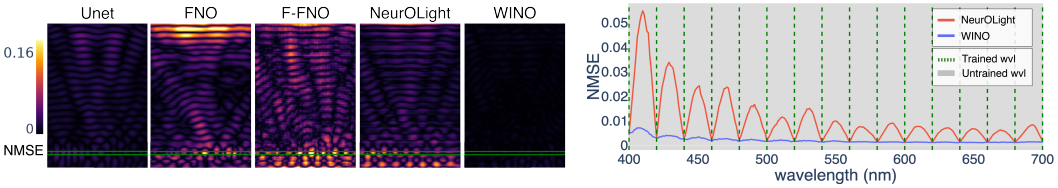

Figure 2: Field interpolation comparisons. (left) Comparison of model errors at 410nm, the unobserved wavelength with the highest error. Darker colors indicate lower errors. Other models show generally high errors in both the design region (green) and the overall area, our model. Nevertheless, WINO demonstrates near-zero errors across the entire area, including the design region. (right) Comparison of entire simulation region errors in the 400-700nm wavelength range.

simulation. Detailed information on model architectures is in the Appendix section titled “Detailed model architectures."

In EM simulation, while predicting the entire E-field is important, accurately predicting the E-field within the design region is even more crucial. For studies involving surrogate solvers to predict E-fields and subsequently optimize designs, error-free predictions in the design region are essential for designing high-performance photonic devices. However, existing models struggle with the nonlinear resonances observed in design regions. These resonances can vary with different source wavelengths, even for the same structure, leading to significant errors in interpolation.

As shown in Figure 2 (left), the overall models, except WINO, exhibit significant errors in the design region. Such resonance patterns can occur in the metalens used in our experiments and other photonic structures designed at the sub-wavelength scale. Therefore, our model can accurately predict resonance patterns during wavelength interpolation in the design region, making it broadly applicable. Futhermore, we aimed to improve the overall error compared to the SOTA and significantly reduce the error in the design region. Our wave interpolation results achieved SOTA performance in both test loss across the entire region and loss in the design region, compared to existing models Li et al. [2020], Tran et al. [2021], Gu et al. [2022] as shown in Figure 2 (right), Table 1 and Figure 3. Specifically, in the design region, which is represented in Table 1 and Figure 9, our model showed an impressive 80.5% improvement over the previous SOTA model, NeurOLight, for untrained wavelengths and an improvement of 13.2% for trained wavelengths. Moreover, we aimed to drastically reduce the number of parameters, achieving a SOTA model with only 0.426 million parameters, a 74.18% reduction compared to NeurOLight and a 96.33% reduction compared to Unet.

4 Conclusion

In this study, we introduce the Wave Interpolation Neural Operator (WINO), a novel surrogate solver designed to address the limitations of existing models in predicting electric fields across a continuous spectrum of wavelengths. By leveraging Fourier Group Convolution Shuffling (FGCS) and a new conditioning method, WINO significantly improves parameter efficiency and interpolation performance. Our experimental results demonstrate that WINO achieves state-of-the-art performance in both test loss and design region test loss compared to existing models such as FNO, F-FNO, and the previous SOTA, NeurOLight. Furthermore, WINO accurately predicts the nonlinear resonance patterns in photonic structures. This capability and SOTA performance position WINO as a powerful tool, paving the way for broadband photonic applications and enabling the design of highly complex photonic structures with a surrogate solver.

References

Tyler W Hughes, Ian AD Williamson, Momchil Minkov, and Shanhui Fan. Forward-mode differentiation of maxwell's equations. *ACS Photonics*, 6(11):3010–3016, 2019.

Tyler W Hughes, Momchil Minkov, Ian AD Williamson, and Shanhui Fan. Adjoint method and inverse design for nonlinear nanophotonic devices. *ACS Photonics*, 5(12):4781–4787, 2018.

Karl S Kunz and Raymond J Luebbers. *The finite difference time domain method for electromagnetics*. CRC press, 1993.

Ardavan F Oskooi, David Roundy, Mihai Ibanescu, Peter Bermel, John D Joannopoulos, and Steven G Johnson. Meep: A flexible free-software package for electromagnetic simulations by the fdtd method. *Computer Physics Communications*, 181(3):687–702, 2010.

Chanik Kang, Chaejin Park, Myunghoo Lee, Joonho Kang, Min Seok Jang, and Haejun Chung. Large-scale photonic inverse design: computational challenges and breakthroughs. *Nanophotonics*, 2024a. doi: doi:10.1515/nanoph-2024-0127.

Jiaqi Jiang, David Sell, Stephan Hoyer, Jason Hickey, Jianji Yang, and Jonathan A Fan. Free-form diffractive metagrating design based on generative adversarial networks. *ACS nano*, 13(8): 8872–8878, 2019.

Chanik Kang, Dongjin Seo, Svetlana V Boriskina, and Haejun Chung. Adjoint method in machine learning: a pathway to efficient inverse design of photonic devices. *Materials & Design*, 239: 112737, 2024b.

Nikola Kovachki, Zongyi Li, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: Learning maps between function spaces with applications to pdes. *Journal of Machine Learning Research*, 24(89):1–97, 2023.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

Alasdair Tran, Alexander Mathews, Lexing Xie, and Cheng Soon Ong. Factorized fourier neural operators. *arXiv preprint arXiv:2111.13802*, 2021.

Jiaqi Gu, Zhengqi Gao, Chenghao Feng, Hanqing Zhu, Ray Chen, Duane Boning, and David Pan. Neurolight: A physics-agnostic neural operator enabling parametric photonic device simulation. *Advances in Neural Information Processing Systems*, 35:14623–14636, 2022.

Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. Aggregated residual transformations for deep neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1492–1500, 2017.

John Guibas, Morteza Mardani, Zongyi Li, Andrew Tao, Anima Anandkumar, and Bryan Catanzaro. Adaptive fourier neural operators: Efficient token mixers for transformers. *arXiv preprint arXiv:2111.13587*, 2021.

Myungjoon Kim, Junhyung Park, and Jonghwa Shin. Efficient fourier neural operators by group convolution and channel shuffling. In *ICLR 2024 Workshop on AI4DifferentialEquations In Science*, 2024. URL `https://openreview.net/forum?id=TyglZCofE3`.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. *arXiv preprint arXiv:2010.08895*, 2020.

Jayesh K Gupta and Johannes Brandstetter. Towards multi-spatiotemporal-scale generalized pde modeling. *arXiv preprint arXiv:2209.15616*, 2022.

