

# Inverse Design of Structural Color: Finding Multiple Solutions via Conditional Generative Adversarial Networks

Peng Dai<sup>1</sup>, Kai Sun<sup>1</sup>, Xingzhao Yan<sup>1</sup>, Otto L. Muskens<sup>1</sup>, C.H. (Kees) de Groot<sup>1</sup>, Xupeng Zhu<sup>3</sup>, Yueqiang Hu<sup>2</sup>, Huigao Duan<sup>2</sup>, Ruomeng Huang<sup>1\*</sup>

<sup>1</sup>Faculty of Engineering and Physical Sciences, University of Southampton, Southampton, United Kingdom, SO17 1BJ

<sup>2</sup>National Engineering Research Center for High Efficiency Grinding, College of Mechanical and Vehicle Engineering, Hunan University, Changsha, China, 410082

<sup>3</sup>School of Physics Science and Technology, Lingnan Normal University, Zhanjiang, China, 5240481

\*r.huang@soton.ac.uk

These authors contributed equally: Peng Dai, Kai Sun

## Abstract

The “one-to-many” problem is a typical challenge that faced by many machine learning aided inverse nanophotonics designs where one target optical response can be achieved by many solutions (designs). Although novel training approaches, such as tandem network, and network architecture, such as the mixture density model, have been proposed, the critical problem of solution degeneracy still exists where some possible solutions or solution spaces are discarded or unreachable during the network training process. Here, we report a solution to the “one-to-many” problem by employing a conditional generative adversarial network (cGAN) that enables generating sets of multiple solution groups to a design problem. Using the inverse design of a transmissive Fabry-Pérot-cavity-based color filter as an example, our model demonstrates the capability of generating an average number of 3.58 solution groups for each color. These multiple solutions allow the selection of the best design for each color which results in a record high accuracy with an average index color difference  $\Delta E$  of 0.44. The capability of identifying multiple solution groups can benefit the design manufacturing to allow more viable designs for fabrication. The capability of our cGAN is verified experimentally by inversely designing the RGB color filters. We envisage this cGAN-based design methodology can be applied to other nanophotonic structures or physical science domains where the identification of multi-solution across a vast parameter space is required.

**Keywords:** Inverse design, one-to-many problem, generative adversarial networks, deep learning, structural color, Fabry-Pérot cavity.

## Introduction

Retrieval of structural arrangements for a desired optical performance has been the fundamental challenge in nanophotonics inverse design.<sup>[1]</sup> This challenging task requires searching a range of geometrical parameters in an ample design space. For the past decades, sophisticated methods have been developed to arrive at optimized solutions using heuristics-based approaches such as particle swarm,<sup>[2,3]</sup> genetic algorithms,<sup>[4–6]</sup> and gradient-based approaches.<sup>[7,8]</sup> Often, the intuition and experience of the designer have been critical in obtaining rational operational principles with the advantage that the performance of such structures can be understood.<sup>[7,9–12]</sup> However, such a rational approach can be rather unsatisfactory as human intuition will not always guarantee an optimal design with the highest possible performance.

The recent remarkable development of deep learning (DL) has offered great promise in solving some of the critical challenges faced in nanophotonics.<sup>[1,13–15]</sup> DL is a subset of machine learning technology inspired by animal brains' biological neural networks.<sup>[16]</sup> Through a training process that involves iterative optimization of network parameters *via* gradient descent,<sup>[17]</sup> the DL model, which contains a few artificial neuron layers, can learn the non-linear relations between the photonic structure and optical response. In recent years, it has found many successful applications in photonic research such as chiral metamaterials,<sup>[18,19]</sup> wavefront engineering,<sup>[20–22]</sup> and structural color.<sup>[23–27]</sup> An exhaustive overview of this field of study can be found in the review literatures.<sup>[14,15]</sup>

In the inverse design of nanophotonic structures, the DL regression models have been utilized to predict the potential solution of optical response and achieved success to some extent under some specific regions and constraints.<sup>[9,25,28–31]</sup> One of the main difficulties for inverse design is the “one-to-many” mapping characteristic where one target optical response can be produced by multiple structures. A direct inversion of the artificial neural network (ANN) to predict the design usually fails as the training will not converge due to the existence of several possible solutions.<sup>[32]</sup> To circumvent the challenges related to the “one-to-many” mapping, tandem networks architecture is proposed.<sup>[9]</sup> This type of network consists of two sub-networks: an actual inverse network that takes the optical response and outputs the desired design and a pre-trained forward network that evaluates the outputs and generates optical responses. The tandem architecture allows direct comparison in the physics domain and avoids confusion from different designs that lead to a similar optical response. In previous work, we adopted such an architecture for the inverse design of a Fabry-Pérot-cavity-based structural color with high design accuracy.<sup>[33]</sup> Despite the success in tackling the convergence problem using this tandem network, the one-to-many problem leaves us with some critical issues. A prominent limitation of the tandem approach resides in the existence of *inaccessible solutions*<sup>[15]</sup> or a *dead zone*<sup>[34]</sup>, which means that some possible solutions or solution spaces are discarded or unreachable during the network training process. In consequence, only a singular solution can be identified for each target optical response. The collapse of the network onto a set of singular solutions significantly limits the solution diversity and could even prevent reaching an optimal solution as the selected solution is not necessarily the best performing one. Therefore, developing an approach that enables identifying multiple solutions in different local/global optima can be extremely beneficial in achieving higher inverse design accuracy/performance. In addition, it will also provide more freedom in the actual fabrication process to select designs that fits the manufacturing requirement or facility limitation. For instance, a solution with the smaller dimensions could be selected to realize denser device integration while a large-dimension solution could be more suitable when the available facilities

have certain resolution limitations. Furthermore, the performance decay induced by the fabrication error can also be mitigated *via* manufacturing a device with a less parameter-sensitive solution.

Compared with other works that employed the generative models (e.g., generative adversarial networks (GAN) and variational autoencoder (VAE)) to design the 2D metasurface,<sup>[35–39]</sup> we introduced the GAN model<sup>[40]</sup> into the parameter-based 1D nanophotonic inverse design, which provides an approach fully addressing the one-to-many problem. GAN is one of the most popular generative models that has been widely used in computer vision and image synthesis.<sup>[41]</sup> It consists of two networks, a generator and a discriminator, which contest with each other during the training process to create authentic images.<sup>[42]</sup> Due to its unique ability to reveal the hidden distribution behind enormous training datasets, GAN has been used by optical scientists to inversely design metasurfaces and diffractive optical devices with a high degree of freedom.<sup>[22,43–46]</sup> Here, we investigate the potential of applying GAN in identifying multiple solutions for one target optical response. We argue that the geometrical distribution in the training dataset is a key hidden parameter that can be utilized in GAN to avoid *dead zone* and identify multiple solutions.

We conduct an inverse design of a transmissive Fabry-Pérot (F-P) cavity-based structural color filter device using GAN as a demonstration. The structural color filter can display colors by selectively transmitting or reflecting the light with a specific wavelength by varying geometry. It has received enormous interest from research communities owing to its potential applications in display, full-color printing, optical encryption and solar cells.<sup>[47–49]</sup> The reasons for using such a system are threefolds: Firstly, the inverse design of the tri-layer F-P cavity structure (Ag-SiO<sub>2</sub>-Ag in this work) is a typical one-to-many problem as the transmissive peaks would appear periodically at the same wavelength as the increase of resonant cavity length;<sup>[50,51]</sup> Secondly, as a relatively simple structure with the well-established analytical model,<sup>[52]</sup> the performance of the developed GAN can be effectively investigated and evaluated. Thirdly, the design quality can be precisely quantified by the color difference index  $\Delta E_{2000}$ <sup>[53]</sup> (denoted as  $\Delta E$  below), which is calculated from two CIELAB vectors and has a physical meaning in statistics compared to mean squared error (MSE). After training, our GAN produces an average of 3.58 solution groups for each color in the testing set, which covers 93.9% of all ground truths. These multiple solutions allow the selection of the best design for each color which results in the lowest average color difference  $\Delta E$  of 0.44 that has ever been reported for the inverse design of structural color. The identification of multiple solution groups is verified experimentally by fabricating different multilayer structures to demonstrate the same color.

