

---

# Physics-Constrained Comprehensive Optical Neural Networks

---

**Yanbing Liu<sup>1,\*</sup>**  
ybliu@bupt.edu.cn

**Jianwei Qin<sup>2,\*</sup>**  
qinjw3@sjtu.edu.cn

**Yan Liu<sup>2</sup>**  
liu\_yan@sjtu.edu.cn

**Xi Yue<sup>1</sup>**  
yuexi@bupt.edu.cn

**Xun Liu<sup>3</sup>**  
liuxun\_laby@163.com

**Guoqing Wang<sup>4</sup>**  
gqwang0420@hotmail.com

**Tianyu Li<sup>4</sup>**  
cosmos.yu@hotmail.com

**Fangwei Ye<sup>2,†</sup>**  
fangweiye@sjtu.edu.cn

**Wei Li<sup>3,†</sup>**  
wei\_li\_bj@163.com

<sup>1</sup> Beijing University of Posts and Telecommunications, Beijing, 100876, China

<sup>2</sup> Shanghai Jiao Tong University, Shanghai 200240, China

<sup>3</sup> Beijing Institute of Space Mechanics and Electricity, Beijing, 100094, China

<sup>4</sup> University of Electronic Science and Technology of China, Chengdu, 611731, China

## Abstract

With the advantages of low latency, low power consumption, and high parallelism, optical neural networks (ONN) offer a promising solution for time-sensitive and resource-limited artificial intelligence applications. However, the performance of the ONN model is often diminished by the gap between the ideal simulated system and the actual physical system. To bridge the gap, this work conducts extensive experiments to investigate systematic errors in the optical physical system within the context of image classification tasks. Through our investigation, two quantifiable errors—light source instability and exposure time mismatches—significantly impact the prediction performance of ONN. To address these systematic errors, a physics-constrained ONN learning framework is constructed, including a well-designed loss function to mitigate the effect of light fluctuations, a CCD adjustment strategy to alleviate the effects of exposure time mismatches and a ‘physics-prior-based’ error compensation network to manage other systematic errors, ensuring consistent light intensity across experimental results and simulations. In our experiments, the proposed method achieved a test classification accuracy of 96.5% on the MNIST dataset, a substantial improvement over the 61.6% achieved with the original ONN. For the more challenging QuickDraw16 and Fashion MNIST datasets, experimental accuracy improved from 63.0% to 85.7% and from 56.2% to 77.5%, respectively. Moreover, the comparison results further demonstrate the effectiveness of the proposed physics-constrained ONN learning framework over state-of-the-art ONN approaches. This lays the groundwork for more robust and precise optical computing applications.

## 1 Introduction

In recent years, optical neural networks (ONNs) have garnered significant research attention for inference tasks such as object detection and object classification [1, 2, 3, 4], attributed to their advantages

---

\*Equal contribution

†Corresponding Author

of low energy consumption, high transmission speed, and large information capacity[5, 6, 7]. To establish ONNs, they are typically simulated as Deep Neural Networks(DNNs)[8, 9, 10] on electronic devices and trained using the backpropagation algorithm[11, 12, 13], after which the trained model parameters are deployed to physical ONN systems (as shown in Fig 1(a)). Theoretically, ONNs can maintain prediction performance comparable to that of their simulated electronic counterparts. However, during experimental implementation, errors are inevitably introduced, unexpectedly reducing their prediction accuracy[14, 15].

Some measurable errors in the optical system, such as light field perturbations caused by the scattering of impurity particles in the environment and optical distortions due to lens aberrations, can be mitigated by explicitly modeling them and integrating these error models into electronic training process[16, 17] (as shown in Fig 1(b)), while other unmeasurable errors, such as ambient light effects, laser instability, and crosstalk between light fields, are difficult to correct through physical modeling, posing significant challenges in compensating for discrepancies between simulations and experiments[18, 19, 20].

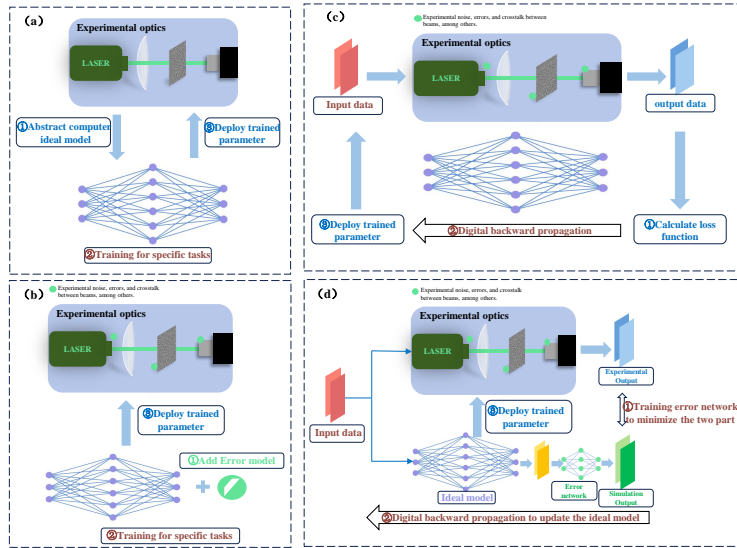


Figure 1: (a) Experimental schematic diagram of the ONN and corresponding simulated DNN on the electronic device. (b) Error compensation method based on the physical error model (c) Hybrid training method with pure digital DNN with error compensation functionality (d) Error compensation method with an ideal physical model and a digital error compensation DNN

The introduction of “hardware-in-the-loop” training techniques[21, 18, 22, 23] has opened up new possibilities for tackling the aforementioned challenges. Unlike in silico training, this hybrid training method incorporates the actual ONN physical response in each update loop, to mitigate the impact of system errors[24], as shown in Fig 1(c). This approach allows the simulated electronic neural network to capture the physical dynamic changes of the input light field along its propagation path, including the propagation process of the light field itself, the influences of various error sources and disturbances encountered. However, the simulated electronic neural network, especially for the primarily linear ONNs, might be too simplistic to learn the complex input-output light field mapping relationship, including light propagation, various errors, and the coupling between light fields.

