PHOTONICS Research

FreeformNet: fast and automatic generation of multiple-solution freeform imaging systems enabled by deep learning

Boyu Mao,[1] **Tong Yang,**[1,2,*] **Huiming Xu,**[1] **Wenchen Chen,**[1] **Dewen Cheng,**[1,3] (iD) **and Yongtian Wang**[1,2]

[1]*Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China*
[2]*Beijing Key Laboratory of Advanced Optical Remote Sensing Technology, School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China*
[3]*e-mail: cdwlxk@bit.edu.cn*
Corresponding author: yangtong@bit.edu.cn

Using freeform optical surfaces in lens design can lead to much higher system specifications and performance while significantly reducing volume and weight. However, because of the complexity of freeform surfaces, freeform optical design using traditional methods requires extensive human effort and sufficient design experience, while other design methods have limitations in design efficiency, simplicity, and versatility. Deep learning can solve these issues by summarizing design knowledge and applying it to design tasks with different system and structure parameters. We propose a deep-learning framework for designing freeform imaging systems. We generate the data set automatically using a combined sequential and random system evolution method. We combine supervised learning and unsupervised learning to train the network so that it has good generalization ability for a wide range of system and structure parameter values. The generated network FreeformNet enables fast generation (less than 0.003 s per system) of multiple-solution systems after we input the design requirements, including the system and structure parameters. We can filter and sort solutions based on a given criterion and use them as good starting points for quick final optimization (several seconds for systems with small or moderate field-of-view in general). The proposed framework presents a revolutionary approach to the lens design of freeform or generalized imaging systems, thus significantly reducing the time and effort expended on optical design. © 2023 Chinese Laser Press

1. INTRODUCTION

Optical design and imaging optics play an important role in technological and social development. In their long history, imaging systems have mainly consisted of spherical and aspherical elements because of their rotational shape and ease of fabrication; however, their aberration correction ability is limited, particularly in nonsymmetric systems. To overcome the limitations of traditional spherical and aspherical systems, non-rotational symmetric freeform optical surfaces can be used, which improve system performance and specifications while reducing the volume and number of elements. The use of the freeform optical surface is considered to be a revolution in imaging optical design [1,2]. In the last 15 years, the development of advanced fabrication technologies has promoted the use of freeform optics in many fields, such as astronomical telescopes [3], head-mounted and head-up displays [4–8], cameras [9,10], off-axis imagers [11,12], and imaging spectrometers [13–15].

Advanced freeform surfaces may improve the performance of imaging optics but also further increase design difficulty and time cost significantly because of the complexity of the surface shape and nonsymmetric system structure, in addition to the limitation of the scarcity of existing reference systems and the difficulty of understanding freeform optics. Traditional optical design or lens design methods usually start with finding a proper starting point and then performing multiparameter optimization. The starting point is typically obtained by searching literature or accessing the lens databases within optical design software. However, this process takes a lot of time and may not find a feasible starting point, especially for the freeform optical systems. Without a good starting point, the design process will rely on extensive human effort and design skills, and extensive time may be spent on tedious trial and error, especially for beginners in optical design with limited or no knowledge of aberration theory or other design experiences. Nodal aberration theory has been used to guide the design and optimization

of freeform imaging systems [16–18]. Direct or point-by-point design methods have been proposed to construct systems based on given design requirements [19–25]. However, these methods also have limitations on the design efficiency, simplicity, and generality, especially for a system with advanced system specifications such as wide field-of-view. The above-mentioned design methods need to be tailored to a certain design task. For other design tasks, the methods may need to be reapplied, and even the optimization strategy may need to be adjusted accordingly. In addition, the time cost for a single design task is high. Deep learning (DL) can be considered as a solution to these issues, as it can effectively summarize design knowledge and apply this knowledge to design tasks with a wide range of system and structure parameters. In 2019, Côté _et al._ used DL to obtain lens design databases to produce high-quality starting points for coaxial spherical objectives [26]. Then, this was improved by the introduction of more design forms [27,28]. However, the above framework is limited to coaxial spherical systems. In 2019, Yang _et al._ [29] proposed the preliminary design framework of freeform reflective imaging systems, which Chen _et al._ [30] improved by increasing the range of system specifications. However, the design method is limited because only one solution can be generated, which may not be optimal to fulfill the design requirements for system structures. It is of great significance while challenging to realize ultrafast generation of multiple-solution freeform imaging systems based on given design requirements on system and structure parameters, which can dramatically improve optical design efficiency and reduce human effort.