Haejun Chung and Owen D Miller. High-na achromatic metalenses by inverse design. *Optics Express*, 28(5):6945–6965, 2020.

Sajan Shrestha, Adam C Overvig, Ming Lu, Aaron Stein, and Nanfang Yu. Broadband achromatic dielectric metalenses. *Light: Science & Applications*, 7(1):85, 2018.

Joonhyuk Seo, Jaegang Jo, Joohoon Kim, Joonho Kang, Chanik Kang, Seongwon Moon, Eunji Lee, Jehyeong Hong, Junsuk Rho, and Haejun Chung. Deep-learning-driven end-to-end metalens imaging. *arXiv preprint arXiv:2312.02669*, 2023.

Arseniy I Kuznetsov, Mark L Brongersma, Jin Yao, Mu Ku Chen, Uriel Levy, Din Ping Tsai, Nikolay I Zheludev, Andrei Faraon, Amir Arbabi, Nanfang Yu, et al. Roadmap for optical metasurfaces. *ACS photonics*, 11(3):816–865, 2024.

Zhijie Li, Wenhui Peng, Zelong Yuan, and Jianchun Wang. Fourier neural operator approach to large eddy simulation of three-dimensional turbulence. *Theoretical and Applied Mechanics Letters*, 12 (6):100389, 2022.

Jaideep Pathak, Shashank Subramanian, Peter Harrington, Sanjeev Raja, Ashesh Chattopadhyay, Morteza Mardani, Thorsten Kurth, David Hall, Zongyi Li, Kamyar Azizzadenesheli, et al. Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators. *arXiv preprint arXiv:2202.11214*, 2022.

Rahul Trivedi, Logan Su, Jesse Lu, Martin F Schubert, and Jelena Vuckovic. Data-driven acceleration of photonic simulations. *Scientific reports*, 9(1):19728, 2019.

Sensong An, Bowen Zheng, Mikhail Y Shalaginov, Hong Tang, Hang Li, Li Zhou, Jun Ding, Anuradha Murthy Agarwal, Clara Rivero-Baleine, Myungkoo Kang, et al. Deep learning modeling approach for metasurfaces with high degrees of freedom. *Optics Express*, 28(21):31932–31942, 2020.

Jiaqi Jiang, Mingkun Chen, and Jonathan A Fan. Deep neural networks for the evaluation and design of photonic devices. *Nature Reviews Materials*, 6(8):679–700, 2021.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pages 234–241. Springer, 2015.

Prafulla Dhariwal and Alexander Nichol. Diffusion models beat gans on image synthesis. *Advances in neural information processing systems*, 34:8780–8794, 2021.

Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707, 2019.

Joowon Lim and Demetri Psaltis. Maxwellnet: Physics-driven deep neural network training based on maxwell's equations. *Apl Photonics*, 7(1), 2022.

Mingkun Chen, Robert Lupoiu, Chenkai Mao, Der-Han Huang, Jiaqi Jiang, Philippe Lalanne, and Jonathan A Fan. High speed simulation and freeform optimization of nanophotonic devices with physics-augmented deep learning. *ACS Photonics*, 9(9):3110–3123, 2022.

Kamyar Azizzadenesheli, Nikola Kovachki, Zongyi Li, Miguel Liu-Schiaffini, Jean Kossaifi, and Anima Anandkumar. Neural operators for accelerating scientific simulations and design. *Nature Reviews Physics*, pages 1–9, 2024.

Zongyi Li, Daniel Zhengyu Huang, Burigede Liu, and Anima Anandkumar. Fourier neural operator with learned deformations for pdes on general geometries. *Journal of Machine Learning Research*, 24(388):1–26, 2023.

Chenkai Mao, Robert Lupoiu, Tianxiang Dai, Mingkun Chen, and Jonathan A Fan. Towards general neural surrogate solvers with specialized neural accelerators. *arXiv preprint arXiv:2405.02351*, 2024.

Enze Xie, Wenhai Wang, Zhiding Yu, Anima Anandkumar, Jose M Alvarez, and Ping Luo. Segformer: Simple and efficient design for semantic segmentation with transformers. *Advances in neural information processing systems*, 34:12077–12090, 2021.

Joon-Suh Park, Shuyan Zhang, Alan She, Wei Ting Chen, Peng Lin, Kerolos MA Yousef, Ji-Xin Cheng, and Federico Capasso. All-glass, large metalens at visible wavelength using deep-ultraviolet projection lithography. *Nano letters*, 19(12):8673–8682, 2019.

Alexander Dorodnyy, Jasmin Smajic, and Juerg Leuthold. Mie scattering for photonic devices. *Laser & Photonics Reviews*, 17(9):2300055, 2023.

Manuel Nieto-Vesperinas. Fundamentals of mie scattering. In *Dielectric Metamaterials*, pages 39–72. Elsevier, 2020.

Elyas Bayati, Alan Zhan, Shane Colburn, Maksym Viktorovich Zhelyeznyakov, and Arka Majumdar. Role of refractive index in metalens performance. *Applied optics*, 58(6):1460–1466, 2019.

Jean Kossaifi, Nikola Kovachki, Kamyar Azizzadenesheli, and Anima Anandkumar. Multi-grid tensorized fourier neural operator for high-resolution pdes. *arXiv preprint arXiv:2310.00120*, 2023.

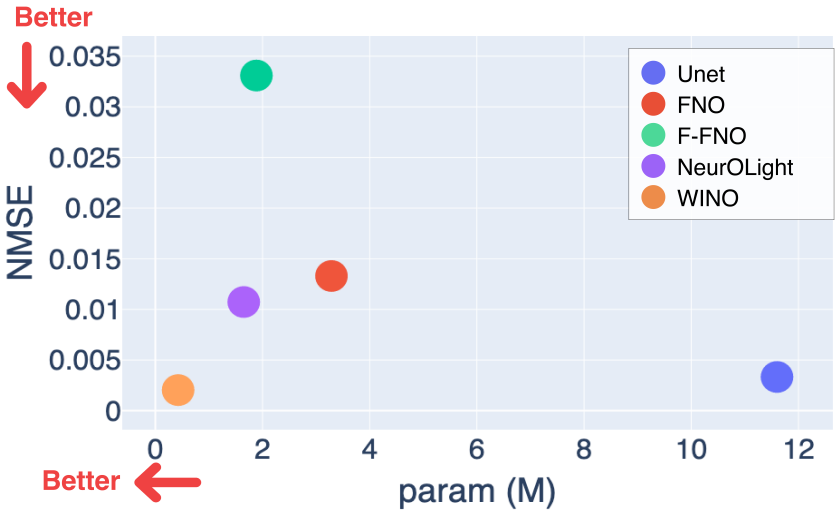

Figure 3: Comparison of WINO with various surrogate solvers in terms of the number of parameters and prediction performance for untrained wavelengths.