Therefore, error compensation neural networks are proposed to better simulate the physical optical system [25], as illustrated in Fig 1(d). The addition of the error compensation neural network in the hybrid training scheme is expected to bridge the gap between simulation and physical systems. However, due to the large optimization space and the lack of physical constraints in the simulation system, the training process can be slow and may even diverge.

Inspired by the concept of Physics-Informed Neural Networks (PINNs)[26, 27, 28, 29], which integrate physical information into the network for optimization, we investigate systematic errors in the optical physical system first and then propose a physics-constrained ONN learning framework for image recognition tasks. In our approach, critical physical information is quantitatively integrated

into a 'physics-prior-based' error compensation network[30, 31]. This narrows the search space and reduces the complexity of the required DNNs. In our experiments, the error compensation network converges rapidly and effectively minimizes the disparity between simulation results and actual observations, which significantly enhances the image recognition accuracy of experimental ONNs, leading to state-of-the-art (SOTA) performance on several datasets.

The main contributions of this work are as follows:

- (1) To describe the transmission equations of complex optical systems, we combine quantifiable physical information with machine learning. Given known ideal physical models and key parameters, we use minimal data and lightweight neural networks, leveraging both physics-driven and data-driven approaches for rapid and precise modeling of complex optical systems.
- (2) We informed the network of two significant quantitative errors—laser source instability and camera exposure mismatch—as physical prior information, greatly enhancing convergence efficiency.
- (3) Under multiple physical constraints, our network can focus more efficiently on learning other unmeasurable system errors beyond the two aforementioned quantitative errors, such as crosstalk between beams, device imperfections, and alignment errors. Consequently, it achieves state-of-the-art results across multiple datasets.

## 2 Related Work

**Training base on the ideal optical 4f system** The 4f-ONN is constructed with two Fourier lenses and a Spatial light modulator (SLM) in the focal plane, modulating the frequency spectrum of the light field, which allows for the automatic realization of optical convolution operations and makes it well-suited for optical neural networks[32, 33, 3, 34]. Taking the optical 4f system as an example, in the absence of errors, its ideal transformation process adheres to the Fresnel diffraction[35, 36] integral under ideal conditions[15]. The loss between the simulated output optical field image and the target optical field is backpropagated through gradient descent, affecting the frequency domain phase distribution of the 4f system. Through multiple training iterations on the electronic device, an optimized spectral phase distribution is obtained. This phase distribution is then loaded onto the focal plane of the 4f system via the SLM, enabling image classification tasks.

**Fitting errors based on the physical model.** [17, 16] aim to model the primary disturbances and errors in optical systems, integrating the disturbance model appropriately into the ideal simulated physical transformation model as an accurate description of the optical system. While this approach enhances the robustness of the model against certain interferences, it only considers one or a few disturbances existing in the physical system, thus unable to account for all disturbances, unmeasurable quantities, and their coupling and crosstalk within the model. Therefore, its description of the physical system model is not yet precise enough.

**ONNs auto-learning** [24, 18] aims to describe real optical systems through autonomous learning using ONNs. By feeding input and output signals obtained from experiments into the ONNs, the ONNs learn the functionality of the optical system in a data-driven manner. Ultimately, this enables the network to accurately reflect the functionality of the actual optical system. However, learning the input-output relationships in experiments might be overly complex for optical neural networks, particularly for simple linear optical neural networks.

**Error compensation network** [37, 25] integrates ideal physical models with deep neural networks (DNNs) in optical system modeling to compensate system errors and accelerate convergence speed. However, the lack of physical parameters and physics-constrained loss functions requires the network to learn this information from more data during training. Consequently, it necessitates the design of more complex DNN architectures and the acquisition of additional experimental data to support the training process. In contrast, through the integration of physical information into the simulated model, our method is able to achieve high prediction performance using a relatively simple error compensation network.

### 3 Method

#### 3.1 Introducing quantifiable physical information into optical systems

Introducing quantifiable physical information into optical systems is a key strategy for enhancing the accuracy and efficiency of simulations in image recognition applications. The relationship between the input and output signals of the optical system in experiments can be represented by a simplified model, as depicted in Fig 2(a):

$$g(u) = f(u) + \Delta f_{dev}(u) + \Delta f_{jit}(u) + \dots + \eta(u) \quad (1)$$

where  $f(u)$  denotes the ideal transformation process,  $\Delta f_{dev}$  indicates deviations caused by imperfections in optical devices,  $\Delta f_{jit}$  represents deviations due to laser jitter, and  $\eta$  encompasses deviations caused by unmeasurable quantities in the system.

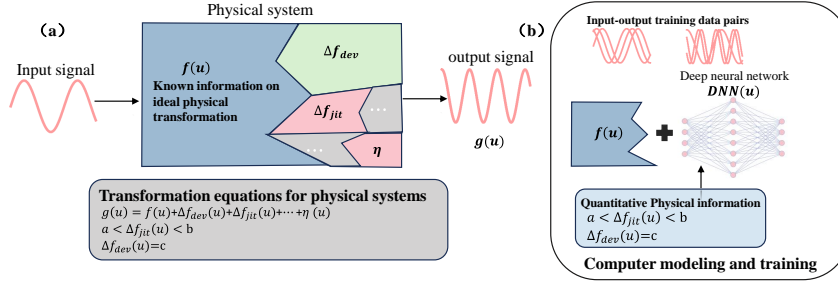


Figure 2: Schematic diagram illustrating the integration of physical information to reduce experimental errors in the precise simulation of physical systems.