## Results and Discussion

### Problems of Tandem Network

We start by elaborating on the critical *dead zone* problem that the tandem network faces when performing inverse design. Taking the inverse design of a transmissive F-P-cavity-based structural color device as an example, the structure (shown in Figure 1a) has two reflective metal layers (Ag), which are separated by a lossless dielectric (SiO<sub>2</sub>) and stacks on a quartz substrate.

The function of color filter is achieved by multiple round-trip phase delays of electromagnetic waves in the F-P resonator. By tuning the thickness of each layer ( $d_1$ ,  $d_2$ , and  $d_3$ , denoted as a single design vector  $\mathbf{D}$  below), the resonant peak can be controlled to produce different colors. In a typical one-to-many problem, one color, represented by the CIELAB vector  $\mathbf{Lab}$ , could be produced by several different structures  $\mathbf{D}_1$ ,  $\mathbf{D}_2$ , and  $\mathbf{D}_3$ . When this color ( $\mathbf{Lab}$ ) is fed to a tandem network, only one of these solutions will be selected. The other solutions will be discarded to minimize the loss during the training, as illustrated in Figure 1b. This solution degeneracy is due to the inherent contradiction between the one-to-one mapping nature of ANNs and the one-to-many mapping problem.

One potential approach to resolve this contradiction is to introduce a latent vector  $\mathbf{z}$ . The value of this latent vector can be sampled from a specific distribution (e.g., uniform or normal distribution) and assigned to the  $\mathbf{Lab}$  (as shown in Figure 1c). Therefore, the one color can be split into three inputs by the variation of this latent vector and generates three one-to-one mappings that the tandem network can handle. However, these three inputs could still degenerate into very similar solutions within one solution group. The cartoon schematic of such degeneracy is illustrated in Figure 1d where the loss vs  $\mathbf{D}$  landscape is plotted. The three sunken areas represent the three potential solution groups (outlined by blue lines) for this color. Different inputs can still be mapped into one sunken area, resulting in the degeneracy of the solution groups (depicted as the red dots). This behavior will be demonstrated and discussed later in the paper. Therefore, more constraint is required to force the network to search in the entire  $\mathbf{D}$  space so all three solution groups can be identified (shown in Figure 1e). This additional constraint can be realized by generating a dataset distributed across the entire  $\mathbf{D}$  space and forcing the network to learn such distribution. This means that alongside the effort to minimize the color difference, the distribution distance between the true thickness and the generator's predicted thickness should also be minimized during the training process.

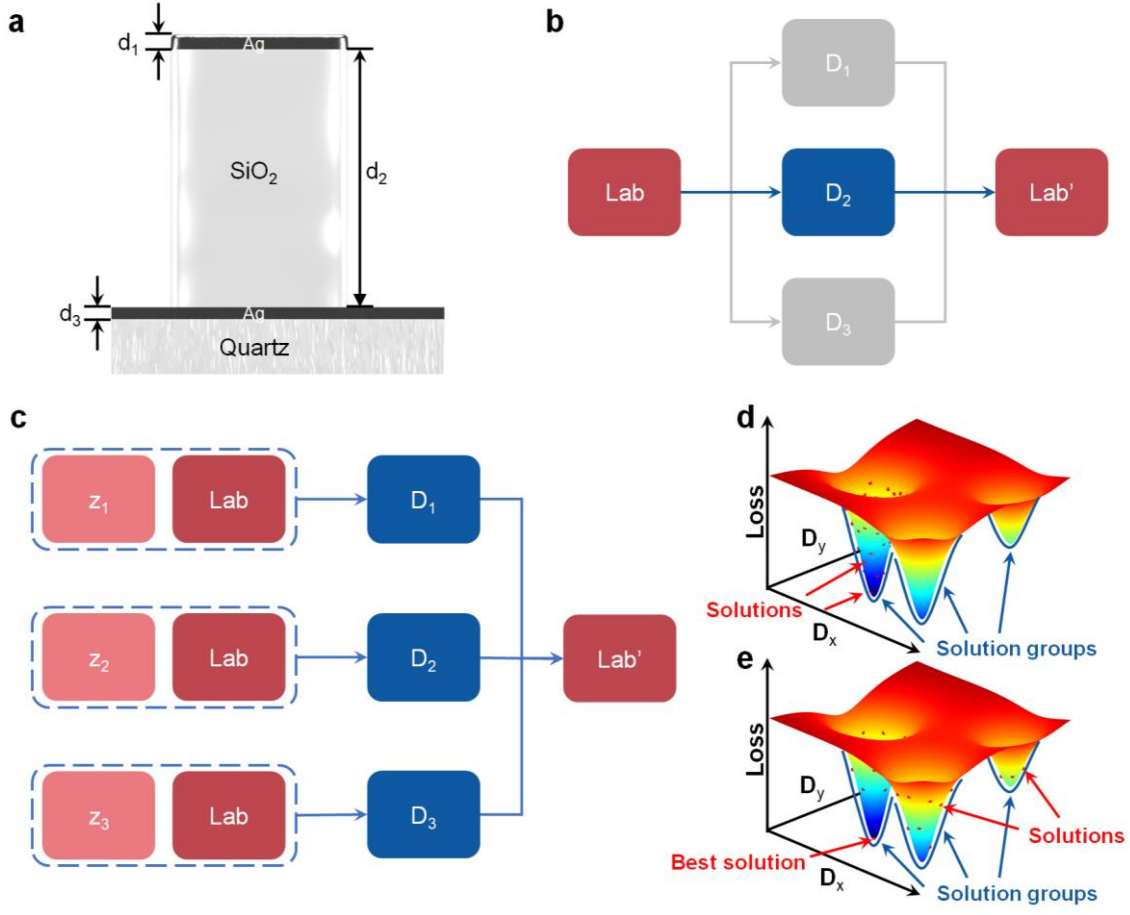


Figure 1. The schematics elaborate on the approach to tackling the one-to-many problem. a) The schematic of the transmissive color filter structure in this demonstration. b) the convergence schematic of the tandem network. c) the schematic of converting the one-to-many problem to three one-to-one problems. d-e) The cartoon loss landscape schematics of loss vs.  $\mathbf{D}$  when the output is degenerated (d), and the problem is well addressed (e).

### cGAN Network Design and Training

The conditional generative adversarial networks<sup>[54]</sup> (cGAN) is a perfect tool for this task. cGAN usually has two parts, a discriminator and a generator. The generator takes a latent vector and the design target (condition) as inputs and outputs the generated data (also known as fake data). The discriminator evaluates the performance and distribution of the fake data and produces losses for backpropagation to shape the fake data close to the real data (training data). Figure 2 illustrates the network architectures used in this paper (details are shown in Figure S1). The generator comprises three sub-networks: Lab upraising layers,  $\mathbf{z}$  upraising layers, and thickness regressor. The upraising layers are responsible for raising the dimensions of **Lab** and  $\mathbf{z}$  from 3 and 2 to 128. The two upraised vectors are concatenated and put into the thickness regressor, which has 256 neurons per hidden layer and a sigmoid function at the end of the output layer, to generate the predicted thicknesses. The discriminator has two parallel networks: evaluator and Lab regressor. The evaluator evaluates the thickness data from the dataset (real) and generator (fake) and outputs two distribution scores. The difference between real and fake scores measures the distribution difference between ground truths and predicted data. The pre-trained Lab regressor takes thickness as input and returns the corresponding color (**Lab**).