A Related Work

A.1 Neural Network-Based Approximations

Light is an electromagnetic wave, and its behavior is governed by PDEs known as Maxwell's equations. Maxwell's equations describe the electromagnetic behavior of light such as the mutual induction of electric and magnetic fields, the propagation of light waves, and the interaction of light with matter. Previous efforts in electromagnetic field prediction have focused on training neural networks to approximate field values. Data-driven neural network research Jiang et al. [2021], Trivedi et al. [2019], An et al. [2020] attempted to learn and predict Maxwell's equations using simulated field data. However, these approaches often struggled to accurately capture the physical characteristics of Maxwell's equations. To address this issue, Physics-Informed Neural Networks (PINNs) were employed in Physics-Augmented Deep Learning research Chen et al. [2022]. This method combined data loss and Maxwell loss as the total loss, integrating Maxwell's equations into the field prediction process for given structures. While this approach improved prediction speed compared to traditional FDFD solvers, it still had significant limitations. Specifically, it required datasets for each individual wavelength and could only predict fields for the trained wavelengths. Additionally, the field predictions were confined to small and simple photonic structures.

A.2 Advancements in Operator Learning

Previous research using data-driven approaches has struggled to accurately learn and infer the characteristics of Maxwell's equations. As a result, researchers have begun exploring neural network-based operators designed to model and predict complex physical systems Azizzadenesheli et al. [2024], Li et al. [2020]. Recent advancements in the field of Operator Learning offer promising solutions that effectively learn PDE characteristics while preserving discretization invariance. The FNO proposed by Li et al. [2020] presents a deep learning architecture designed to learn mappings between infinite-dimensional function spaces. This approach significantly enhances the efficiency and accuracy of solving PDEs. The FNO achieves resolution invariance and supports zero-shot super-resolution, establishing it as a pioneering method in operator learning. After FNO has been published, the Factorized-FNO Tran et al. [2021], named F-FNO, improves the performance of the original FNO by using separable spectral layers and enhanced residual connections, significantly reducing errors in various PDE problems. Additionally, Multi-Grid Tensorized Fourier Neural Operator (MG-TFNO) Kossaifi et al. [2023]. MG-TFNO introduces tensor decomposition techniques that significantly reduce the memory footprint while maintaining model expressivity. Specifically,

tensorization might help optimize our model's performance when predicting the electric field over a broader range of wavelengths, without compromising the computational efficiency.

In electromagnetic field prediction, the state-of-the-art model NeurOLight Gu et al. [2022] leverages the FNO. Gu et al. [2022] encodes wave prior by combining wavelength information and material permittivity, effectively learning the PDEs of Maxwell's equations. The model has demonstrated superior parameter efficiency and higher accuracy compared to previous studies. However, it still has limitations as it predicts fields for a limited narrow range of wavelengths and grid steps, making it less suitable for broadband wavelength applications.

B Problem Setting

Variable	Meaning	Dimensionality
$R_h^{(l)}$	Horizontal component weights in Fourier domain	$M_h C^2/4G^2$
$R_v^{(l)}$	Vertical component weights in Fourier domain	$M_v C^2/4G^2$
$K^{(l)}$	Kernel integral operator in the original FNO	$C^2 M_h M_v$
$z^{(l)}$	Input chunk at layer l	Infinite
L	Number of WINO layers	1
$\mathcal{F}$	Fourier transform	Infinite
$\mathcal{F}^{-1}$	Inverse Fourier transform	Infinite
σ	Non-linear activation function	Infinite
b	Bias vector	
$W_{ch}^{(l)}$	Channel shuffling weights	C^2
G	Number of groups in Fourier space	Between 1 and $\frac{1}{2}\min\{s_1, \cdots, s_d\}$
λ	Wavelength of light	1
$\boldsymbol{\epsilon}_r$	Relative permittivity	Infinite
$\mathcal{W}_x, \mathcal{W}_z$	Wave prior in the x and z directions	Infinite
$\Delta l_x, \Delta l_z$	Simulation step length in x and z directions	1
a	Input function	Infinite
u	Output function	Infinite
d_a, d_u	Dimension of input/output co-domain	1

Table 2: Table of notation

Let $\Omega \subset \mathbb{R}^d$, $\mathcal{A} = \mathcal{A}(\Omega; \varepsilon^{d_a})$ where $\varepsilon = \{\epsilon_{\text{air}}, \epsilon_{\text{material}}\}$, and $\mathcal{U} = \mathcal{U}(\Omega; \mathbb{C}^{d_u})$ be a bounded open set of the underlying domain, infinite-dimensional spaces of the relative permittivity and field of the simulation. ϵ_{air} and $\epsilon_{\text{material}}$ represent the relative permittivity values of air and material. Additionally, we assume that $\mathcal{W} \subset \mathbb{R}_{\geq 0}$ is the broadband range, such as the visible light spectrum, and $\tilde{\mathcal{W}} \subset \mathcal{W}$ is a discrete collection of wavelengths evenly distributed at a specific interval from $\mathcal{W}$. Our model G_θ learns an ideal electromagnetic simulator $G^\dagger : \mathcal{A} \to \mathcal{U}$ for continuous wavelengths $\mathcal{W}$ by mapping between infinite-dimensional function spaces using a finite set of Maxwell PDE input-output pairs $\{(w_j, a_j), u_j\}_{j=1}^N$, where $w \in \tilde{\mathcal{W}}$. Similar to a prior work Kovachki et al. [2023], WINO has an iterative process to map between infinite-dimensional function spaces.

$$u = \mathcal{G}_\theta(a) = (\mathcal{Q} \circ \mathcal{L}^{(L)} \circ \cdots \circ \mathcal{L}^{(1)} \circ \mathcal{P})(a)$$

where $\mathcal{L}^{(l)}$ is a l-th operator layer, L indicates the number of layers, and $\mathcal{P}$, $\mathcal{Q}$ are lifting and projection operators, respectively.

