

Supplementary information for
Wavefront-aberration-tolerant diffractive
deep neural networks using volume
holographic optical elements

Ikuo Hoshi^{1*}, Koki Wakunami¹, Yasuyuki Ichihashi¹
and Ryutaro Oi¹

^{1*}Applied Electromagnetic Research Center, National Institute of
Information and Communications Technology, Nukui-Kitamachi,
Koganei, 184-8795, Tokyo, Japan.

*Corresponding author(s). E-mail(s): hoshi@nict.go.jp;
Contributing authors: k.wakunami@nict.go.jp;
y-ichihashi@nict.go.jp; oi.ryutaro@nict.go.jp;

This pdf file includes:

Figure S1 – S9.

Algorithm S1 – S3.

Source code S1.

1 diffraction efficiency

Figure S1 (a), (b), and (c) show the transmission spectrum of each layer used for obtaining diffraction efficiencies. The diffraction efficiencies were obtained by calculating $\frac{S_{\text{Base}} - S_{\text{Peak}}}{S_{\text{Base}}}$ from S_{Base} and S_{Peak} in these graphs [1]. For off-axis D²NN with vHOEs shown in Figs. 2 and 3 in the main text, the diffraction efficiencies were 90.1% for the first layer, 90.0% for the second layer, and 88.8% for the third layer. Thus, the overall diffraction efficiency of the three layers was 72.0%. This diffraction efficiency and transparency should be improved by further optimization of the parameters of recording and after processing or by improving the performance of the hologram-recording material.

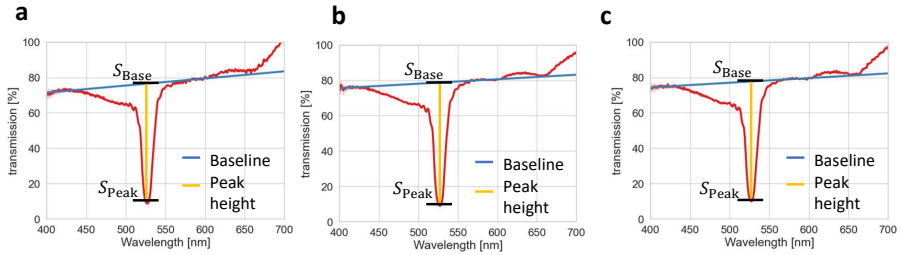


Fig. S1 Transmission spectrum of each layer: (a) first layer, (b) second layer, and (c) third layer.

2 experiment results

Figures S2 and S3 show all the output wavefronts in the optical experiment shown in Fig. 3 without and with aberration-adaptive learning, respectively. A red frame means the correct answer, and numbers below the output wavefront are prediction results.