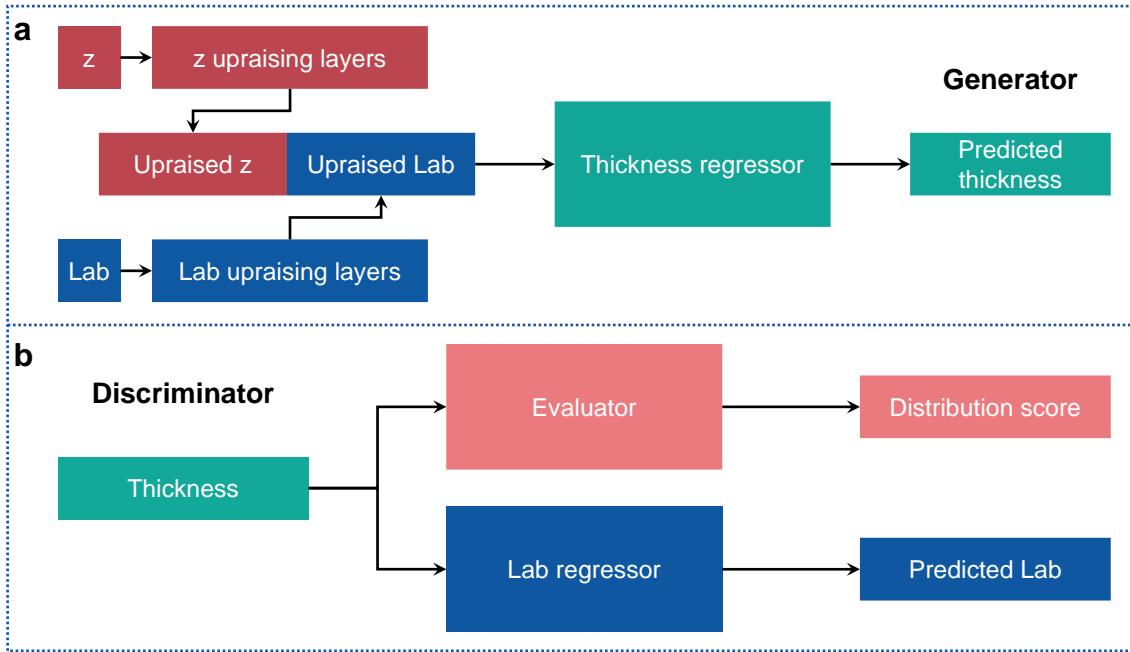


Figure 2. The architectures of the cGAN are used in the work. a) generator. b) discriminator.

As mentioned above, during the training of cGAN, a reference data (real data) distribution is required so the generator could learn the distribution from these data *via* the evaluator. The dataset was generated by randomly sampling 50,000 different combinations of layer thicknesses (**D**) using a uniform distribution. We set the sampling range of 0 to 50 nm for the Ag films and 0 to 1,000 nm for the SiO<sub>2</sub> film. As a thicker Ag layer will significantly reduce the color brightness while too thick SiO<sub>2</sub> can cause many resonances in the visible spectrum and induce color degeneration and crosstalk, we believe that the sampling ranges and distribution are sufficient to cover the requested **D** space for diverse solution groups. The thickness distributions are shown in Figure 3a-3c, in which the thicknesses are uniformly divided into 100 groups, and the bars indicate the probability density of different groups. It is worth mentioning that other thickness distributions may also be applied as the reference distribution, but the uniform one is more straightforward for us due to its accessibility in implementation and excellent color coverage. These structural designs were then turned into transmission spectra *via* the transfer matrix method and converted into HSV (hue, saturation, and value) and CIELAB color models. All generated colors are plotted in the CIE 1931-xy chromaticity diagrams in Figure S2. The discussion of relationships between geometry parameters and the color is provided in section 2 of supplementary materials. The obtained dataset is divided into three groups for training (40,000), validation (5,000), and testing (5,000).

Before training our cGAN model, we firstly trained a Lab regressor to achieve forward modeling of the color **Lab** from a structure **D**. The detailed training process of this Lab regressor is provided in supplementary materials (Table S1 and Figure S4). The trained Lab regressor demonstrates an average color difference ( $\Delta E$ ) of 0.19 from the testing set. The significance of  $\Delta E$  can be classified into five groups: 1)  $\Delta E < 1$ , no discernable color difference; 2)  $1 < \Delta E < 2$ , the difference can be observed by experienced persons; 3)  $2 < \Delta E < 3.5$ , the difference can be observed by inexperienced persons; 4)  $3.5 < \Delta E < 5$ , clear difference can be noticed; 5)  $5 < \Delta E$ , two different colors are observed.<sup>[55]</sup> In the training of cGAN, both the color and thickness distribution

differences are used in the loss function to apply backpropagation. We used Wasserstein distance as the distribution loss and applied spectral normalization to both evaluator and generator to improve the training stability.<sup>[56]</sup> The updating in the training process was alternatively performed on the generator and evaluator to minimize the hinge adversarial loss.<sup>[57–59]</sup> Details of the training process and the definitions of loss functions are provided in the supplementary materials (Table S2 and Equation S1-S3).

The training loss curves of the generator and evaluator are shown in Figure 3d. The loss of generator demonstrates a clear drop to close to zero, indicating the significant improvement of its predicting capability during the training process. The loss of evaluator compares the fake and real score and remains steadily at *ca.* 2 during training. To better understand this behavior, we plot the value of real and fake scores as a function of the training epoch in Figure 3e. It can be observed that the real and fake scores are close, indicating a balanced performance improvement for the generator and evaluator. This property is essential in cGAN training to ensure a balanced development of the generator and evaluator, which will be discussed in detail in the following section. Figure 3f displays the evolution curve of the generator. The Lab loss (red line), which is the MSE between prediction and ground truth of **Lab**, dropped fast in the first ten thousand epochs and then gradually converged to 0.02, meaning that the generator's prediction accuracy improved quickly then converged. The generator's fake score (blue line) in Figure 3f fluctuates around zero, which projects that the generator followed the development of the evaluator and had no model collapse during the training process.

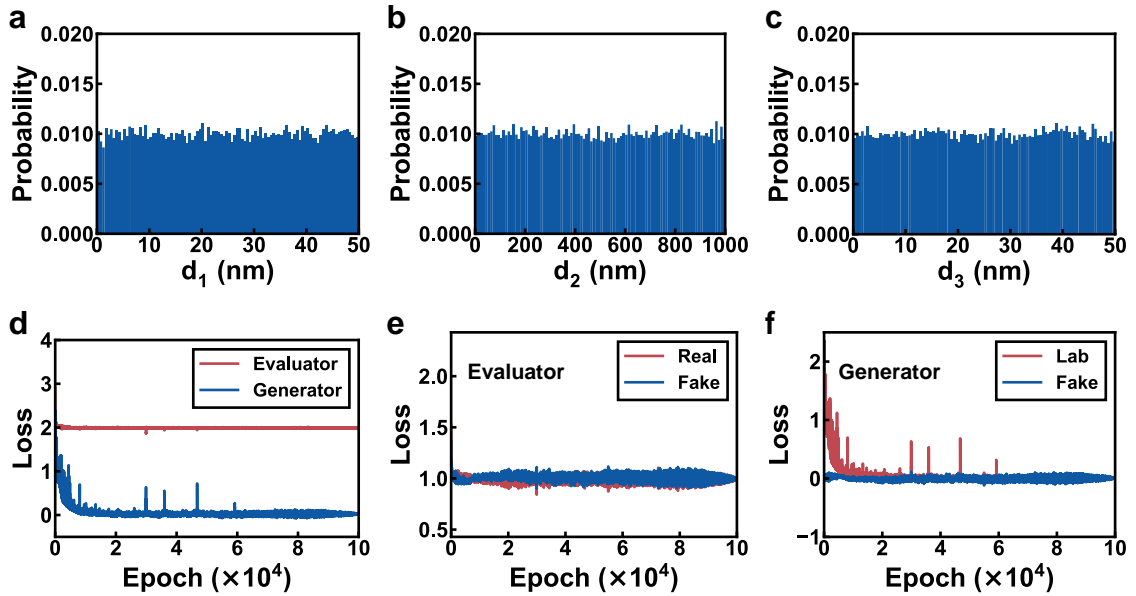


Figure 3. The dataset distribution and cGAN training loss curves. a-c) The thickness histograms of the entire dataset. d) The training loss curves of the generator (blue) and evaluator (red). e) The curves of evaluator loss components including the real (red) and fake (blue) scores. f) The MSE between **Lab** predicted by generator and ground truth (red), and generator's fake score (blue).