C Experiment Setup

C.1 Comprehending Optical Field Simulation

To collect training and validation datasets, we employ the FDFD method which is one of the numerical methods used to solve Maxwell's equations. The FDFD method discretizes the spatial domain into coordinates and solves electromagnetic fields in the frequency domain, making it particularly effective for steady-state analysis. This method is especially effective for generating field data for single frequencies as light passes through the metalens. Resonances can occur in subwavelength photonic

structures and significantly affect their performance. However, predicting the resonance of electric fields in subwavelength structures becomes challenging when the wavelength changes due to the nonlinearity of resonance Dorodnyy et al. [2023], Nieto-Vesperinas [2020]. Therefore, surrogate solvers trained on specific wavelengths struggle to predict fields for different wavelengths, even when applied to the same structure. For a detailed explanation of the nonlinearity of resonance, refer to the Appendix section titled “Nonlinearity of Subwavelength Photonic Structure Resonance."

C.2 FDFD Simulation Setup

In our study, we utilize Ceviche Hughes et al. [2019], an open-source FDFD simulation tool, to generate electromagnetic field data for photonic structures. Specifically, we analyze the field of a metalens structure, simulating only half of the cylindrical symmetry lens to reduce computation time. We implement the FDFD simulation with a grid-based approach at a spatial resolution of 40 points per μm. The simulation domain measures 5.2 $\mu m \times$ 6.85 μm, excluding a 1-μm thick perfectly matched layer (PML) at the boundaries. We use SiO_2, a material commonly employed in the design of metalenses Park et al. [2019]. As demonstrated in Figure 4 (right), we randomly place either SiO_2 or air within the minimum design grid size, a width of 75nm, in the design area. This simulation setup generates data for the E and H fields when visible light passes through the structure.

C.3 Datasets

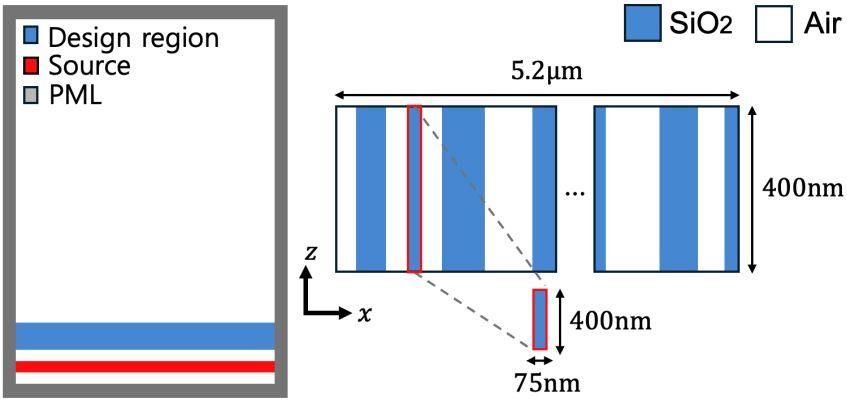

Figure 4: FDFD simulation setup overview. (left) The entire simulation space includes the metalens structure. The blue region represents the design region of the lens, red line represents the visible light source, and the gray region represents the PML used to minimize reflections. (right) The design region of the metalens, shown as the blue area in (left).

We aim to predict the E-field for all wavelengths from discrete data in the 400-700nm (visible light) range. To achieve this, we generate field data for discrete wavelengths as our dataset. Different lens structures are randomly placed in the design region for each data, and the field data resulting from FDFD simulations, along with the permittivity(ε) of the materials, wavelength(λ), and structural information, are used as pairs for training. Specifically, the training set consists of 12,000 data sampled randomly at 20nm intervals within the 400-700nm wavelength range. As mentioned above, our goal is to achieve seamless field prediction across the broadband wavelength. Therefore, to use unseen data for testing, we use 6,020 data each for the test and validation sets, composed of 20 data for each wavelength, sampled at 1nm intervals between 400 and 700nm.

C.4 Training Details

Our model training runs for 200 epochs with 32 batch size. The best model is selected based on the lowest loss in predicting previously unseen wavelengths during the training phase. We employ the GELU (Gaussian Error Linear Unit) activation function and the AdamW optimizer with the following parameters: a learning rate of 0.002, beta values of (0.9, 0.000), epsilon set to 10^{-8}, and a weight decay factor of 0.0001. To dynamically adjust the learning rate, we utilize a cosine annealing learning rate scheduler with a minimum learning rate of 0.00001. Mode is set to (50, 60). We leverage

PyTorch, and the training process is conducted using a single RTX 6000 Ada and an AMD EPYC 7763 CPU.

C.5 Training Objective and Evaluation Metric

Fields typically exhibit different statistics despite the fixed source power. We employ the normalized mean squared error (N-MSE) objective, $L(G_\theta(a), G^\dagger(a)) = (||G_\theta(a) - G^\dagger(a)||_2^2 / ||G^\dagger(a)||_2^2)$, to distribute the optimization effort evenly across several field data. We also use N-MSE for evaluation.

D Ablation Study

D.1 Hyperparameters of FGCS Operator

Weight sharing	Channel shuffling	Groups	Channels	Params	Test loss	
					Untrained	Trained
		1	32	0.423	0.00413	0.00361
		2	48	0.562	0.00278	0.00193
		4	64	0.654	0.00251	0.00165
✓		2	48	0.372	0.00398	0.00342
✓		4	64	0.401	0.00234	0.00199
	✓	2	48	0.576	0.0033	0.00276
	✓	4	64	0.679	0.00211	**0.00153**
✓	✓	2	48	0.386	0.00366	0.00302
✓	✓	4	64	0.426	**0.00204**	0.00177
		8	128	1.898	0.00162	0.00092
✓		8	128	1.306	0.00123	0.00091
	✓	8	128	1.996	0.00145	**0.00081**
✓	✓	8	128	1.405	**0.00105**	**0.00081**

Table 3: Ablation of design components of proposed FGCS operator. The ✓indicates the use of the design components.