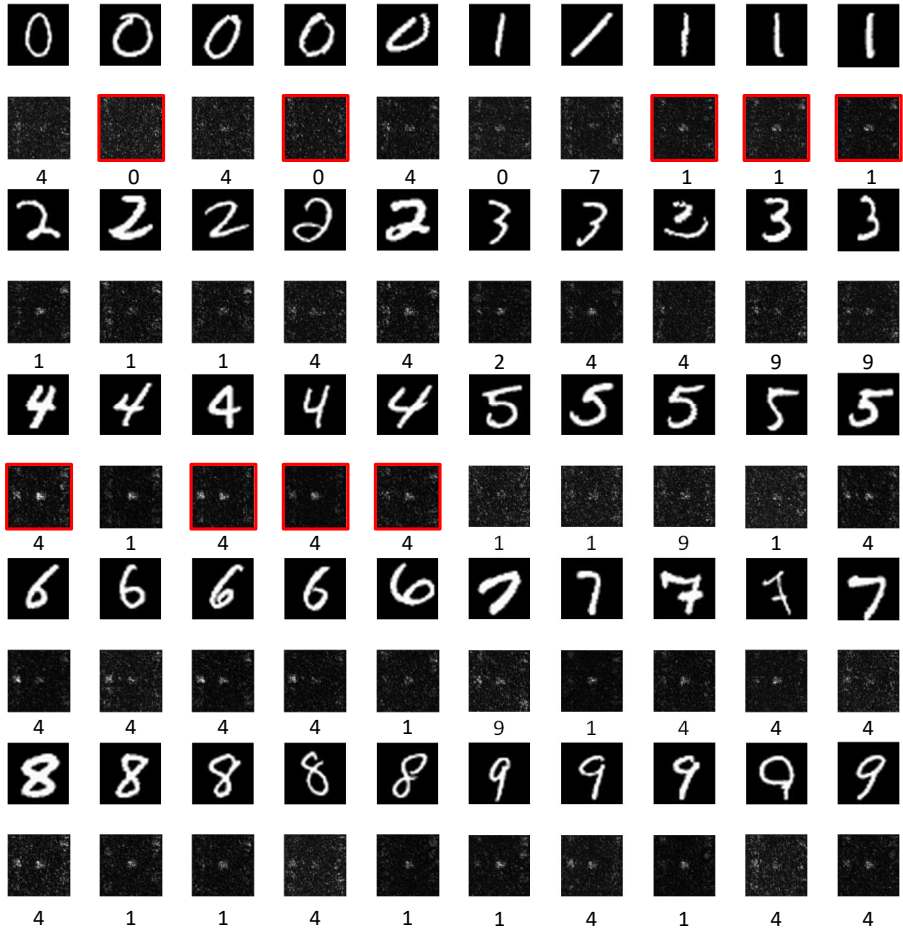


Fig. S2 All output wavefronts in the optical experiment without aberration-adaptive learning.