In this study, we propose a deep-learning framework for the intelligent design of freeform imaging optics. The design framework is viable for generalized off-axis reflective, refractive, and catadioptric systems with multiple freeform surfaces. We generate the training data set automatically using a combined sequential and random freeform system evolution method. We further propose a special feedback strategy and use it to improve

diversity of the systems in the data set. We also combine supervised learning and unsupervised learning based on freeform surface differential ray tracing to obtain a neural network with high performance. The generated network, which we call FreeformNet, enables the fast generation (less than 0.003 s per system) of multiple-solution freeform imaging systems after we input the design requirements that include the system and structure parameters. We can filter and sort the output systems based on a given criterion and use them as good starting points for the quick final optimization (several seconds for systems with small or moderate field-of-view in general).

We used the design of a freeform off-axis three-mirror imaging system with wide range and advanced system specifications to demonstrate the effect of FreeformNet. The efficiency of the freeform optical design improved significantly and human effort was minimized. The proposed framework provides a new and generalized approach for complicated imaging optical design, thus facilitating the development of revolutionary and generalized optical design software.

2. METHODS

A. Overall Design Framework

The design framework uses a deep neural network (DNN) to generate the freeform systems. The DNN has multiple hidden layers between the input and output layers, which can achieve a complex mapping relationship from the input space to the output space. For the starting point generation task of freeform imaging optical systems, the inputs are specific system parameters and structure parameters, and the outputs are one or more freeform surface imaging systems that meet the design requirements, which can be used for subsequent optimization. This can be considered as design knowledge, as the network "knows" the systems corresponding to the design inputs. These knowledges are obtained through combined supervised and unsupervised learning during DNN training. Figure 1 shows the entire
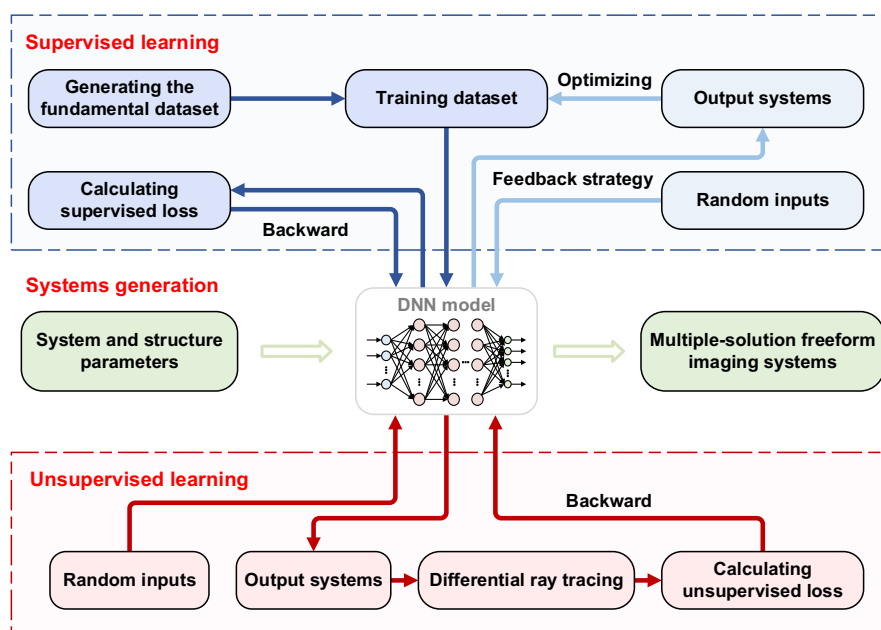

Fig. 1. Whole optical design framework based on deep learning.

design framework. For supervised learning, a considerable number of freeform systems with different system and structure parameters, in addition to good imaging performance, are required as the training data set. A combined sequential and random freeform system evolution method is proposed and used to automatically generate the fundamental data set for the pretraining of the DNN in supervised learning mode. A feedback strategy is further used to enrich the data set and improve the supervised training result of the DNN. Unsupervised learning based on differential ray tracing is then introduced. Combinations of various system and structure parameters are input into the DNN to obtain the output systems, and their imaging performance and constraints are integrated into the total loss function. Supervised and unsupervised trainings are combined and cooperate to obtain the final DNN with good performance and generalization ability. The use of a feedback strategy and unsupervised learning reduce the pressure of obtaining extensive systems for constructing data sets. When network training is completed, the DNN can be considered to have enough knowledge to output systems: one or more systems can be output quickly and directly according to certain design requirements, and the systems can be regarded as good design starting points for quick further optimization.