Accurate simulation of the actual optical signal transmission process necessitates not only simulating the ideal transformation of the light path but also compensating for system errors, as illustrated in Fig 2(b). Here,  $f(u)$  symbolizes the ideal transformation used in the simulation, and DNN is employed as the neural network for error compensation, with  $\Delta f_{dev}$  and  $\Delta f_{jit}$  representing quantifiable system errors. Methods such as the Finite Difference Time Domain[38, 39, 40] (FDTD) for solving the Helmholtz equation or the paraxial Schrödinger equation using the split-step Fourier transform method[41, 42, 43] are utilized to effectively simulate the ideal transformation  $f(u)$  of the input light field. The primary role of the DNN is to learn from the difference between the output of  $f(u)$  and the experimental output, focusing on experimental errors beyond  $f(u)$  and their coupling effects—a complex task. To reduce the complexity of the DNN and ensure the convergence of our model, quantifiable physical information, such as the range of laser jitter and the grayscale value range of images received by cameras, is incorporated into the DNN. Through carefully designed loss functions[44, 45, 46] and by adjusting the overall data bias, this quantifiable physical information is integrated into the DNN model, thereby reducing its complexity and significantly increasing convergence speed.

#### 3.2 Architecture of optical neural networks and measurement of quantifiable experimental errors

We employ the method described above to precisely simulate the light propagation process within an image-classification optical neural network, where the experimental errors contribute to a reduction in classification accuracy. The experimental architecture is illustrated in Fig 3(a), we established an experimental framework for an optical neural network with error compensation capabilities. Detailed descriptions of the experimental setup can be found in the supplementary materials. Using the Fresnel diffraction integral, we can derive the ideal transformation process within this optical system (i.e., the function  $f$  in Equation.1). This setup allows for precise control and manipulation of the light field, facilitating detailed investigations into the dynamics of image processing within optical neural networks.

$$E_{out} = \frac{\psi_{sph}}{i\lambda f_1 f_2} \cdot [t(x, y) * e^{i\phi(x', y')}] \quad (2)$$

$$f = \min\{|E_{out}|^2, I_{max}\} \quad (3)$$

where  $*$  denotes the convolution operation,  $f_1, f_2$  are focal length of L1 and L2, respectively.  $t(x, y)$  is the input image in DMD, loading in the form of a binary transmission function.  $\phi(x', y')$  represents for the phase distribution on SLM and  $\psi_{sph}$  is a spherical wave phase factor. Due to the inherent limitations of CCD which can only detect light intensity, and given that the maximum intensity that the CCD can read is  $I_{max}$ , the output light field must transform as dictated by Equation.3. After transforming the optical field of the input image using the optical 4f system and SLM, the output optical field is obtained. Taking the MNIST dataset as an example, the output optical field is divided into 10 equally sized regions for recognition and classification. The light intensity  $I_i (i = 1, 2, \dots, 10)$  in each region is measured to form the intensity sequence  $[I_1, I_2, \dots, I_{10}]$ . The region with the highest light intensity corresponds to the classification result.

In this optical system, beyond the ideal transformation process  $f$ , some quantifiable physical quantities can be utilized to simplify the complexity of the error compensation network. Two primary quantifiable experimental errors include the instability of laser intensity and the consistency of light intensity between the output image and the simulated image. To quantitatively analyze and mitigate these errors, we measured the instability of laser intensity and the impact of camera exposure time on the experimental output image.

### 3.2.1 Quantitative compensation of laser intensity instability

For the instability of laser intensity  $\Delta f_{jit}$ , we used a CCD to measure the overall grayscale values within a fixed area to assess the stability of the laser. Over 700 minutes, the variation of pixel values changed randomly, as depicted in Fig 3(b). To incorporate errors caused by intensity instability into our error compensation network, A well-designed loss function is used to increase the gap between the maximum and second maximum intensity values within the classification area. The designed loss function is shown below:

$$Loss = ReLu\{W_{Gap} - |I_{max} - I_{2ndmax}|\} + \sum_{i=1}^N y_i \log(\hat{y}_i) \quad (4)$$

Here, the  $W_{Gap}$  denotes the light intensity gap,  $N$  denotes the total number of classes,  $y$  represents a one-hot encoded vector that indicates the true class labels, with  $y_i$  being the  $i$ -th element of the vector  $y$ . The term  $\hat{y}$  corresponds to the network's output probabilities, which are typically derived through the application of a softmax function, with  $\hat{y}_i$  representing the probability that the model assigns to the likelihood that the sample pertains to class  $i$ . In the experiment, the gap is configured to substantially exceed the range of instability variations, thereby mitigating the reduction in experimental accuracy attributable to laser instability. However, the value of  $W_{Gap}$  cannot be set too high, as this would reduce the network's fitting ability, leading to decreased classification accuracy. Here, we have experimentally measured the impact of different  $W_{Gap}$  values on experimental accuracy (as shown in the inset of Fig 3(b)). When  $W_{Gap}$  is set to 10, the experimental accuracy is optimal. In this setting, the network still maintains good fitting ability, and the instability of laser intensity is also compensated. This approach ensures that despite fluctuations in laser output, the performance of the optical neural network remains robust, providing reliable and precise experimental outcomes.

### 3.2.2 Quantitative compensation of exposure time mismatches

For the noise caused by optical devices  $\Delta f_{dev}$ , experimentally, the primary and measurable error source comes from the intensity of light received by the CCD. The distribution of pixel values read by the CCD corresponds to the distribution of light intensity it receives. However, inherent biases in the CCD itself result in discrepancies between simulated and experimental pixel values, leading to reduced experimental accuracy. The CCD's pixel values depend on the exposure time and the intensity of the emitted laser light. To eliminate this error, we fixed the output laser intensity and adjusted the exposure time of CCD to measure the gap between the simulated and experimental CCD pixel values. For each exposure time, we loaded 1000 images from the MNIST dataset onto the DMD and calculated the difference in pixel values across 10 classification regions, as shown in Fig 3(c). In the ideal case, the error is completely eliminated ( $\delta I = 0$ , as indicated by the red bars). In actual experiments, we first coarsely adjust the exposure time to bring the average  $\delta I$  close to zero (green and blue bars in Fig 3(c)), and then finely adjust the exposure time to minimize the variance of  $\delta I$