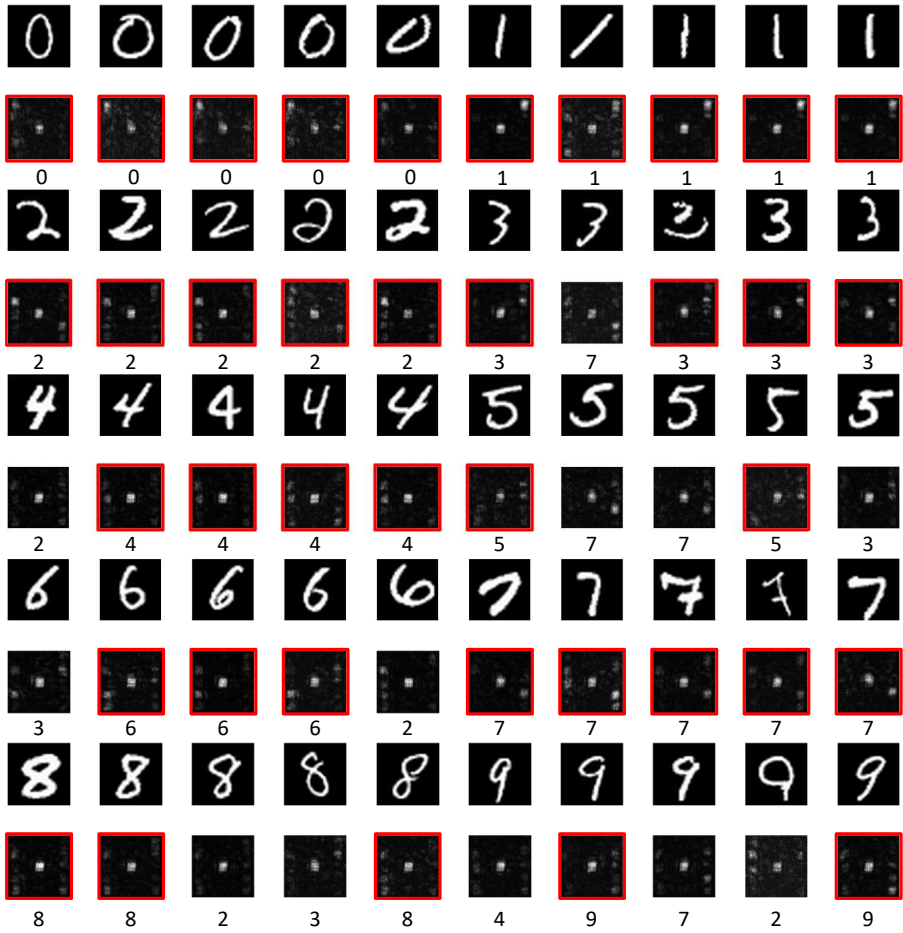


Fig. S3 All output wavefronts in the optical experiment with aberration-adaptive learning.

3 Additional simulations

Figures S4 – S6 show the additional simulations in the case of using fashion mnist as a dataset. Figure S4 corresponds to Section 2.2 in the main text, Figure S5 corresponds to Section 2.3 in the main text, and Figure S6 corresponds to Section 2.4 in the main text, with the same conditions except for the dataset. The training data were resized to 896×896 pixel and then zero-padded to 1024×1024 pixel.

Although the accuracy is degraded in certain classes in Fig. S4, the results are generally similar to those of the mnist case in all simulations. Regarding the performance degradation of certain classes, Figure S4 shows that other classes are classified correctly with aberration-adaptive learning and that the accuracies of classes 2, 5, and 6 are lower than for other classes even without aberration-adaptive learning. In particular, classes 2 and 5 are less confident in their answers because the output wavefronts have less light in the area of the correct answer than the others. This indicates that the classification of these classes is not easy even with the D²NN without aberration-adaptive learning. Therefore, we consider that our D²NN requires base performance improvement by increasing the amount of training data or the number of pixels as neurons.

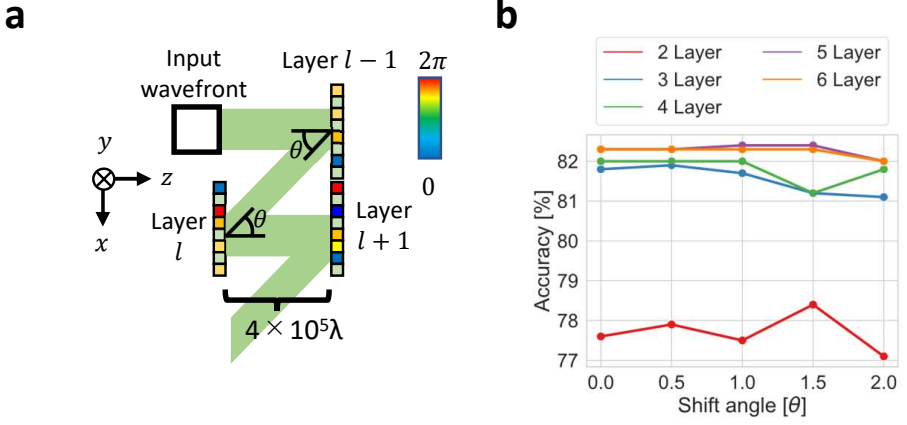


Fig. S5 The simulation result of Section 2.3 in the main text when using fashion mnist as dataset. (a) Simulation model of off-axis D²NN. Multiple models were created, trained, and evaluated for each angle and number of layers. (b) Graph of accuracies for shift angle θ , where the shift angle θ corresponds to θ in (a), and each color represents the number of model layers. The accuracy was plotted for $\theta = [0.0^\circ, 0.5^\circ, 1.0^\circ, 1.5^\circ, 2.0^\circ]$.