### cGAN Performance Evaluation

After training, we used the testing set to evaluate the performance of the generator. The testing set contains 5,000 data with the same uniform distribution as the training set (shown in Figure 4a-

4c). The trained generator takes the 5,000 **Lab** and latent vectors  $\mathbf{z}$  sampled from the standard normal distribution as inputs and predicts 5,000 inverse designed thicknesses (denoted as  $\mathbf{D}_p$ ). The distributions of the three thicknesses in  $\mathbf{D}_p$  are plotted in Figure 4d-4f. To quantify the distribution distance between predicted thickness and ground truth, the Jensen-Shannon distance (JSD) is introduced. The corresponding JSD computation process and reference values are provided in supplementary materials section 4 (Figure S5 and Equation S4-S8). The measured JSDs for our case are 0.069 for  $d_1$ , 0.067 for  $d_2$ , and 0.066 for  $d_3$ , reflecting that the predicted thickness and ground truth have great similarity in distribution. This proves that the trained generator is able to produce designs with restricted thickness distribution.

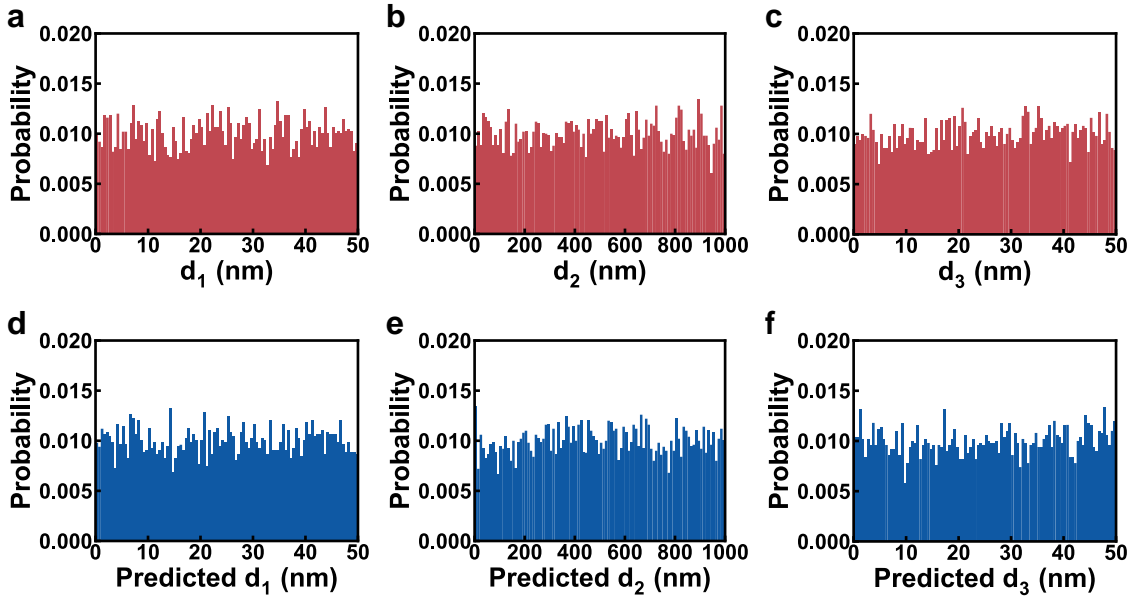


Figure 4. The distribution comparison between ground truth and prediction. The thickness histograms of a-c) the testing set and d-f) predicted by the generator.

Next, we investigated the capability of the generator in achieving diverse solutions. As mentioned above, the latent vector  $\mathbf{z}$  is introduced to generate multiple solutions for each **Lab**. Currently, only one  $\mathbf{z}$  is sampled for each **Lab**, hence only one solution is predicted for each color. In order to obtain multiple solutions, we randomly sampled 1,000 different  $\mathbf{z}$  from the standard normal distribution for each **Lab** and classified the obtained 1,000 solutions into different solution groups by the density-based spatial clustering of applications with noise (DBSCAN), an unsupervised machine learning clustering algorithm. Details of the classification process can be found in section 5 of the supplementary materials. Figure 5a-5b present the distributions of solution group numbers when 100 and 1,000  $\mathbf{z}$  were assigned to each **Lab**, respectively. The aggregation of solution group numbers moves from 1-4 to 2-5 as the  $\mathbf{z}$  number grew from 100 to 1,000, showing a significant solution diversity improvement. An average of 3.58 solution groups can be identified for these 5,000 **Lab** when each target color is assigned with 1,000 different  $\mathbf{z}$ , proving the capability of our network in achieving diverse solution groups. Figure 5c plots the solution group number as a function of different  $\mathbf{z}$  numbers used for each color. It can be observed that the solution group number initially increases sharply with increasing  $\mathbf{z}$  number and eventually saturates when the  $\mathbf{z}$  number reaches over 1,000 (highlighted by the red dot in Figure 5c). A maximum average solution number of 3.66 was obtained when 2,100  $\mathbf{z}$  was used for each color (highlighted by the red triangle



in Figure 5c). The following slight decrease is related to limitations in the clustering algorithm. As the  $z$  number increases, the different solution groups may slightly overlap with each other, which cannot be distinguished by DBSCAN.

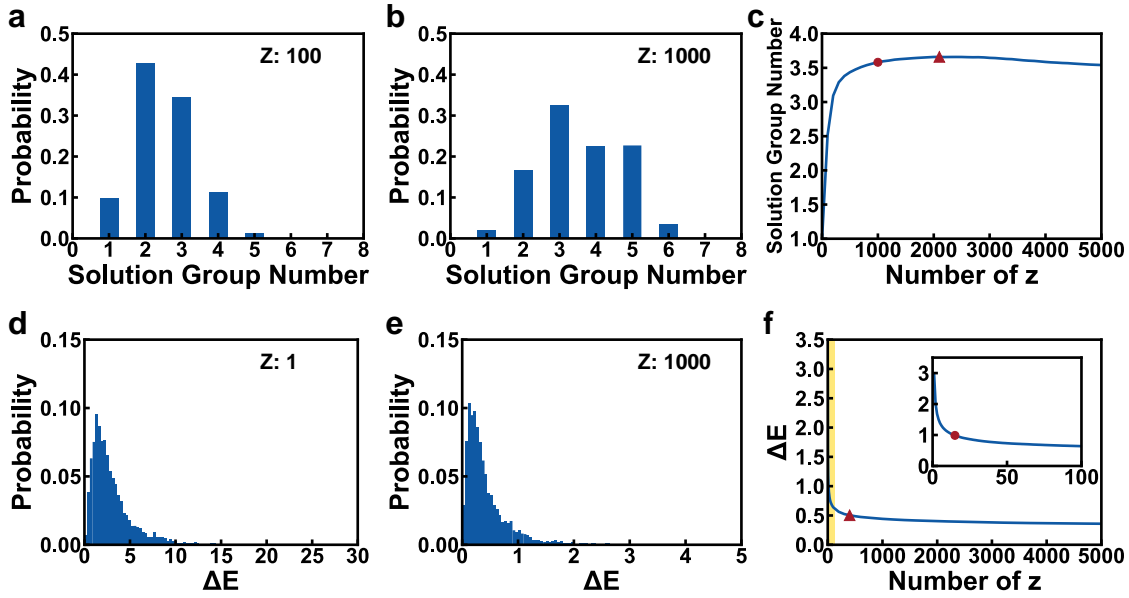


Figure 5. The tendencies of solution group number and  $\Delta E$  as  $z$  sampling number varies. a-b) The solution group number histograms when each **Lab** combines 100 (a) and 1,000  $z$  (b). c) The curve of average testing solution group number against  $z$  number, the circle and triangle markers refer to the solution group numbers when each **Lab** assigned with 1,000 and 2,100  $z$ . d-e) The testing  $\Delta E$  histograms when each **Lab** is assigned with 1 (d) and 1,000 (e)  $z$ . f) The curve of average testing  $\Delta E$  against  $z$  number, in which the insert is the enlarge of yellow shading region, the circle and triangle markers refer to the  $\Delta E$  while each **Lab** assigned with the  $z$  number of 15 and 400.