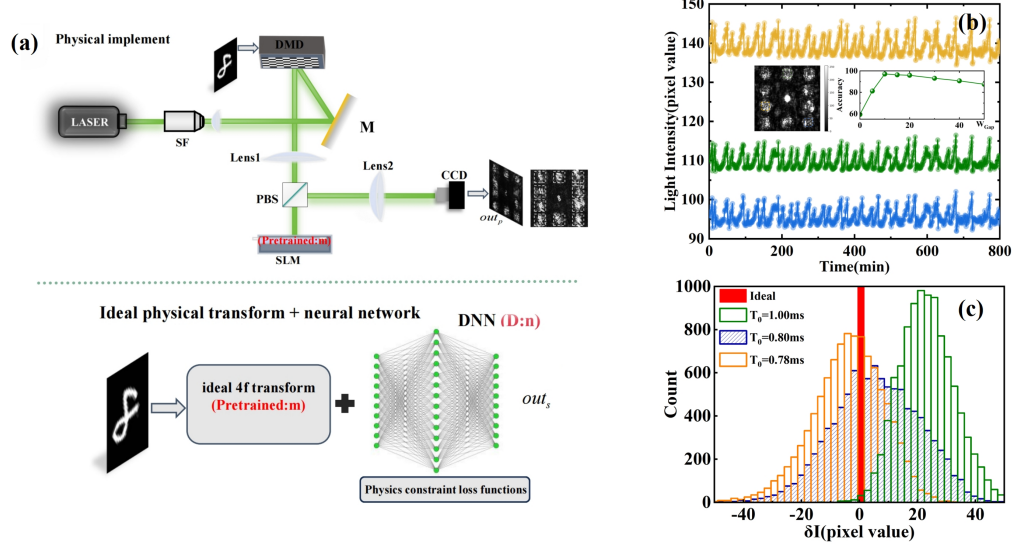


Figure 3: (a) Schematic of an image classification optical neural network with an error-compensating DNN incorporating quantitative physical information. (b) The random fluctuations of output light intensity, measured in three classification regions over 700 minutes. Inset: the experimental accuracy as a function of light intensity gap  $W_{Gap}$ . (c) The difference between simulated and experimental CCD reading values under various exposure times.

(orange bars in Fig 3(c)). This method effectively eliminates errors caused by the CCD and improves classification accuracy.

### 3.3 Training process of physics-prior-based error compensation network

In addition to the quantifiable system errors[47, 48, 49] previously discussed, several types of errors within optical systems are unmeasurable and can be categorized into fixed and coupling errors. Fixed errors arise from experimental deviations or mismatches, including the misalignment of the spatial light modulator, rotation of the input image, and diffraction effects due to lens apertures. Coupling errors, on the other hand, originate from the interaction between the light field and environmental impurities. As the phase map in the SLM is adjusted, while the light field is accordingly altered, the fixed errors remain unchanged, whereas the coupling errors vary. These experimental errors can affect the output light field of the 4f system, causing it to deviate from the output image in a perfect scenario, thereby reducing the recognition accuracy of ONN.

To mitigate the impact of such errors, a DNN network is used to model various environmental noises and experimental errors, aligning the ideal output image of the 4f system with the actual output image. The specific error compensation process is as below:

(i)*Pre-training*: Based on the Fresnel diffraction integral, train the simulational phase distribution  $m_0$  of SLM in the ONN without considering any experimental errors and environmental noises.

(ii)*Error compensation*: Load the phase map  $m_{k-1}$  onto the SLM to obtain the actual output image. Train the error compensation DNN  $n_k$  to minimize the difference between the ideal and actual output image.

(iii)*Re-training*: Load DNN to compensate for the errors and retrain a new phase map  $m_k$  of ONN to minimize the well-designed loss function (eq.4) in simulation. These two networks, ONN and DNN, are connected by residual connections.

(iv)*Iteration*: The coupling errors are altered when the phase map changes from  $m_{k-1}$  to  $m_k$ . To compensate for that, repeat steps (ii) and (iii), iterating until the experimental classification accuracy is maximized.

Through such an iteration process, the error compensation network can model effectively the environmental noises and experimental errors, significantly improving the classification accuracy.

Incorporating quantifiable physical data, such as laser intensity fluctuations and camera exposure effects, directly into the network’s architecture enables it to adapt more robustly to the inherent variability of optical systems. For more detailed information on the training process, please refer to the supplementary materials.

## 4 Result

### 4.1 Dataset

The MNIST (Modified National Institute of Standards and Technology)[50, 51] dataset consists of 70,000 28x28 pixel grayscale images of handwritten digits, widely used as a benchmark for training and testing image processing systems in machine learning and computer vision. The QuickDraw16 dataset, a subset of Google’s Quick, Draw! project includes 16 categories of hand-drawn images that mimic natural variations in handwritten and sketched drawings, serving as a valuable resource for image classification and recognition tasks. The Fashion MNIST dataset[52], provided by Zalando, contains 70,000 28x28 pixel grayscale images categorized into 10 different fashion items, such as T-shirts, trousers, shoes, and bags. Designed as a direct replacement for the traditional MNIST dataset, Fashion MNIST is extensively used for research and evaluation in image classification tasks. During training, we used the entire training set. To evaluate accuracy, we randomly selected 1,000 images from the test sets of each of the three datasets to assess experimental accuracy.

### 4.2 Error compensation network without quantifiable physical information

In comparison to our approach of incorporating quantifiable physical information into the error compensation network, we also conducted an experiment where a DNN was trained to model the system errors without introducing any quantifiable physical data. As illustrated in Fig 4(a1), the network trained only on the ideal transformation process  $f$  of the light field achieved an experimental accuracy of only 61.6%. This lower accuracy is due to several factors, including the instability of the laser  $\Delta f_{jit}$ , mismatches in exposure time  $\Delta f_{dev}$ , and various unmeasurable errors  $\eta$ , which resulted in significant deviations between the output light field images in experiments and simulations, as shown in Fig 4(a2-a3). The accuracy of image recognition depends on the distribution of light intensity across ten regions within the output image; thus, the experimental accuracy was significantly reduced.