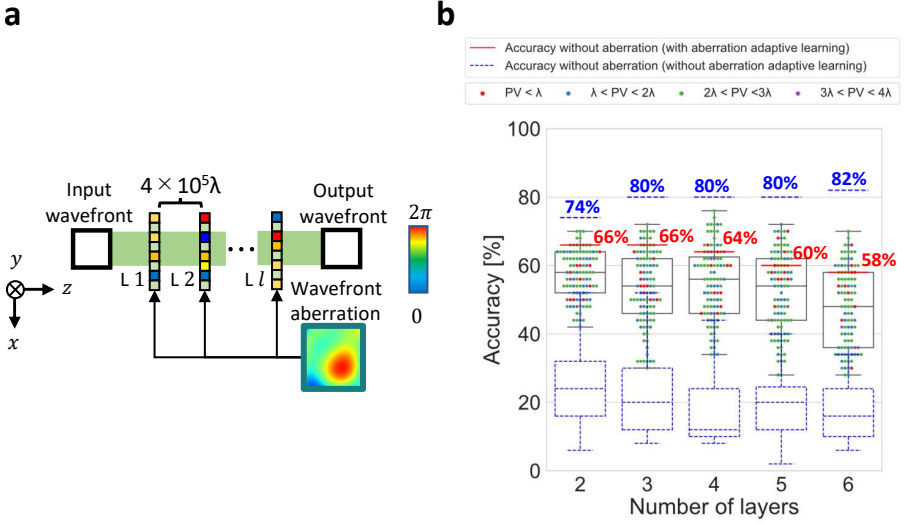


Fig. S6 The simulation result of Section 2.4 in the main text when using fashion mnist as dataset. (a) Inline D²NN model for performance evaluation. The distances between each layer are $4.0 \times 10^5 \lambda$. (b) Simulation results of aberration-adaptive learning. The red lines represent the accuracies without wavefront aberration, each colored point represents the classification accuracy of one out of 100 trials and the strength of the added wavefront aberration, and the box plots represent the results of 100 trials. The blue dashed lines and the box plots are results without aberration-adaptive learning.

4 Aberration-adaptive learning

4.1 Model architecture

Figure S7 shows two-step learning. In the case of single-step learning, the initial phases are set instead of the phases obtained by learning in the first step. In aberration-adaptive learning, the wavefront aberration added layer is inserted after the diffraction layer. By doing this, the variable of the modulation amount to be learned can be adapted to wavefront aberrations since the learning process proceeds with the wavefront aberration overlapping the modulation amount. A sample source code of the model definition is shown in source code S1 and the algorithm for each layer in Fig. S7 is shown in Algorithm S1 – S3.

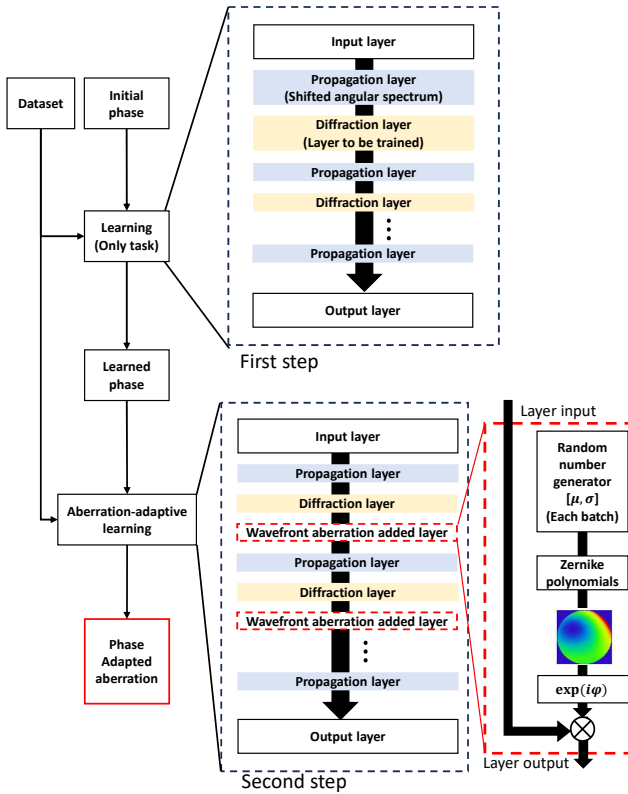


Fig. S7 Network structure of aberration-adaptive learning.