We assess the validity of the hyperparameters (groups and channels) and design choices (weight sharing and channel shuffling) of the proposed FGCS operator by conducting an extensive ablation study, as shown in Table 3. We increase the number of groups in proportion to the growth of channels to consider a consistent number of parameters in all cases. Row 9 presents the FGCS operator used in our model. Rows 1-3, which do not utilize weight sharing and channel shuffling in the Fourier domain, follow the same operational form as described in a previous study Kim et al. [2024]. When only weight sharing is employed (rows 4-5), the prediction of fields for both unobserved and observed wavelengths during the training phase demonstrates worse performance compared to that in the previous study Kim et al. [2024] due to a substantially small number of parameters. However, as the number of groups and channels increases, the performance becomes competitive. Notably, the performance in predicting the fields for unobserved wavelengths during training is better. Thus, the weight-sharing scheme significantly aids in generalization for the unobserved spectral parameters. In addition, comparing the previous method Kim et al. [2024] (rows 2-3) with the case where only channel shuffling is utilized (rows 6-7), the field prediction performances for both the untrained and trained wavelengths are better. This suggests that employing the channel shuffling method with the grouping method significantly improves the overall performance. Following the design configuration of the FGCS operator (rows 8-9), we can construct a parameter-efficient model that achieves an overall performance improvement as the number of groups and channels is heightened. While the FGCS operator shows slightly lower prediction performance for trained wavelengths than when only a channel shuffling scheme is employed, it demonstrates superior performance for untrained wavelengths. Therefore, the proposed FGCS operator contributes to improving interpolation performance for unseen broadband spectrum parameters during training. Rows 10-13 present varying performances when increasing the model width. The performance improvement of the FGCS operator becomes increasingly significant with the addition of more channels. In particular, there is a substantial improvement when the number of groups and channels increases to 8 and 128, respectively. We believe that excessively sparse constructions of weights tensor and information compensating with channel shuffling may improve the generalization capacity of large models.

D.2 Conditioning Methods

We conduct a comparative analysis of our proposed method and two other conditioning methods to evaluate our conditioning method's efficacy.

Conditioning method	Params	Test loss	
		Untrained	Trained
Spectral parameter conditioning	1.086	0.13761	0.002577
Concatenating wave prior	0.326	0.002821	0.002022
Our conditioning method	0.426	**0.002035**	**0.001774**

Table 4: Ablation study of several conditioning methods.

Table 4 shows the results of the ablation study of several conditioning methods employed on WINO. As in the previous experiments, The terms "untrained" and "trained" represent the setups in which fields are predicted using untrained and trained wavelength data, respectively. The conditioning methods used for comparison are spectral parameter conditioning Gupta and Brandstetter [2022], Mao et al. [2024] and the concatenating wave prior Gu et al. [2022]. The spectral parameter conditioning method transforms the scalar wavelength into a vector via sinusoidal embedding Vaswani et al. [2017]. It then employs a two-layer feed-forward network to project the embeddings onto a higher-dimensional space of $4 \times$ hidden channels. The embeddings are mapped to the Fourier space using the module named FreqLinear module described in Gupta and Brandstetter [2022]. Finally, the embeddings in the Fourier space are mode-wise multiplied with the input of the Fourier integral kernel to perform parameter conditioning in the spectral domain. The concatenating wave prior injects the condition information by simply concatenating the given refined wave prior to the input data. Although the spectral parameter conditioning method effectively predicts the fields of wavelengths observed during training, it fails to perform interpolation in the parameter space of the broadband spectrum. Thus, the spectral parameter conditioning method is incapable of learning the patterns of fields with continuously varying wavelengths. Our conditioning method thoroughly outperforms the concatenating wave prior conditioning method in both the trained and untrained wavelength cases, achieving a 12.3% improvement in trained wavelengths and a 27.9% improvement in untrained wavelengths. Consequently, the proposed method is effective for injecting conditions that are substantially similar to the fields in terms of the spatial and structural features. Furthermore, we conducted an additional comparative analysis between our conditioning method and other methods using various models. For a detailed explanation of the more ablation studies, refer to the Appendix section titled "Additional Ablation Studies for Conditioning Methods."

As the influence of simulation parameters like higher relative permittivity increases, the nonlinear optics effects become more pronounced, and the precluded assumption weakens. However, we found that our method still works reasonably well even in these cases. For a detailed experiments, refer to the Appendix section titled "Additional Experiments for the Weakened Precluded Assumption."

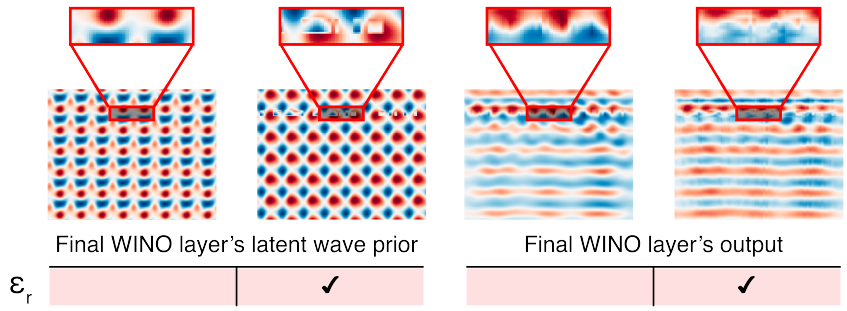

Figure 5: Visualization of the features of condition information and outputs from the final WINO layers for a specific part of the entire region for 400nm wavelength data. (left) The latent wave prior in the final WINO layer, excluding and including the ϵ_r term. (right) The output of the final WINO layer without and with considering the ϵ_r term. The magnified sections are part of the region where structures exist.

D.3 Permittivity in the Wave Prior

ϵ_r term	Test loss	
	Untrained	Trained
✓	0.002372	0.002148
	0.002035	**0.001774**

Table 5: Ablation study for ϵ_r term in wave prior.

As shown in Table 5, we demonstrate performance enhancement resulting from the exclusion of the ϵ_r term from the wave prior through an ablation study. To identify the reason for the improvement, we investigate the latent wave prior and the output of the last WINO layer. Figure 5 (left) shows the latent wave prior before element-wise multiplication in the last WINO layer. Sharp features emerge in regions with structures when the wave prior reflects the ϵ_r term because our conditioning method utilizes shallow 1×1 convolutional layers to ensure the discretization invariance property and maintain the spatial structure of the refined wave prior. Figures 5 (right) show the outputs of the final WINO layer. The result in Figure 5 (right), where the ϵ_r term is included in the wave prior, has bizarre features present throughout the entire region compared to when the ϵ_r term is not included. Therefore, we conclude that the ϵ_r term in the wave prior introduces unintended bias.

E Detailed Model Architectures

E.1 Unet

We use a 4-level modern convolutional Unet architecture with an initial channel size of 16. Each level consists of two residual blocks with GELU activation function and Group Normalization. We concatenate the refined wave prior with the input for the conditioning method.

E.2 FNO

We set the number of Fourier layers and the number of channels in each Fourier layer to 5 and 32, respectively. The frequency modes for the z and x axes are set to [32, 10]. In addition, the lifting operator is a linear operator, and the projection operator is a 2-layer feed-forward network with GELU activation and a dimension of 128. We concatenate the refined wave prior with the input to inject condition information.