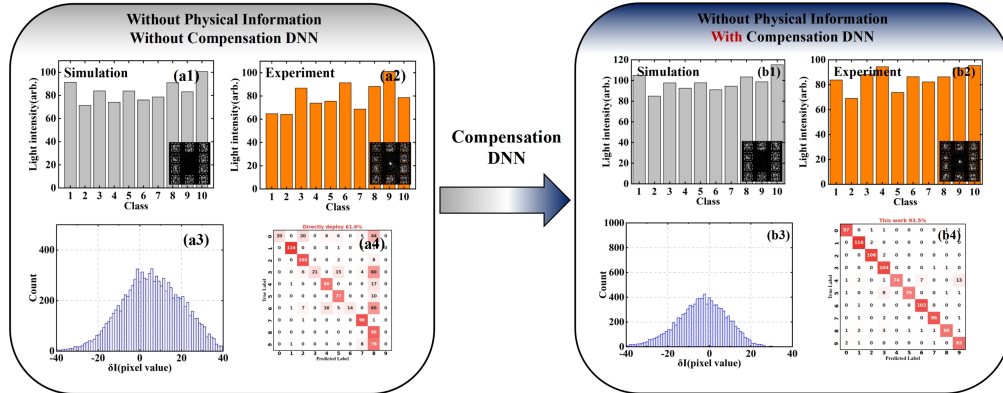


Figure 4: The simulation and experiment result with and without compensation DNN on MNIST dataset, maintaining without quantifiable physical information. (a1,b1) The simulational light intensity in ten classification regions is depicted in the inset. (a2,b2)The experimental light intensity in ten classification regions is depicted in the inset. (a3,b3) Histogram of light intensity difference  $\delta I$  between simulation and experiment. (a4,b4)Experimental confusion matrix.

After incorporating an error compensation network without quantifiable physical information, the discrepancies in light intensity distribution between experiments and simulations were mitigated, as shown in Fig 4(b4), boosting the experimental accuracy to 93.5%. As indicated in Fig 4(b3), the role of the error compensation DNN at this point was twofold: to modulate the overall light intensity and to adjust the local distribution of light intensity, thereby narrowing the gap between



simulation and experiment. However, as observed in Fig 4(b1-b2), the modulation resulted in a small difference between the highest and second-highest light intensities, reducing the robustness of the neural network.

### 4.3 Error compensation network with quantifiable physical information

By integrating quantifiable physical information, we effectively compensated for system errors caused by laser instability  $\Delta f_{jit}$ , and mismatches in exposure time, represented as  $\Delta f_{dev}$ . Consequently, the DNN now primarily focuses on compensating for unmeasurable errors  $\eta$ , which are inherently more challenging to predict and correct. This approach has significantly improved the convergence speed and accuracy of the error compensation network. For the MNIST dataset, we mitigated the error  $\Delta f_{dev}$  by adjusting the exposure of the CCD, thus ensuring that the overall light intensity between the simulation and experiment was consistent. This adjustment allowed the DNN to focus solely on modulating the local distribution of light intensity, thereby simplifying its complexity. As shown in Fig 4.3(a4), even without an error compensation network, the classification accuracy increased to 83.7% due to the consistency of the overall light intensity distribution.

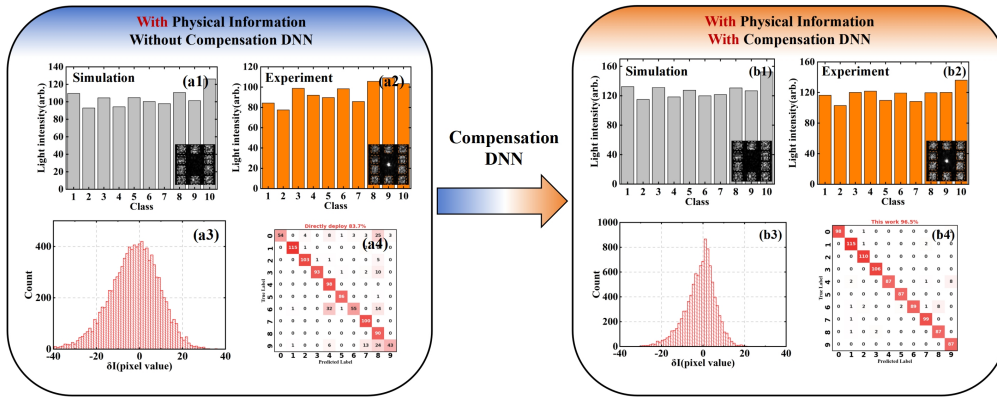


Figure 5: The simulation and experiment result with and without compensation DNN on MNIST dataset, maintaining with quantifiable physical information. (a1,b1) The simulation light intensity in ten classification regions is depicted in the inset. (a2,b2)The experimental light intensity in ten classification regions is depicted in the inset. (a3,b3) Histogram of light intensity difference  $\Delta I$  between simulation and experiment. (a4,b4)Experimental confusion matrix.

Additionally, a well-designed loss function was used to address the instability of laser intensity,  $\Delta f_{jit}$ , ensuring that the difference in light intensity between the maximum and the second maximum values within the classification regions was significantly greater than the range of fluctuations in laser intensity. After training, this error compensation network, enhanced with quantifiable physical information, increased the classification accuracy of the MNIST dataset to 96.5%.

### 4.4 Further experiment in different datasets

For the more complex Quickdraw16 dataset and FMNIST dataset, applying the same method of adjusting the CCD exposure to ensure consistency in overall light intensity, and using a well-designed loss function to address the instability of laser intensity significantly enhanced the recognition accuracy. The classification accuracy for the Quickdraw16 dataset increased from 63.0% to 85.7% through these adjustments. Similarly, the classification accuracy for the FMNIST dataset improved from 56.2% to 77.3% through these adjustments. The analysis of the convergence speed of the error compensation network can be found in the supplementary materials. Our approach can also handle more challenging datasets like CIFAR-10. To do this, we simply replace the dataset loaded onto the DMD with CIFAR-10. The initial classification accuracy with CIFAR-10 is 30%, which improves to 57% after optimization using our method. However, since current research[25, 37] on spatial 2D light-based ONNs predominantly uses the datasets mentioned in the paper, we do not present a detailed and specific demonstration of the results of CIFAR-10 in this paper due to the lack of comparative benchmarks.