source code S1 sample code of model definition

[illegible]

```

106
107     distance = distance_list[-1]
108     x_offset = x_offset_list[-1]
109     y_offset = y_offset_list[-1]
110
111     x = ZeroPadding2D(padding = (input_shape[2] // 2,
112                                input_shape[3] // 2),
113                       data_format = "channels_first")(x)
114     x = ShiftedAngularSpectrum(wave_lambda,
115                               distance,
116                               pitch,
117                               x_offset = x_offset,
118                               y_offset = y_offset)(x)
119     x = Cropping2D(cropping = (input_shape[2] // 2,
120                               input_shape[3] // 2),
121                   data_format = "channels_first")(x)
122
123     model = Model(inputs = inputs, outputs = x)
124     model.compile(optimizer=optimizer,
125                 loss=loss)
126     return model
127

```

Algorithm S1 Propagation layer (Shifted angular spectrum [2])

Input: $\tilde{x}_l, x_0, y_0, z_0, width, height, p, \lambda, \Delta u, \Delta v$

 // $width$ and $height$ are values after padding.

Output: $\tilde{x}_{l+1}$

 1: **Initialization:**

 2: $u = -\frac{1}{2p}, -\frac{1}{2p} + \Delta u, \dots, \frac{1}{2p}$
 $v = -\frac{1}{2p}, -\frac{1}{2p} + \Delta v, \dots, \frac{1}{2p}$
 $u_{limit}^{\pm} = \left[\left(x_0 \pm \frac{1}{2\Delta u} \right)^{-2} z_0^2 + 1 \right]^{-\frac{1}{2}} \lambda^{-1}$
 $v_{limit}^{\pm} = \left[\left(y_0 \pm \frac{1}{2\Delta v} \right)^{-2} z_0^2 + 1 \right]^{-\frac{1}{2}} \lambda^{-1}$
for u, v **do**
 $w = (\lambda^{-2} - u^2 - v^2)^{1/2}$

 //The case when the w becomes imaginary is not
cared about because

 //the w does not become imaginary in the parameters
of this study.

 $H(u, v) = \exp[i2\pi(x_0 u + y_0 v + z_0 w)]$

 if $\frac{width}{2}p < x_0$
for u **do if** not $(u_{limit}^- \leq u \leq u_{limit}^+)$ **then**
 $H(u, v) = 0$

 else if $-\frac{width}{2}p \leq x_0 \leq \frac{width}{2}p$
for u **do if** not $(-u_{limit}^- \leq u \leq u_{limit}^+)$ **then**
 $H(u, v) = 0$

 else if $x_0 < -\frac{width}{2}p$
for u **do if** not $(-u_{limit}^- \leq u \leq -u_{limit}^+)$ **then**
 $H(u, v) = 0$

 if $\frac{height}{2}p < y_0$
for v **do if** not $(v_{limit}^- \leq v \leq v_{limit}^+)$ **then**
 $H(u, v) = 0$

 else if $-\frac{height}{2}p \leq y_0 \leq \frac{height}{2}p$
for v **do if** not $(-v_{limit}^- \leq v \leq v_{limit}^+)$ **then**
 $H(u, v) = 0$

 else if $y_0 < -\frac{height}{2}p$
for v **do if** not $(-v_{limit}^- \leq v \leq -v_{limit}^+)$ **then**
 $H(u, v) = 0$

//band-limited

 3: $X(u, v) = \mathcal{F}[\tilde{x}_l]$ // $\mathcal{F}$ denotes Fourier transform.

 4: **return** $\mathcal{F}^{-1}[X(u, v) \times H(u, v)]$

Algorithm S2 Diffraction layer

Input: $\tilde{x}$ // Wavefront. Data type is Complex.**Output:** $\tilde{x}'$ // Wavefront after phase modulation.**Initialization:** $\psi = \theta_0$

// Learnable variable

return $\tilde{x} \times e^{i\psi}$

Algorithm S3 Wavefront aberration added layer. (A bimodal Gaussian distribution is used.)