E.3 F-FNO

We employ 12 Fourier layers, each consisting of 64 channels. The frequency modes for the z and x axes are set to [50, 60]. The lifting operator is a linear operator, and the final projection operator is a 2-layer feed-forward network with GELU activation function and a dimension of 256. To incorporate condition information, we concatenate the refined wave prior with the input.

E.4 NeurOLight

We use 12 Fourier layers with 64 channels each. The frequency modes for the z and x axes are set to [50, 60]. The channel expansion factor for the convolutional modules in the NeurOLight layers is configured to be 2. To inject condition information, we concatenate the wave prior with the input. During training, we implement stochastic network depth with a rate of 0.1 to the residual NeurOLight layers to alleviate overfitting.

E.5 WINO

We use 12 WINO layers with 64 channels each and specify the number of groups as 4. The frequency modes for the z and x axes are configured as [50, 60]. We implement a channel expansion scheme with a factor of 2 for the 2-layer feed-forward network in the WINO layers, inspired by Xie et al. [2021], Gu et al. [2022]. The lifting operator is linear, and the final projection operator is a 2-layer feed-forward network with GELU activation and a dimension of 256.

Model	Param (M)	Test loss		Design region test loss	
		Untrained	Trained	Untrained	Trained
Unet-Bottleneck	11.6371	0.11786	0.03303	0.059753	0.05574
Unet-AdaGN	11.779	0.09342	0.03548	0.047393	0.034908
Unet	11.5999	0.003332	0.001732	0.00619	0.003775
FNO	3.2868	0.0133	0.007719	0.028339	0.025067
FNO-Ours	3.2974	0.009966	0.008369	0.027739	0.028831
F-FNO	1.887	0.03309	0.02204	0.043035	0.037091
F-FNO-Ours	1.92	0.007606	0.006705	0.014702	0.014724

Table 6: Comparisons of field prediction and interpolation with various conditioning methods and models. Unet, FNO, and F-FNO refer to concatenating the wave prior with the input data. “Ours" represents our conditioning method.

Model	Param (M)	Test loss		Design region test loss	
		Untrained	Trained	Untrained	Trained
Unet	11.5999	0.017165	0.017799	0.027481	0.02874
FNO	3.2868	0.120722	0.120426	0.333563	0.33381
FNO-Ours	3.2974	0.120246	0.122924	0.351337	0.348711
F-FNO	1.887	0.100605	0.106164	0.192996	0.192103
F-FNO-Ours	1.92	0.067897	0.067897	0.140872	0.140648
NeurOLight	1.65	0.017074	0.011558	0.033887	0.026682
WINO	**0.426**	**0.007932**	**0.008131**	**0.018456**	**0.017944**

Table 7: Comparisons of field prediction and interpolation with high relative permittivity.

F Additional Ablation Studies for Conditioning Methods

To verify the effectiveness of our conditioning method, we conduct an additional ablation study by applying the proposed conditioning method to other models and comparing it with other conditioning methods. “Unet-Bottleneck" maps the scalar wavelength value into a vector using a simple three-layers feed-forward network with GELU and then adds the induced condition vector to the input of the bottleneck in Unet. “Unet-AdaGN" reflects the wavelength condition in the model using Adaptive Group Normalization (AdaGN) Dhariwal and Nichol [2021]. The given wavelength value is transformed using sinusoidal embedding Vaswani et al. [2017]. Subsequently, a two-layer feed-forward network is used to project the embeddings into a higher-dimensional space of $4 \times$ hidden channels. The condition embeddings are incorporated into the Unet through AdaGN. “Unet, FNO, and F-FNO" refers to the method of concatenating the wave prior with the input data. Finally, “Ours" refers to our proposed conditioning method. According to Table 6, our method significantly improves the interpolation performance for unobserved wavelengths compared with the various conditioning methods of other models. When our conditioning method is applied to the FNO model, the field prediction performance for the wavelengths used during training is marginally reduced compared to when the method is not applied. However, there is a substantial performance improvement throughout the entire region for the unobserved wavelength data. In particular, the performance enhancement is dramatic in the case of F-FNO, which has conspicuously fewer parameters than Unet and FNO, factorizes the Fourier transforms over the problem dimensions, and is composed of many layers. Applying our conditioning method to F-FNO leads to a 77% increase in performance for unobserved wavelengths and a 67.6% increase for observed wavelength data across the entire simulation region compared with not using the conditioning method. In addition, in the design region, there is a 65.8% performance improvement for unobserved wavelengths and 60% improvement for observed wavelengths. We highlight that the error in the design region is consistent for both observed and unobserved wavelengths. Thus, we verify that our conditioning method substantially contributes to the interpolation capability.

G Additional Experiments for the Weakened Precluded Assumption

We conducted additional experiments to evaluate the performance of our proposed method when the precluded assumption (that the field and condition data have high similarity) becomes weak. We

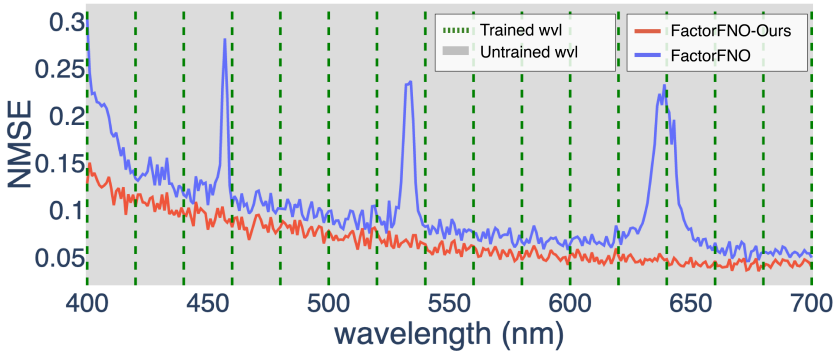

Figure 6: Field interpolation comparisons of F-FNO with simply concatenating conditioning and our conditioning method. Comparison of entire simulation region errors in the 400-700nm wavelength range.

Model	Param (M)	Test loss		Design region test loss	
		Untrained	Trained	Untrained	Trained
WINO	0.426	0.007932	0.008131	0.018456	0.017944
WINO (l=1)	0.417	0.007626	0.007741	0.017677	0.01782
WINO (l=2)	0.409	0.007534	0.007552	0.017484	0.017566
WINO (l=3)	0.401	0.008208	0.008391	0.019047	0.019464

Table 8: Field prediction and interpolation comparisons between injecting the condition into all but the last l WINO layers when relative permittivity is high.

set up the simulation with a more complex field using a material with higher relative permittivity (lithium niobate, refractive index $n \approx 2.3$)[3].