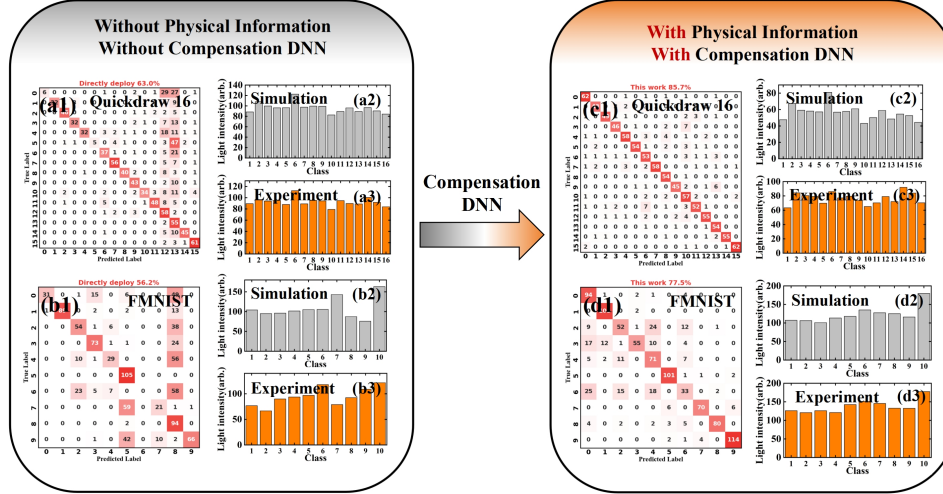


Figure 6: (a1) Experimental confusion matrix in Quickdraw 16 dataset without physical information and compensation DNN. (b1) Experimental confusion matrix in FMNIST dataset without physical information and compensation DNN. (c1) Experimental confusion matrix in Quickdraw 16 dataset with physical information and compensation DNN. (d1) Experimental confusion matrix in FMNIST dataset with physical information and compensation DNN. (a2,b2,c2,d2) The simulation light intensity in classification regions. (a3,b3,c3,d3) The experimental light intensity in classification regions.

#### 4.5 Accuracy comparison with other error compensation network

We have compared our work with other studies and found that our experimental classification accuracy on the MNIST dataset aligns closely with the results from Tsinghua University[25] and those published in Nature magazine; on the QuickDraw16 dataset, we achieved an experimental classification accuracy of 85%. These results indicate that, unlike models that do not incorporate physical information, our approach of quantitatively introducing physical information has enabled our optical-DNN to be fed with known measurable experimental perturbations and errors. This allows the DNN that follows the ideal model to learn more about unknown perturbations, errors, and imperfections in optical devices, and their interdependencies. Consequently, compared to neural networks that do not incorporate physical information quantitatively, our optical-DNN(optica-Deep Neural Network) is able to more accurately represent real experimental systems, enhancing the consistency between simulation and experimental results, and further improving experimental accuracy. Additionally, our error compensation network is very lightweight, with approximately 5,000 learnable parameters(For more details, please refer to the supplementary materials). Therefore, the introduction of quantitative physical information into neural networks is crucial, especially when dealing with complex datasets that require highly precise simulations of actual physical processes.

Architecture	Hybrid CNN[14]	DAT[25]		PAT[37]		This work Qualitative		This work Quantifiable	
		Directly deployed	optimize	Directly deployed	optimize	Directly deployed	optimize	Directly deployed	optimize
MNIST		24.9%	92.4%	24.9%	61.4%	61.6%	93.5%	83.7%	96.5%
FMNIST		8.4%	77.3%					56.2%	77.5%
Quickdraw16	72.0%							63.0%	85.7%

Table 1: The accuracy comparison between other error compensation network architecture in various datasets.

## 5 Conclusion

In this study, we explore the mapping relationship between input and output of a physical ONN system, which can be modeled as the sum of ideal transformation process, measurable errors (laser-

related deviations, CCD-related deviations) and other unmeasurable errors. To reduce the disparity between the ideal simulated ONN model and the real physical ONN model, a physics-constrained ONN learning framework is constructed for image classification tasks. Specifically, we introduce a well-designed loss function to mitigate the laser-related deviations and a CCD adjustment strategy to reduce the CCD-related deviations. In addition, a 'physics-prior-based' error compensation network is proposed to manage other unmeasurable errors. The effectiveness of our approach is demonstrated through extensive experiments on the MNIST, Fashion MNIST and QuickDraw16 datasets.

Although we aim to reduce the gap between the ideal simulated ONN model and the real physical ONN model, it is important to note that our research primarily focuses on the image classification task. Specifically, the intensity gap loss function we propose may not be applicable to other AI tasks. However, we believe that the quantifiable error analysis and processing methods could serve as valuable references for other applications, as the two quantifiable errors—light source instability and exposure time mismatches—are inevitable in current optical systems. For future work, we plan to develop a more generalizable approach to address systematic errors in ONN systems and investigate its application in other AI tasks, such as image restoration and image segmentation.

## Acknowledgments and Disclosure of Funding

This work was supported by the Shanghai Outstanding Academic Leaders Plan (No. 20XD1402000) and the Scientific and Technological Innovation Funds of Shanghai Jiao Tong University. Additionally, it received partial support from the National Natural Science Foundation of China under grant U23B2011.