Input: $\tilde{x}, z_4(x, y), z_5(x, y), z_6(x, y), \mu, \sigma$ // z_4, z_5, z_6 are denotes Zernike olynomials corresponding to oblique astigmatism, defocus, and vertical astigmatism.**Output:** $\tilde{x}'$ // Aberration added wavefront.1: **for** $i \leftarrow 4$ to 6 **do**2: $choice = rand.uniform(min = 0, max = 1)$ // A uniform random number of 0 or 1 is generated.3: $choice_{inverse} = 1 - choice$ 4: $coefficient_1 = rand.normal(mean = \mu, std = \sigma)$ 5: $coefficient_2 = rand.normal(mean = -\mu, std = \sigma)$ 6: $c_i = coefficient_1 \times choice + coefficient_2 \times choice_{inverse}$ 7: **end for**8: $\varphi = c_4 \times z_4(x, y) + c_5 \times z_5(x, y) + c_6 \times z_6(x, y)$ 9: **return** $\tilde{x} \times e^{i\varphi}$

4.2 Random number generator

Figure S8 shows the generated random number and peak-to-valley distance (PV) value distributions. Aberration-adaptive learning uses random numbers to generate the coefficients of the Zernike polynomials to generate the various wavefront aberrations with PVs in a specific range. The PV range should be close to the PV of the wavefront aberrations overlapped on the system, which is often expected to be normally distributed. However, the general Gaussian distribution generates many values near zero when trying to generate equal positive and negative values. As a result, while the range of PVs is extended, the mode is close to zero, which prevents aberration-adaptive learning from working well in real space. To avoid this, and to generate positive and negative coefficients equally, we adopted the probability distribution shown in Fig. S8 (a) for generating the coefficients of the Zernike polynomials. The PV value distribution of the wavefront aberration generated from the probability distribution shown in Fig. S8 (a) is shown in Fig. S8 (b).

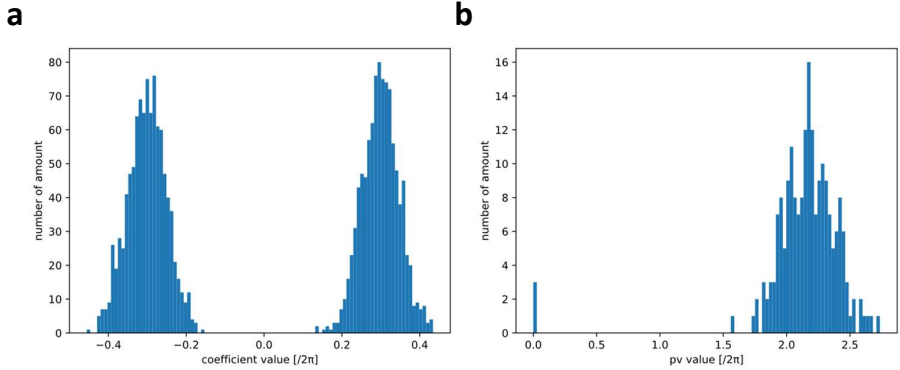


Fig. S8 (a) Probability distribution used in this study. The D²NN constructed in this study needed to be adapted to the PV distribution of the wavefront aberration shown in (b). Therefore, we used the bimodal normal distribution to obtain the wavefront aberration distribution. (b) Distribution of PV values of wavefront generated from the probability distributions shown in (a) and Zernike polynomials.