FNO is inefficient in terms of the number of parameters. Despite having fewer layers and widths and lacking additional modules like fully connected layers, it has the highest number of parameters among the neural operators compared. As shown in Table 7[4], FNO struggles to learn complex optical properties even when our conditioning method is applied due to the limited learning capacity caused by the given problems. In contrast, F-FNO has sufficient capacity because of its superior parameter efficiency, which allows for many layers, widths, and additional MLPs. F-FNO demonstrates inferior performance despite its superior learning capacity if concatenating the wave prior with input data for conditioning. In particular, it exhibited peaks in the loss at specific wavelengths (Figure 6). On the other hand, applying our conditioning method to F-FNO significantly improves the field prediction performance and resolves the peak problem. This demonstrates that our method effectively helps to learn complex optical phenomena for a narrow range of wavelengths when sufficient learning capacity is available. While NeurOLight exhibits superior performance in field inference, it struggles to predict E-field across the broadband spectrum. However, WINO still demonstrates the best performance even when the precluded assumption becomes weak.

G.1 Weakening the Precluded Assumption in WINO

WINO element-wisely multiplies latent condition data at every layer. While the conditioning method serves as a powerful regularization means for interpolation on a wide range of conditions, it can have

[3]Permittivity is the square of the refractive index, denoted by n, which indicates how much light is refracted when it passes through a medium. Therefore, materials with higher permittivity have higher refractive indices. Consequently, when light passes through a material with high permittivity, it slows down more than in a material with lower permittivity, resulting in a more complex field pattern.

[4]The inference performance for untrained wavelengths often appears better than for trained wavelengths. This is because nonlinear optics phenomena (diffraction, resonance) caused by the high epsilon occur strongly, resulting in the loss exhibiting a slightly exponential change with respect to the wavelength. Consequently, if spectrum (condition) interpolation is successful, it can result in lower loss values for untrained wavelengths.

adverse effects when the precluded assumption is weakened due to more complex fields simulated from materials with high relative permittivity.

To assess the effects of weakening the assumption in WINO, we conduct an ablation study, injecting information into layers except for the last l layers.

As shown in Table 8, injecting the condition into all but the last 1 and 2 WINO layers ($l = 1, 2$) reflects weakened assumption well and leads to performance improvement. However, injecting the condition into all but the final 3 WINO layers ($l = 3$) results in a decrease in both field inference performance and interpolation performance. This demonstrates the trade-off between the powerful regularization of our conditioning method and reflecting the weakened assumption.

H Nonlinearity of Subwavelength Photonic Structure Resonance

Mie scattering Dorodnyy et al. [2023], Nieto-Vesperinas [2020], a fundamental mechanism describing the interaction of electromagnetic waves with particles smaller than the wavelength of light, is pivotal in understanding resonance phenomena in subwavelength photonic structures. The mathematical formulation of Mie scattering is expressed through an infinite series of spherical harmonics, where each term represents a different mode of scattering:

$$\sigma(\lambda, r) = \frac{2\pi}{k^2} \sum_{n=1}^{\infty} (2n+1) \left(|a_n|^2 + |b_n|^2 \right)$$

where λ is the wavelength, r is the radius of the particle, $k = \frac{2\pi}{\lambda}$ is the wave number, and a_n, b_n are the scattering coefficients for electric and magnetic modes, respectively. The nonlinearity arises from the complex interaction between these modes, which can be highly sensitive to structural and material parameters, making accurate predictions difficult.
Furthermore, high Q-factors are indicative of sharp resonance peaks and narrow bandwidths in the response of photonic structures, which implies that the system has a high selectivity in frequency response. While beneficial for many applications, high Q-factors also mean that the resonance is extremely sensitive to slight deviations in system parameters, such as changes in material properties or geometric alterations. This sensitivity leads to significant challenges in predicting the behavior of the system under slightly altered conditions:

$$Q = \frac{\omega_0}{\Delta\omega}$$

where ω_0 is the resonance frequency and $\Delta\omega$ is the bandwidth of the resonance.

At subwavelength scales, interactions between light and matter involve complex phenomena like electric and magnetic Mie resonances. Therefore, surrogate solvers trained at a single wavelength struggle to predict nonlinear resonances in regions with matter across broadband wavelengths. Our WINO model overcomes this challenge, accurately predicting resonances arising from structures.

We employ a wave prior based on the solution of the wave equation as the condition data ($\mathcal{W}_x = e^{j\frac{2\pi\sqrt{\epsilon_r}}{\lambda} x \mathbf{1}^T \Delta l_x}$ and $\mathcal{W}_z = e^{j\frac{2\pi\sqrt{\epsilon_r}}{\lambda} \mathbf{1} z^T \Delta l_z}$). However, when the relative permittivity, ϵ_r, reflects the materials used in the simulation, physically implausible features appear in the regions where the material is present (Figure 8). To resolve this, we set ϵ_r to 1, assuming the entire materials are air. This approach allows us to obtain physically plausible features in free space.

I Limitations and Future Work

Our weight grouping and sharing method reduces the number of MAC (Multiply–ACcumulate) in the matrix multiplication within the Kernel Integral Operator by a factor of $1/G$ compared to conventional FNOs and decreases the number of parameters by a factor of $1/G^2$. However, WINO requires more computation time (Table 9). This increase in time complexity is attributed to the model's employment of a greater number of layers than the vanilla FNO, as well as the reshaping and concatenation of tensors within each layer in the complex domain to construct the cross-shaped architecture and FCGS operation. On the other hand, despite the additional tensor reshaping processes in the complex domain, WINO still achieves faster inference speeds compared to NeurOLight, which uses a nonlinear Lifting operator and complicated feed-forward networks. Furthermore, our proposed

Model	Runtime (s)	Param (M)
Unet	0.018	11.5999
FNO	0.006	3.2868
FNO-Ours	0.008	3.2974
F-FNO	0.035	1.887
F-FNO-Ours	0.041	1.92
NeurOLight	0.101	1.65
WINO	0.08	0.426

Table 9: Comparisons of the number of parameters and runtime. Runtime is averaged over ten trials with 16 batch sizes.