## References

- [1] Xiangyan Meng, Guojie Zhang, Nuannuan Shi, Guangyi Li, José Azaña, José Capmany, Jianping Yao, Yichen Shen, Wei Li, Ninghua Zhu, et al. Compact optical convolution processing unit based on multimode interference. *Nature Communications*, 14(1):3000, 2023.
- [2] Ziyu Gu, Yesheng Gao, and Xingzhao Liu. Optronic convolutional neural networks of multi-layers with different functions executed in optics for image classification. *Optics Express*, 29(4):5877–5889, 2021.
- [3] Yanbing Liu, Shaochong Liu, Tao Li, Tianyu Li, Wei Li, Guoqing Wang, Xun Liu, Wei Yang, and Yuan'an Liu. Towards constructing a doe-based practical optical neural system for ship recognition in remote sensing images. *Signal Processing*, page 109488, 2024.
- [4] Xing Lin, Yair Rivenson, Nezih T Yardimci, Muhammed Veli, Yi Luo, Mona Jarrahi, and Aydogan Ozcan. All-optical machine learning using diffractive deep neural networks. *Science*, 361(6406):1004–1008, 2018.
- [5] Tianyu Wang, Shi-Yuan Ma, Logan G Wright, Tatsuhiro Onodera, Brian C Richard, and Peter L McMahon. An optical neural network using less than 1 photon per multiplication. *Nature Communications*, 13(1):123, 2022.
- [6] HH Zhu, Jun Zou, Hengyi Zhang, YZ Shi, SB Luo, N Wang, H Cai, LX Wan, Bo Wang, XD Jiang, et al. Space-efficient optical computing with an integrated chip diffractive neural network. *Nature communications*, 13(1):1044, 2022.
- [7] Ryan Hamerly, Liane Bernstein, Alexander Sludds, Marin Soljačić, and Dirk Englund. Large-scale optical neural networks based on photoelectric multiplication. *Physical Review X*, 9(2):021032, 2019.
- [8] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436–444, 2015.
- [9] Andreas Kamilaris and Francesc X Prenafeta-Boldú. Deep learning in agriculture: A survey. *Computers and electronics in agriculture*, 147:70–90, 2018.
- [10] Yanming Guo, Yu Liu, Ard Oerlemans, Songyang Lao, Song Wu, and Michael S Lew. Deep learning for visual understanding: A review. *Neurocomputing*, 187:27–48, 2016.

- [11] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *nature*, 323(6088):533–536, 1986.
- [12] Ali Momeni, Babak Rahmani, Matthieu Malléjac, Philipp Del Hougne, and Romain Fleury. Backpropagation-free training of deep physical neural networks. *Science*, 382(6676):1297–1303, 2023.
- [13] Sunil Pai, Zhanghao Sun, Tyler W Hughes, Taewon Park, Ben Bartlett, Ian AD Williamson, Momchil Minkov, Maziyar Milanizadeh, Nathnael Abebe, Francesco Morichetti, et al. Experimentally realized in situ backpropagation for deep learning in photonic neural networks. *Science*, 380(6643):398–404, 2023.
- [14] Julie Chang, Vincent Sitzmann, Xiong Dun, Wolfgang Heidrich, and Gordon Wetzstein. Hybrid optical-electronic convolutional neural networks with optimized diffractive optics for image classification. *Scientific reports*, 8(1):1–10, 2018.
- [15] Mario Miscuglio, Zibo Hu, Shurui Li, Jonathan K George, Roberto Capanna, Hamed Dalir, Philippe M Bardet, Puneet Gupta, and Volker J Sorger. Massively parallel amplitude-only fourier neural network. *Optica*, 7(12):1812–1819, 2020.
- [16] Deniz Mengü, Yifan Zhao, Nezih T Yardimci, Yair Rivenson, Mona Jarrahi, and Aydogan Ozcan. Misalignment resilient diffractive optical networks. *Nanophotonics*, 9(13):4207–4219, 2020.
- [17] Hoda Sadeghzadeh and Somayyeh Koochi. Translation-invariant optical neural network for image classification. *Scientific Reports*, 12(1):17232, 2022.
- [18] Hans-Christian Ruiz Euler, Marcus N Boon, Jochem T Wildeboer, Bram van de Ven, Tao Chen, Hajo Broersma, Peter A Bobbert, and Wilfred G van der Wiel. A deep-learning approach to realizing functionality in nanoelectronic devices. *Nature nanotechnology*, 15(12):992–998, 2020.
- [19] Yichen Shen, Nicholas C Harris, Scott Skirlo, Mihika Prabhu, Tom Baehr-Jones, Michael Hochberg, Xin Sun, Shijie Zhao, Hugo Larochelle, Dirk Englund, et al. Deep learning with coherent nanophotonic circuits. *Nature photonics*, 11(7):441–446, 2017.
- [20] Julian Bueno, Sheler Maktoobi, Luc Froehly, Ingo Fischer, Maxime Jacquot, Laurent Larger, and Daniel Brunner. Reinforcement learning in a large-scale photonic recurrent neural network. *Optica*, 5(6):756–760, 2018.
- [21] Yuchi Huo, Hujun Bao, Yifan Peng, Chen Gao, Wei Hua, Qing Yang, Haifeng Li, Rui Wang, and Sung-Eui Yoon. Optical neural network via loose neuron array and functional learning. *Nature Communications*, 14(1):2535, 2023.
- [22] Darcy Bullock, Brian Johnson, Richard B Wells, Michael Kyte, and Zhen Li. Hardware-in-the-loop simulation. *Transportation Research Part C: Emerging Technologies*, 12(1):73–89, 2004.
- [23] Sebastian Schmitt, Johann Klähn, Guillaume Bellec, Andreas Grübl, Maurice Guettler, Andreas Hartel, Stephan Hartmann, Dan Husmann, Kai Husmann, Sebastian Jeltsch, et al. Neuromorphic hardware in the loop: Training a deep spiking network on the brainscales wafer-scale system. In *2017 international joint conference on neural networks (IJCNN)*, pages 2227–2234. IEEE, 2017.
- [24] James Spall, Xianxin Guo, and Alexander I Lvovsky. Hybrid training of optical neural networks. *Optica*, 9(7):803–811, 2022.
- [25] Ziyang Zheng, Zhengyang Duan, Hang Chen, Rui Yang, Sheng Gao, Haiou Zhang, Hongkai Xiong, and Xing Lin. Dual adaptive training of photonic neural networks. *Nature Machine Intelligence*, 5(10):1119–1129, 2023.
- [26] Shengze Cai, Zhiping Mao, Zhicheng Wang, Minglang Yin, and George Em Karniadakis. Physics-informed neural networks (pinns) for fluid mechanics: A review. *Acta Mechanica Sinica*, 37(12):1727–1738, 2021.