Concurrently with the increasing number of solution groups, the color difference  $\Delta E$  can also be significantly reduced. Figure 5d plots the  $\Delta E$  distribution of 5,000 solutions when one  $z$  is sampled for each **Lab**. The majority of the  $\Delta E$  is located between 0 to 5, with an average  $\Delta E$  value of 2.98. While 1,000  $z$  are sampled for each **Lab**, the lowest  $\Delta E$  for each color almost falls in 0-1 (can be considered no color difference, as shown in Figure 5e), with an average value of 0.44. To the best of our knowledge, this is the lowest color difference reported for structural color inverse design. The dependence of the  $z$  number on the average lowest  $\Delta E$  is presented in Figure 5f. An average lowest  $\Delta E$  can reach below 1 when the  $z$  number is over 15 (highlighted by a red dot in Figure 5f inset) and below 0.5 when the  $z$  number is over 400 (highlighted by the red triangle in Figure 5f). To elucidate the high design accuracy achieved by our network, we compared the absolute  $d_2$  differences ( $\Delta d_2$ ) between the solutions obtained from each color (using 1,000  $z$ ) and its ground truth (global optimum). We discovered that 93.9% of the  $\Delta d_2$  are within 5 nm, implying the majority of the ground truths have been successfully identified. Detailed distribution information of minimum  $\Delta d_2$  is provided in Figure S6.

All of these results suggest that both the quantity and quality of solutions can be improved using our network by increasing the number of  $z$ . It is worth mentioning that although multiple  $z$  sampling is required to achieve a high solution group number and low  $\Delta E$ , the time required for

such prediction is straightforward and largely effortless. For example, applying inverse design and solution optimization for 5000 colors with 1,000  $z$  sampling for each color will only consume 16.18 s in our computation environment, which is equivalent to 3.24 ms per color.

In order to verify the fabrication applications of multiple solution groups, we applied the generator to inversely design two colors generated by the F-P cavity. Table 1 lists the best solution for each obtained solution group. Both target colors have achieved more than two different solution groups. The ground truths (target designs) have been successfully identified as one of the solutions in both cases. In particular, five solution groups have been identified for color #2 with the  $d_2$  value spanning from 39 to 985 nm, proving the powerful searching ability of our network. These multiple solution groups allow the selection of suitable designs for the manufacturing requirement. Films that are too thin can be challenging to fabricate with high quality. For example, a design with a very thin Ag layer (e.g., 3 nm in design 3 for color #2) could increase the fabrication relative error and might need to be discarded despite its low  $\Delta E$ . Instead, a different design (e.g. design 4) could be selected due to its relatively thicker films but still competitive  $\Delta E$ . The sensitivity of the solution to fabrication errors is also an important factor to consider. We calculated the sensitivity here as the  $\Delta E$  when  $d_2$  was added or subtracted by 1 nm (denoted as  $\delta_1$  and  $\delta_{-1}$ , respectively). Higher  $\delta_{\pm 1}$  means that the solution is more sensitive to fabrication error and requires a more precise fabrication process. Although the ground truth of color #1 in Table 1 has been successfully identified in design 1, the sizeable  $\delta_{\pm 1}$  inflicts extreme requirements on the fabrication accuracy. On the other hand, design 3 with a slightly larger  $\Delta E$  but the smallest  $\delta_{\pm 1}$  would be more suitable for the actual application. The capability of our network in identifying multiple solution groups for each color is clearly advantageous for providing viable designs for fabrication.

Table 1 The inverse design demonstrations of two colors. The target items are the ground truths. The designed items are from the solutions of the generator.

Target D ( $d_1$ - $d_2$ - $d_3$ ) nm	Target Color (sRGB)	Designed D ( $d_1$ - $d_2$ - $d_3$ ) nm	Designed Color (sRGB)	Comparison (Target, Design)		$\Delta E$	( $\delta_1, \delta_{-1}$ )
33-118-27	(-1.549)-0.534-0.426	34-118-26	(-1.481)-0.532-0.421			0.45	(2.03, 2.22)
		16-284-32	(-1.475)-0.536-0.420			0.73	(1.03, 1.08)
		25-460-19	(-1.544)-0.528-0.417			0.69	(0.60, 0.63)
3-550-41	0.302-0.274-0.453	21-39-18	0.193-0.268-0.470			7.65	(0.19, 0.19)
		7-183-35	0.298-0.262-0.451			1.08	(0.44, 0.45)
		3-550-41	0.302-0.275-0.453			0.12	(0.30, 0.29)
		33-722-28	0.296-0.273-0.446			0.53	(0.42, 0.42)
		28-985-30	0.305-0.272-0.443			0.75	(0.27, 0.27)

We have proved from the discussion above that the proposed cGAN architecture can completely

solve the “one-to-many” issue for this nanophotonics question. We argue that the vital component in this architecture is the evaluator, which enables the generator to learn the distribution of the dataset. Figure S8 presents the performance of the same networks with the evaluator disabled. The distributions of the three thicknesses designed for the testing set show no uniformity (Figure S8a-S8c). The measured JSDs between these designed thicknesses and the testing set are 0.21 for  $d_1$ , 0.64 for  $d_2$  and 0.14 for  $d_3$ , respectively. More importantly, with 1,000  $\mathbf{z}$  sampled for each color, the network failed to identify multiple solution groups. Instead, the distribution of  $d_2$  is centralized between the range of 800 to 1000 nm, which leads to a relatively sizeable  $\Delta E$  distribution (Figure S8d) with an average  $\Delta E$  of 7.88 and a limited color coverage (Figure S8e). This proves that the evaluator plays a vital role in the solution quality and diversity of the generator. Finally, it is also worth pointing out that the balance between the generator and evaluator is also essential for the network performance (Figure S9-S11). A stronger discriminator will lead to a divergent generator, while a weaker discriminator will slow the training. Details are discussed in section 4 of supplementary materials.

### Experimental Validation of Designed sRGB Color Filters

To further evaluate the network generalization and search performance of our network, we performed the inverse design of three color filters with sRGB values of red (0.5, 0, 0), green (0, 0.5, 0), and blue (0, 0, 0.5) using our trained generator with 1,000  $\mathbf{z}$  sampling for each color. Note that these three colors are not presented in our dataset and hence have no ground truth. These examples could test the capability of our cCAN in searching multiply solutions for colors outside the dataset. To better visualize the searchability, we fixed the thicknesses of the top and bottom Ag layers to be 30 nm (this Ag restriction is removed in the cGAN inverse design) while sweeping the thickness of the  $\text{SiO}_2$  layer ( $d_2$ ) from 0 to 1,000 nm to generate the loss (MSE) curves for each color (shown in Figure 6a-6c). The loss between the target and swept colors varies as  $d_2$  changes. A small loss represents a better solution, and a dip towards 0 in the loss curve indicates one locally optimal solution group for the target. Therefore, the solution group number of each target color can be estimated from the number of curve dips towards 0. In this case, we could identify two solution groups for both the red and green colors, and three solution groups for the blue color. In addition, all the solution groups correspond to the different resonant orders for each color. The two solution groups at *ca.* 160 nm and 350 nm correspond to the 1<sup>st</sup> and 2<sup>nd</sup> order resonances for red (Figure 6a); the *ca.* 130 nm and 320 nm designs are the 1<sup>st</sup> and 2<sup>nd</sup> order resonances for green (Figure 6b); and the *ca.* 100 nm, 250 nm, and 410 nm designs are the first three resonant orders for blue (Figure 6c). The inverse design results of these three colors are demonstrated in Figure 6d-6f. Here, we plot the distributions of the 1,000 predicted  $d_2$  values for the three colors, respectively, in which  $d_2$  is uniformly divided into 100 groups in the range of 0-1000 nm and a bar indicates the solution number of each thickness group. Our network has successfully produced multiple solution groups for each color, matching the number and value of the solution groups identified from the loss curves. This suggests that our network can identify all possible solutions for different target colors. The relationships between the value of  $\mathbf{z}$  and solution groups are discussed in supplementary materials (Figure S7).