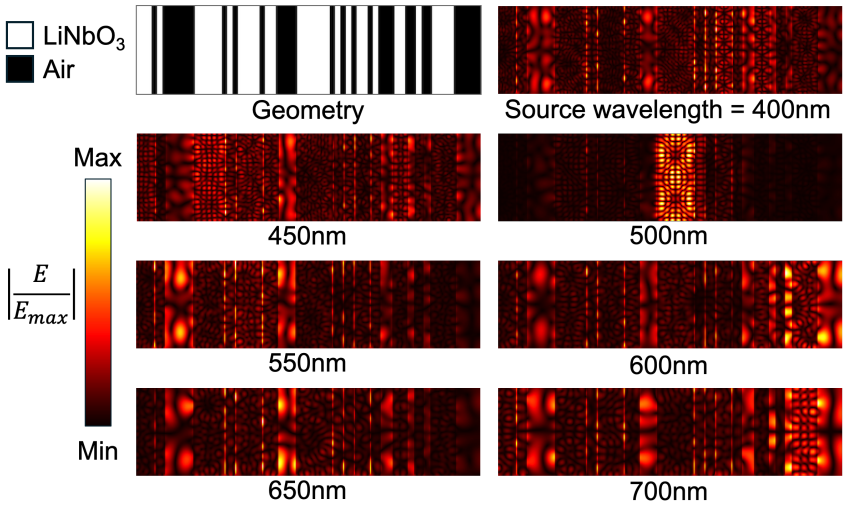

Figure 7: Using the FDTD simulation tool Meep, the E-field at the same time step is visualized for a design region composed of $LiNbO_3$ and Air. This visualization confirms that resonance occurs nonlinearly with respect to wavelength.

conditional method does not significantly add to the computation time, which is one of the critical factors for improving overall field prediction performance.

We leave the task of optimizing the architecture to deal with the time complexity issues for future work. For future work, we plan to extend optical simulations to a centimeter-scale framework and expand the wavelength range from the visible spectrum to the infrared (IR) bands. Additionally, beyond the metalens-based simulations conducted in the present study, we plan to explore other photonic structures, such as mode converters and waveguides.

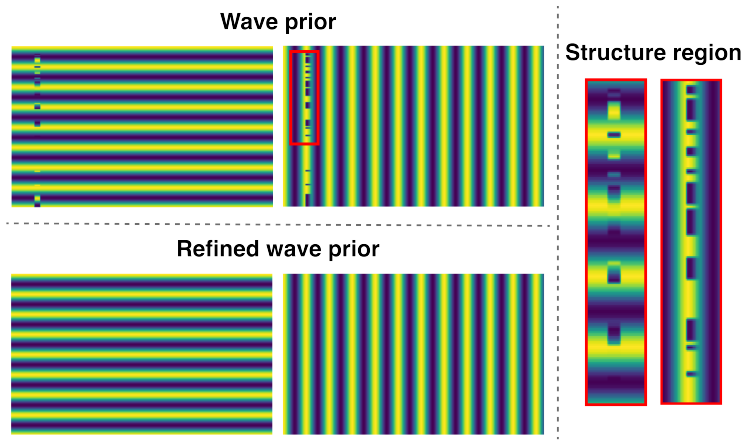

Figure 8: Visualization of (top left) wave prior that includes relative permittivity of materials, (right) physically implausible features of wave prior including relative permittivity of materials, and (bottom left) wave prior that excludes relative permittivity of materials.

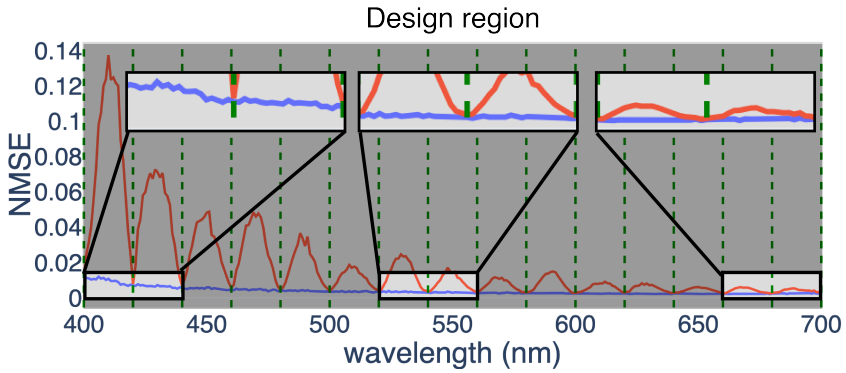

Figure 9: Comparison of design simulation region errors in the 400-700nm wavelength range. The field prediction loss of WINO, particularly in the design region, exhibits a linear change as the wavelengths vary.

D3S3@NeurIPS Paper Checklist (Optional)

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The Abstract and Introduction sections accurately reflect the paper's contributions and scope.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: The Limitation and Future Work section in Appendix discusses the limitations of the proposed work.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory Assumptions and Proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [NA]

Justification: We do not contain any theory.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental Result Reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: The Methodology section in the main paper provides the overall model architecture information. Section C and E in the Appendix demonstrate additional details on the model settings, providing the necessary information for reproducing the main experimental results of this paper.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes] ,

   Justification: We provide our code with detailed instructions as supplementary material. The dataset we used is also in the detailed instructions as the storage link.

   Guidelines:

   - The answer NA means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental Setting/Details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [Yes]

   Justification: Sections C and E of the Appendix demonstrate the experimental setting and details.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment Statistical Significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [No]

   Justification: We did not report error bars because the proposed method demonstrated significantly superior performance on untrained wavelengths, which was the main focus of our study, compared to the given comparison group.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.

- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments Compute Resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: The training details section in the Appendix provides information about the computing resources (GPU and CPU).

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code Of Ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification:

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader Impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer:[NA]

    Justification: Our study utilizes a dataset that strictly involves physical data without any interaction with animals, humans, or any societal components. The work focuses entirely on predicting electric fields across untrained wavelengths using a surrogate solver. Given that the methods developed do not interface with any societal elements and are intended solely for enhancing simulation accuracy and efficiency, there are no foreseeable societal impacts, whether positive or negative. Therefore, this study does not address societal impacts.

Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [NA]

    Justification:

    Guidelines:

    - The answer NA means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: When using existing assets in the code, such as models, we included the license and information about the original owners in that code files.

    Guidelines:

    - The answer NA means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.

- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New Assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes]

    Justification: The markdown files and a license file in the supplementary code materials provide the necessary documentation for the new assets. We documented the MIT license except for copyright holder information to anonymize our assets.

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and Research with Human Subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

    Justification:

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional Review Board (IRB) Approvals or Equivalent for Research with Human Subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]

    Justification:

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.

- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.