- [27] Salvatore Cuomo, Vincenzo Schiano Di Cola, Fabio Giampaolo, Gianluigi Rozza, Maziar Raissi, and Francesco Piccialli. Scientific machine learning through physics-informed neural networks: Where we are and what's next. *Journal of Scientific Computing*, 92(3):88, 2022.
- [28] Zhiping Mao, Ameya D Jagtap, and George Em Karniadakis. Physics-informed neural networks for high-speed flows. *Computer Methods in Applied Mechanics and Engineering*, 360:112789, 2020.
- [29] 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.
- [30] Yin hao Zhu, Nicholas Zabaras, Phaedon-Stelios Koutsourelakis, and Paris Perdikaris. Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data. *Journal of Computational Physics*, 394:56–81, 2019.
- [31] Mo Deng, Shuai Li, Zhengyun Zhang, Iksung Kang, Nicholas X Fang, and George Barbastathis. On the interplay between physical and content priors in deep learning for computational imaging. *Optics Express*, 28(16):24152–24170, 2020.
- [32] Sudhir Cherukulappurath, Georges Boudebs, and André Monteil. 4f coherent imager system and its application to nonlinear optical measurements. *JOSA B*, 21(2):273–279, 2004.
- [33] Nien-An Chang and Nicholas George. Speckle in the 4f optical system. *Applied optics*, 47(4):A13–A20, 2008.
- [34] Hoda Sadeghzadeh and Somayyeh Koochi. High-speed multi-layer convolutional neural network based on free-space optics. *IEEE Photonics Journal*, 14(4):1–12, 2022.
- [35] CJR Sheppard and M Hrynevych. Diffraction by a circular aperture: a generalization of fresnel diffraction theory. *JOSA A*, 9(2):274–281, 1992.
- [36] Yusuf Z Umul. Equivalent functions for the fresnel integral. *Optics express*, 13(21):8469–8482, 2005.
- [37] Logan G Wright, Tatsuhiko Onodera, Martin M Stein, Tianyu Wang, Darren T Schachter, Zoey Hu, and Peter L McMahan. Deep physical neural networks trained with backpropagation. *Nature*, 601(7894):549–555, 2022.
- [38] Karl S Kunz and Raymond J Luebbers. *The finite difference time domain method for electromagnetics*. CRC press, 1993.
- [39] Allen Taflove, Susan C Hagness, and Melinda Picket-May. Computational electromagnetics: the finite-difference time-domain method. *The Electrical Engineering Handbook*, 3(629-670):15, 2005.
- [40] Allen Taflove. Review of the formulation and applications of the finite-difference time-domain method for numerical modeling of electromagnetic wave interactions with arbitrary structures. *Wave Motion*, 10(6):547–582, 1988.
- [41] Simone Gaiarin, Francesco Da Ros, Rasmus T Jones, and Darko Zibar. End-to-end optimization of coherent optical communications over the split-step fourier method guided by the nonlinear fourier transform theory. *Journal of Lightwave Technology*, 39(2):418–428, 2020.
- [42] Oleg V Sinkin, Ronald Holzlohner, John Zweck, and Curtis R Menyuk. Optimization of the split-step fourier method in modeling optical-fiber communications systems. *Journal of lightwave technology*, 21(1):61, 2003.
- [43] GM Muslu and HA Erbay. A split-step fourier method for the complex modified korteweg-de vries equation. *Computers & Mathematics with Applications*, 45(1-3):503–514, 2003.
- [44] Qi Wang, Yue Ma, Kun Zhao, and Yingjie Tian. A comprehensive survey of loss functions in machine learning. *Annals of Data Science*, pages 1–26, 2020.

- [45] Lorenzo Rosasco, Ernesto De Vito, Andrea Caponnetto, Michele Piana, and Alessandro Verri. Are loss functions all the same? *Neural computation*, 16(5):1063–1076, 2004.
- [46] Katarzyna Janocha and Wojciech Marian Czarnecki. On loss functions for deep neural networks in classification. *arXiv preprint arXiv:1702.05659*, 2017.
- [47] Youfang Cao, Anna Terebus, and Jie Liang. State space truncation with quantified errors for accurate solutions to discrete chemical master equation. *Bulletin of mathematical biology*, 78:617–661, 2016.
- [48] Rasool Khadem, Clement C Yeh, Mohammad Sadeghi-Tehrani, Michael R Bax, Jeremy A Johnson, Jacqueline Nerney Welch, Eric P Wilkinson, and Ramin Shahidi. Comparative tracking error analysis of five different optical tracking systems. *Computer Aided Surgery*, 5(2):98–107, 2000.
- [49] Haitao Sun and Jochen Autschbach. Influence of the delocalization error and applicability of optimal functional tuning in density functional calculations of nonlinear optical properties of organic donor–acceptor chromophores. *ChemPhysChem*, 14(11):2450–2461, 2013.
- [50] Alejandro Baldominos, Yago Saez, and Pedro Isasi. A survey of handwritten character recognition with mnist and emnist. *Applied Sciences*, 9(15):3169, 2019.
- [51] Norman Mu and Justin Gilmer. Mnist-c: A robustness benchmark for computer vision. *arXiv preprint arXiv:1906.02337*, 2019.
- [52] Han Xiao, Kashif Rasul, and Roland Vollgraf. Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms. *arXiv preprint arXiv:1708.07747*, 2017.

## NeurIPS paper checklis

### 1. Claims

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

Answer: [Yes]

Justification: [TODO]

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: Section5

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: [Yes]

Justification: **[TODO]**

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: Section4

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: Supplemental material and Section4

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, hyper-parameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [Yes]

Justification: Section4

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: [Yes]

Justification: [TODO]

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: Supplemental material

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?](https://neurips.cc/public/EthicsGuidelines)

Answer: [Yes]

Justification: [TODO]

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: [TODO]

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: [TODO]

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: Section4

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](https://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: [NA]

Justification: [TODO]

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: [No]

Justification: [TODO]

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: [No]

Justification: [TODO]

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.