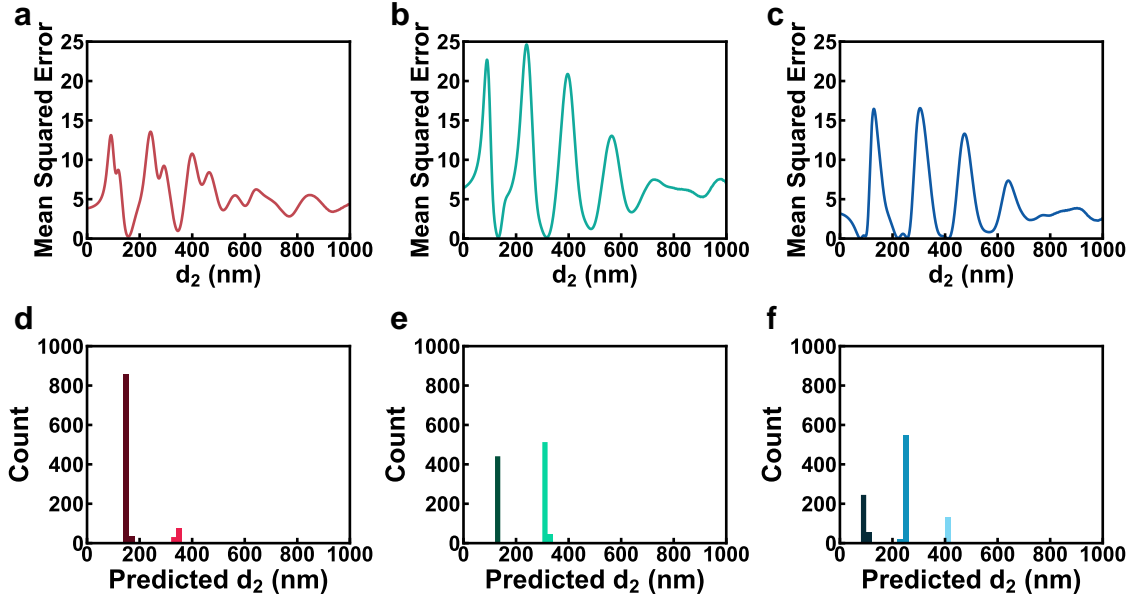


Figure 6. The analysis of sRGB color filter design results. a-c) The MSE curves of the colors with sRGB values of (a) (0.5, 0, 0 red), (b) (0, 0.5, 0 green) and (c) (0, 0, 0.5 blue) when the SiO<sub>2</sub> thickness ( $d_2$ ) was swept from 0 to 1000 nm and Ag thicknesses ( $d_1$  and  $d_3$ ) were fixed at 30 nm. The DBSCAN clustered predicted  $d_2$  histograms for the (d) red, (e) green and (f) blue colors. The dark-colored bar indicates the lower resonant order, and the light-colored bar means the higher order.

We will now experimentally validate our multiple solution groups by fabricating the red, green, and blue color filters previously designed by our cGAN. The parameters of the best solution in each solution group for each color are listed in Table 2. All the designed color filters were fabricated on quartz substrates. Details of the fabrication process are available in the method section. Figure 7a-7c display the cross-sectional SEM of all fabricated samples. A three-layered structure containing Ag (blue shading) top and bottom layers and SiO<sub>2</sub> dielectric is clearly visible for all samples. It is also evident that the dimensions (especially the thickness of the SiO<sub>2</sub> layer) are significantly different for different designs of the same target color. The measured transmission spectra for all samples are compared with their theoretical (calculated from designed parameters) counterparts and shown in Figure 7d-7f. All measured spectra match well with the theoretical ones in terms of both transmission peak wavelength and intensity. The discrepancies in peak intensity at shorter wavelengths may be attributed to the surface oxidization of the Ag films.<sup>[60,61]</sup> Figure 7g-7i display the resultant colors for each design. For each target color (TGT.), we compare the theoretical colors from different designs (DSG.) with the colors converted from corresponding experimental spectra (EXPT. SPT.) and the photographs of the samples taken by the camera under a sunlight-like source (EXPT. PHT.). The discrepancy in the R2 photograph may result from the light source intensity or the camera parameters (e.g., white balance, contrast, exposure time, etc.).

Table 2. The targets and parameters of designed sRGB color filters.

Sample Name	Target Color (sRGB)	Designed D ( $d_1$ - $d_2$ - $d_3$ ) nm	Designed Lab	Designed sRGB	$\Delta E$	( $\delta_1, \delta_{-1}$ )
R1	0.5-0-0	37-158-38	26.2-34.7-26.7	0.449-0.134-0.090	5.20	(1.30, 1.38)
R2	0.5-0-0	37-346-32	29.9-17.4-14.4	0.399-0.233-0.192	13.81	(2.09, 2.22)
G1	0-0.5-0	33-134-36	46.7-(-45.2)-41.4	0.141-0.499-0.130	2.77	(1.74, 1.64)
G2	0-0.5-0	28-318-31	45.9-(-46.7)-37.5	0.076-0.494-0.158	3.92	(0.86, 0.83)
B1	0-0-0.5	30-99-49	13.9-47.3-(-65.6)	(-0.011)-0.012-0.517	0.79	(1.67, 1.44)
B2	0-0-0.5	40-256-31	12.4-49.0-(-65.8)	(-0.014)-(-0.013)-0.503	0.57	(1.02, 0.94)
B3	0-0-0.5	36-414-30	13.8-43.9-(-59.6)	0.080-0.024-0.480	1.48	(0.79, 0.74)