B. Fundamental Data Set Generation

A freeform imaging optical system can be characterized by system parameters, structure parameters, and surface parameters. System parameters describe the field-of-view (FOV), aperture value, and imaging size. Structure parameters include the position and tilt angle of each surface, and surface parameters describe the shape of each surface in the optical system. As a possible network framework, the network accepts system parameters as input and then outputs structure and surface parameters to form only one specific system. Thus, one-to-one mapping can be realized. However, the imaging performance of this specific output system is not optimal in general, and the system structure may not meet the designer's needs. Consequently, the application of this framework is limited. To provide designers with more systems that enable choice and can be filtered or sorted, it is necessary to achieve one-to-multiple mapping, that is, multiple systems should be output when one system parameter combination (or a combination of selected system and structure parameter values) is input. To achieve this goal, structure parameters can also be included in the network input. For the above-mentioned complex one-to-multiple freeform imaging system design framework, the number of input and output parameters is large, and the ranges of the input system and structure parameter values are wide; hence, to obtain a good one-to-multiple network model, previous experience is essential to guide machine learning, particularly in the early phase of the network training process. Therefore, supervised learning should be used, and high-quality freeform systems with combinations of various system and structure parameter values have to be obtained as the "labeled" data set.

The fundamental data set generation framework is shown in Fig. 2. Before the data set is generated for the design framework of one type of system, several representative system parameters that can fully describe the system specifications should be selected. For freeform imaging optical systems,

the FOV, effective focal length (EFL), F-number ($F\#$), and entrance pupil diameter (ENPD) are typically used; however, these parameters are not necessarily used simultaneously because some of them are correlated. Additionally, some parameters can be fixed to a specific value (e.g., EFL), and other systems with various parameter values can be obtained by scaling. Thus, the dimension of the system parameter space, in addition to the number of systems in the data set, can be reduced. The system parameter combination can be expressed as the vector $\boldsymbol{\Phi} = [\Phi_1, \Phi_2, \ldots, \Phi_m, \ldots, \Phi_M]$, where Φ_m denotes a specific parameter type, such as FOV and ENPD. For example, the system parameter combination $\boldsymbol{\Phi}$ of the design example given in Section 3 is [XFOV, YFOV, ENPD], where XFOV and YFOV represent the FOV in x and y directions, respectively. We use a vector $\boldsymbol{\varphi} = [\varphi_1, \varphi_2, \ldots, \varphi_m, \ldots, \varphi_M]$ to denote the specific value of the vector $\boldsymbol{\Phi}$, where φ_m is value of parameter Φ_m. Similarly, the surface shape of a specific system can be characterized by the vector $\mathbf{X} = [X_1, X_2, \ldots, X_v, \ldots, X_V]$, where X_v denotes a specific surface parameter (e.g., surface coefficient). The type of freeform surface is not limited. The locations and tilts of the freeform surfaces can be considered as structure parameters. In this study, only the common case in which the system is symmetric about the YOZ plane is considered. The y and z coordinates of the vertices of the surfaces are considered as structure parameters. If the tilt term (y term) is not used in the surface expression, the partial derivative of y at the surface vertex will be zero, and the surface normal direction at the vertex will coincide with the local z axis of the surface. If the chief ray of the central field is further required to intersect each surface at its vertex, then the surface normal as well as the tilt angle of this surface can be calculated using the coordinates of the surface vertices of this surface and the surface preceding and after this surface. Therefore, the surface tilt angle can be not added to the input structure parameters, and it can reduce the complexity of DNN training. For the output structure parameter, as chief ray of the central field may not intersect each surface at its vertex exactly for the systems generated by the DNN, the surface tilt angle is added as an output structure parameter in order to fully describe the freeform system. The structure parameter combination is denoted by $\boldsymbol{\Psi} = [\Psi_1, \Psi_2, \ldots, \Psi_t, \ldots, \Psi_T]$. We use vectors $\boldsymbol{\chi}$ and $\boldsymbol{\psi}$ to denote the specific value of the vector $\mathbf{X}$ and $\boldsymbol{\Psi}$, respectively. In summary, as shown in Fig. 3, the full input parameters to the DNN are $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$ (excluding the surface tilt angles). The full output parameters are $\boldsymbol{\chi}$ and $\boldsymbol{\psi}$. When $\boldsymbol{\chi}$ and $\boldsymbol{\psi}$ are obtained, along with the input $\boldsymbol{\varphi}$, the output system is fully defined.