In Fig. S8, the mean μ and standard deviation σ are set to $\pm 0.3 \times 2\pi$ and $0.05 \times 2\pi$, respectively, and these values are used in the experiment. These parameters were determined by performance comparison with some parameters added learning rate shown in Fig. S9. To determine the above parameters, we pre-selected and compared some candidates that would generate wavefront aberrations close to those measured experimentally by the wavefront sensor for the vHOEs fabricated. Specifically, we determined the parameters based on median accuracy and number of occurrences for each model. As shown in Fig. S9, the four models as follows have the highest median accuracy at 70% and the highest upper quartile accuracy at 74%: $[\mu, \sigma, \text{learning rate}] = [\pm 0.3 \times 2\pi, 0.05 \times 2\pi, 0.1], [\pm 0.3 \times 2\pi, 0.05 \times 2\pi, 0.05], [\pm 0.35 \times 2\pi, 0.1 \times 2\pi, 0.05], [\pm 0.3 \times$

154 $2\pi, 0.05 \times 2\pi, 0.025]$. Among them, we determined the parameters used in our
 155 experiment to be $[\mu, \sigma, \text{learning rate}] = [\pm 0.3 \times 2\pi, 0.05 \times 2\pi, 0.1]$ since that
 156 model has the highest number of occurrences at 30 in the top quartile group
 of the box plot among the four models.

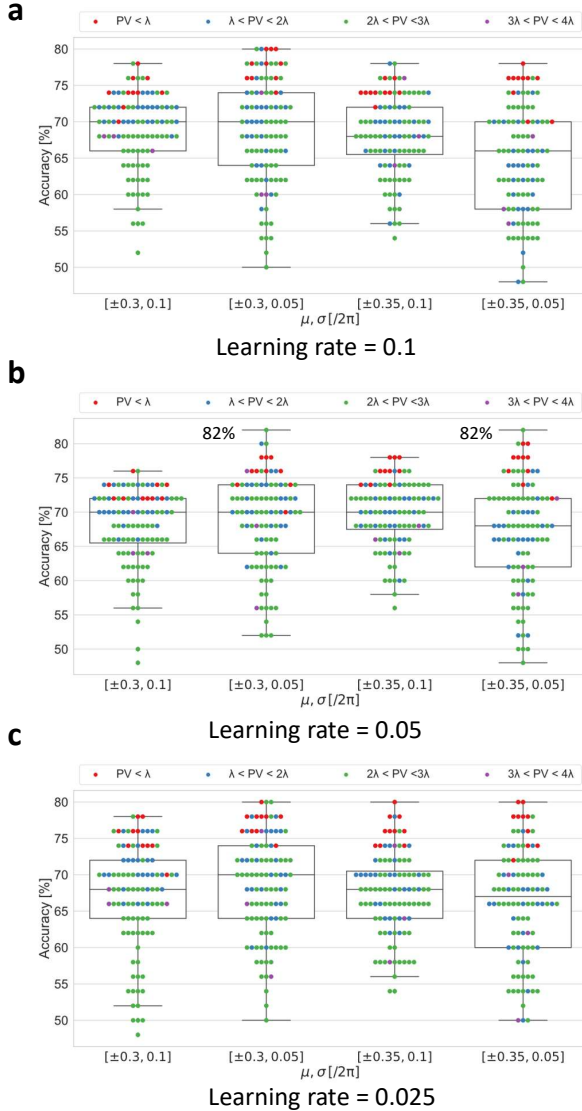


Fig. S9 Evaluation of each model with different parameters for learning rate, mean μ , and standard deviation σ : (a) learning rate 0.1, (b) learning rate 0.05, and (c) learning rate 0.025.

References

- [1] Bruder, F.-K., Frank, J., Hansen, S., Künzel, R., Künzel, J., Lorenz, A., Manecke, C., Meisenheimer, R., Mills, J., Pitzer, L., *et al.*: Reliable photopolymer for new applications: vhoes recorded into bayfol hx film resisting the environment. In: Practical Holography XXXVI: Displays, Materials, and Applications, vol. 12026, p. 1202602 (2022). SPIE
- [2] Matsushima, K.: Shifted angular spectrum method for off-axis numerical propagation. *Optics Express* **18**(17), 18453–18463 (2010)