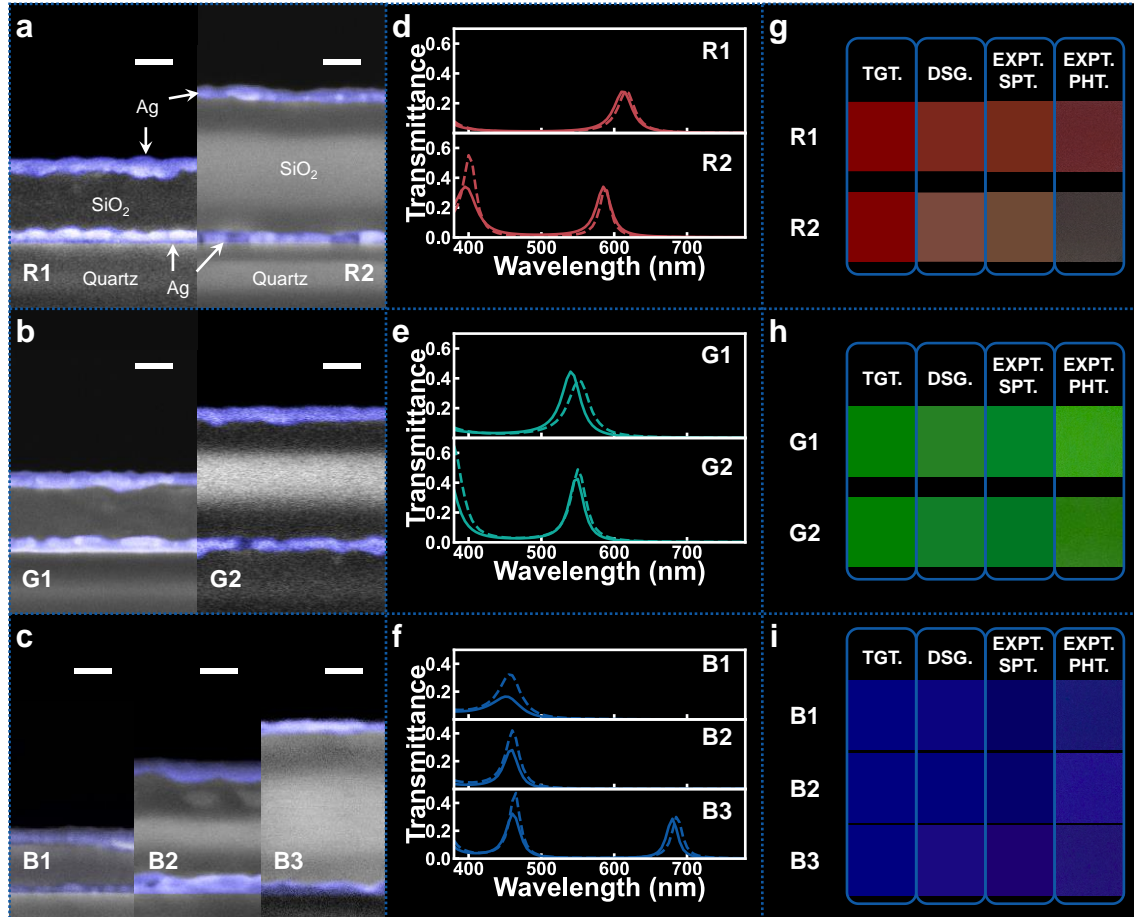


Figure 7. The experimental results of fabricated sRGB color filters. a-c) The cross-sectional SEM images of the fabricated color filters with all the scale bars of 100 nm. d-f) The measured spectra (solid line) and corresponding theoretical ones (dash line). g-i) The color reconstructions, wherein TGT. is the target color; DSG. is the theoretical color of the designed color filter; EXPT. SPT. is the theoretical color of the measured spectrum and EXPT. PHT. is the photograph of the sample taken by the camera.

## Conclusion

In conclusion, a variation of cGAN has been successfully implemented to resolve the typical one-to-many problem in nanophotonics using the inverse design of the F-P-cavity-based color filter as a demonstration. We suggest that the key to avoiding solution group degeneracy is generating solutions with the same parameter distribution of a dataset covering most solution spaces. Equipped with an evaluator that enables learning the dataset distribution, the cGAN developed in this work is able to identify an average of 3.58 solution groups for each color. Simultaneous with the improved solution diversity, the quality of the solutions is also significantly improved. An average  $\Delta E$  of 0.44 has been achieved when sampling 1,000  $\mathbf{z}$  for each color. The high design accuracy could be attributed to identifying 93.9% ground truths of the testing set. Finding multiple solution groups can also benefit design manufacturing to allow more viable designs for fabrication. The capability of our cGAN has also been verified experimentally by inverse designing the sRGB color filters. The measured spectra and photographs of fabricated two red, two green, and three blue color filters present excellent agreement with the designs. It should be acknowledged that the presented work is limited by its dataset range as well as its simple 3-layered structure. However, it is worth highlighting that the 3-layer F-P cavity structure only serves as an example to demonstrate the capability of cGAN and the novelties of this work remain valid. We envisage this cGAN-based design methodology to be applied to other nanophotonic structures or physical science domains where the identification of multi-solution across a vast parameter space is required.

## Method

**Fabrication:** Before fabricating the device, a series of experiments were conducted to identify the deposition rate for both Ag and SiO<sub>2</sub> using both ellipsometry (Woollam M-2000DI) and atomic force microscopy (Park System X7) for calibration. The same recipes were then used to fabricate our designed structural colors to ensure the best experimental reproduction of the designed structures. During the fabrication of structural colors, the bottom Ag thin film was firstly deposited by the electron-beam evaporation system (Leybold LAB700) with a growth rate of 0.5 Å/s. Without breaking the vacuum, a 1 nm thin Ti film was formed by electron-beam evaporation with a growth rate of 0.1 Å/s on top of the Ag film to improve the adhesion between Ag and the following SiO<sub>2</sub>. The SiO<sub>2</sub> layer was subsequently formed on the Ti layer in a reactive sputtering system (Leybold Helios Pro XL) using Si target and O<sub>2</sub> plasma. The growth rate of SiO<sub>2</sub> is 0.52 nm/s. After completing the SiO<sub>2</sub> deposition, the top Ag film was formed with the same process condition as the bottom Ag film and then a 1 nm Ti film on top.

**Characterization:** The refractive index of SiO<sub>2</sub> used in the calculation is measured by an ellipsometer (Woollam M-2000DI). The transmission spectra under normal incidence were measured by a spectrometer (Bentham PVE300) with Air as reference. The photographs were taken using a digital camera (Canon EOS 500D) with a shutter speed of 1/20 s, an aperture of f/5, and the default white balance. In order to produce the source with the correct color temperature,

an Apple iPad Pro (10.5-inch, 2017 year version) was used as a backlight source, with both true tone and night shift switched off and brightness set as maximum. The original photograph is shown in Figure S13. The cross-sectional morphologies of fabricated multiple thin-film structures were characterized using a field-emission scanning electron microscope (JEOL 7500 FEGSEM) at an acceleration voltage of 2 kV and a working distance of 4 mm.

**Calculation:** All the calculations and neural network training were performed on a Windows PC with the hardware: CPU: Intel Core i9-9900K; RAM: 48 GB; GPU: NVIDIA RTX 2070. The involved software packages include Python 3.8, PyTorch 1.9.0, NumPy, Scikit-Learn and Colour-Science. The refractive index of Ag used in the calculation is Ag (Palik 0-2  $\mu\text{m}$ ).<sup>[62]</sup>

## Reference

- [1] W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, Y. Liu, *Nat. Photonics* **2021**, *15*, 77.
- [2] R. Yan, T. Wang, X. Jiang, Q. Zhong, X. Huang, L. Wang, X. Yue, *Nanotechnology* **2020**, *31*, 375202.
- [3] J. C. C. Mak, C. Sideris, J. Jeong, A. Hajimiri, J. K. S. Poon, *Opt. Lett.* **2016**, *41*, 3868.
- [4] B. J. Offrein, G.-L. Bona, R. Germann, I. Massarek, D. Erni, *J. Light. Technol.* **1998**, *16*, 1680.
- [5] J. A. Bossard, L. Lin, S. Yun, L. Liu, D. H. Werner, T. S. Mayer, *ACS Nano* **2014**, *8*, 1517.
- [6] S. So, Y. Yang, S. Son, D. Lee, D. Chae, H. Lee, J. Rho, *Nanophotonics* **2021**, DOI: 10.1515/nanoph-2021-0436.
- [7] J. S. Jensen, O. Sigmund, *Laser Photonics Rev.* **2011**, *5*, 308.
- [8] A. Michaels, M. C. Wu, E. Yablonovitch, *IEEE J. Sel. Top. Quantum Electron.* **2020**, *26*, 1.
- [9] D. Liu, Y. Tan, E. Khoram, Z. Yu, *ACS Photonics* **2018**, *5*, 1365.
- [10] M. M. Elsaywy, S. Lanteri, R. Duvigneau, J. A. Fan, P. Genevet, *Laser Photonics Rev.* **2020**, *14*, 1900445.
- [11] S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, A. W. Rodriguez, *Nat. Photonics* **2018**, *12*, 659.
- [12] S. D. Campbell, D. Sell, R. P. Jenkins, E. B. Whiting, J. A. Fan, D. H. Werner, *Opt. Mater. Express* **2019**, *9*, 1842.
- [13] S. So, T. Badloe, J. Noh, J. Bravo-Abad, J. Rho, *Nanophotonics* **2020**, *9*, 1041.
- [14] J. Jiang, M. Chen, J. A. Fan, *Nat. Rev. Mater.* **2020**, *1*.
- [15] P. R. Wiecha, A. Arbouet, C. Girard, O. L. Muskens, *Photonics Res.* **2021**, *9*, B182.
- [16] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*, MIT Press, **2016**.
- [17] Y. LeCun, Y. Bengio, G. Hinton, *Nature* **2015**, *521*, 436.
- [18] W. Ma, F. Cheng, Y. Liu, *ACS Nano* **2018**, *12*, 6326.
- [19] Y. Li, Y. Xu, M. Jiang, B. Li, T. Han, C. Chi, F. Lin, B. Shen, X. Zhu, L. Lai, *Phys. Rev. Lett.* **2019**, *123*, 213902.
- [20] M. Chen, J. Jiang, J. A. Fan, *ACS Photonics* **2020**, *7*, 3141.
- [21] J. Jiang, J. A. Fan, *Nanophotonics* **2020**, *9*, 1059.
- [22] J. Jiang, D. Sell, S. Hoyer, J. Hickey, J. Yang, J. A. Fan, *ACS Nano* **2019**, *13*, 8872.
- [23] Z. Huang, X. Liu, J. Zang, *Nanoscale* **2019**, *11*, 21748.

- [24] O. Hemmatyar, S. Abdollahramezani, Y. Kiarashinejad, M. Zandehshahvar, A. Adibi, *Nanoscale* **2019**, *11*, 21266.
- [25] L. Gao, X. Li, D. Liu, L. Wang, Z. Yu, *Adv. Mater.* **2019**, *31*, 1905467.
- [26] J. Baxter, A. C. Lesina, J.-M. Guay, A. Weck, P. Berini, L. Ramunno, *Sci. Rep.* **2019**, *9*, 1.
- [27] N. B. Roberts, M. Keshavarz Hedayati, *Appl. Phys. Lett.* **2021**, *119*, 061101.
- [28] X. Xu, C. Sun, Y. Li, J. Zhao, J. Han, W. Huang, *Opt. Commun.* **2021**, *481*, 126513.
- [29] R. Unni, K. Yao, X. Han, M. Zhou, Y. Zheng, *Nanophotonics* **2021**.
- [30] Z. Zhen, C. Qian, Y. Jia, Z. Fan, R. Hao, T. Cai, B. Zheng, H. Chen, E. Li, *Photonics Res.* **2021**, *9*, B229.
- [31] Z. Hou, P. Zhang, M. Ge, J. Li, T. Tang, J. Shen, C. Li, *Nanomaterials* **2021**, *11*, 2672.
- [32] H. Kabir, Y. Wang, M. Yu, Q.-J. Zhang, *IEEE Trans. Microw. Theory Tech.* **2008**, *56*, 867.
- [33] P. Dai, Y. Wang, Y. Hu, C. de Groot, O. Muskens, H. Duan, R. Huang, *Photonics Res.* **2021**, *9*, B236.
- [34] R. S. Hegde, *Nanoscale Adv.* **2020**, *2*, 1007.
- [35] W. Ma, Y. Xu, B. Xiong, L. Deng, R.-W. Peng, M. Wang, Y. Liu, *Adv. Mater.* **2022**, *n/a*, 2110022.
- [36] D. Zhu, Z. Liu, L. Raju, A. S. Kim, W. Cai, *ACS Nano* **2021**, *15*, 2318.
- [37] Z. A. Kudyshev, A. V. Kildishev, V. M. Shalaev, A. Boltasseva, *Nanophotonics* **2021**, *10*, 371.
- [38] W. Ma, Y. Liu, *Sci. China Phys. Mech. Astron.* **2020**, *63*, 284212.
- [39] W. Ma, F. Cheng, Y. Xu, Q. Wen, Y. Liu, *Adv. Mater.* **2019**, *31*, 1901111.
- [40] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, *Commun. ACM* **2020**, *63*, 139.
- [41] L. Gonog, Y. Zhou, in *2019 14th IEEE Conf. Ind. Electron. Appl. ICIEA*, **2019**, pp. 505–510.
- [42] J. Gui, Z. Sun, Y. Wen, D. Tao, J. Ye, *ArXiv Prepr. ArXiv200106937* **2020**.
- [43] Z. Liu, D. Zhu, S. P. Rodrigues, K.-T. Lee, W. Cai, *Nano Lett.* **2018**, *18*, 6570.
- [44] S. An, B. Zheng, H. Tang, M. Y. Shalaginov, L. Zhou, H. Li, M. Kang, K. A. Richardson, T. Gu, J. Hu, *Adv. Opt. Mater.* **2021**, *9*, 2001433.
- [45] S. So, J. Rho, *Nanophotonics* **2019**, *8*, 1255.
- [46] F. Wen, J. Jiang, J. A. Fan, *ACS Photonics* **2020**, *7*, 2098.
- [47] Z. Xuan, J. Li, Q. Liu, F. Yi, S. Wang, W. Lu, *The Innovation* **2021**, *2*, 100081.
- [48] Y. Zhao, Y. Zhao, S. Hu, J. Lv, Y. Ying, G. Gervinskias, G. Si, *Materials* **2017**, *10*, DOI: 10.3390/ma10080944.
- [49] F. Chen, S.-W. Wang, X. Liu, R. Ji, Z. Li, X. Chen, Y. Chen, W. Lu, *Opt. Express* **2016**, *24*, A92.
- [50] Z. Yang, Y. Zhou, Y. Chen, Y. Wang, P. Dai, Z. Zhang, H. Duan, *Adv. Opt. Mater.* **2016**, *4*, 1196.
- [51] Y. Wang, M. Zheng, Q. Ruan, Y. Zhou, Y. Chen, P. Dai, Z. Yang, Z. Lin, Y. Long, Y. Li, *Research* **2018**, *2018*.
- [52] C. C. Katsidis, D. I. Siapkas, *Appl. Opt.* **2002**, *41*, 3978.
- [53] G. Sharma, W. Wu, E. N. Dalal, *Color Res. Appl.* **2005**, *30*, 21.
- [54] M. Mirza, S. Osindero, *ArXiv Prepr. ArXiv14111784* **2014**.
- [55] W. Mokrzycki, M. Tatol, *Mach Graph Vis* **2011**, *20*, 383.



- [56] H. Zhang, I. Goodfellow, D. Metaxas, A. Odena, in *Proc. 36th Int. Conf. Mach. Learn.* (Eds.: K. Chaudhuri, R. Salakhutdinov), PMLR, **2019**, pp. 7354–7363.
- [57] J. H. Lim, J. C. Ye, *ArXiv Prepr. ArXiv170502894* **2017**.
- [58] D. Tran, R. Ranganath, D. M. Blei, *ArXiv Prepr. ArXiv170208896* **2017**.
- [59] T. Miyato, M. Koyama, *ArXiv Prepr. ArXiv180205637* **2018**.
- [60] J.-H. Qiu, P. Zhou, X.-Y. Gao, J.-N. Yu, S.-Y. Wang, J. Li, Y.-X. Zheng, Y.-M. Yang, Q.-H. Song, L.-Y. Chen, *J Korean Phys Soc* **2005**, *46*, 269.
- [61] A. C. Nwanya, P. E. Ugwuoke, B. A. Ezekoye, R. U. Osuji, F. I. Ezema, *Adv. Mater. Sci. Eng.* **2013**, *2013*, 450820.
- [62] E. D. Palik, *Handbook of Optical Constants of Solids*, Academic Press, **1998**.

## Acknowledgments

We acknowledge the support of the International Exchange Scheme (IEC/NSFC/170193) between Royal Society (UK) and the National Natural Science Foundation of China (China). This work is a part of the ADEPT project funded by a program grant from the EPSRC (EP/N035437/1). All data supporting this study are openly available from the University of Southampton repository at DOI: <https://doi.org/10.5258/SOTON/D2182>.

## Author contributions

P.D. and R.H. conceived the project and constructed the research frame. P.D. developed the computer code to run the machine learning algorithms. K.S. and X.Y. performed the experimental fabrications and characterizations. O.L.M. and C.H.D.G helped to interpret the results and improve the research frame. Y.H., H.G. and X.Z. provided constructive suggestions to improve the significance of this project. R.H. supervised and directed the project. P.D. and R.H. wrote the manuscript with inputs from all the authors. All authors attended the result discussion and gave comments and advice on the manuscripts.

## Competing interests

The authors declare no conflicts of interest.