Systems with various system and structure parameter values should be generated as the data set for supervised learning. After the characteristic parameters are selected, it is necessary to determine the appropriate parameter space for parameter values selection, which is a high-dimensional space for system and structure parameters. The system parameter space is abbreviated as $\mathrm{SP}_{\boldsymbol{\Phi}}$ and the structure parameter space as $\mathrm{SP}_{\boldsymbol{\Psi}}$. For selected system parameters, the $\mathrm{SP}_{\boldsymbol{\Phi}}$'s corresponding $\mathrm{SP}_{\boldsymbol{\Psi}}$ is vast; hence, it is impossible to fully sample and generate systems in this space. Additionally, the imaging performance of the many systems may be low and the system structure may be invalid (irregular or having extensive light obstruction). The solution
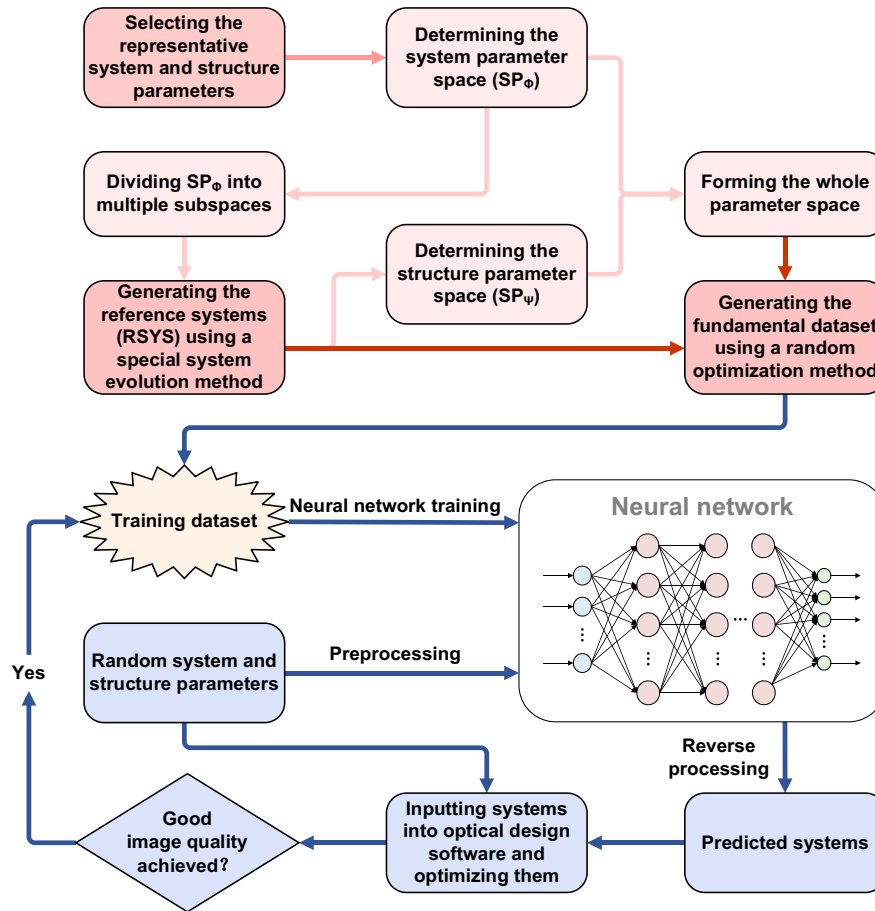
Fig. 2. Illustration of the fundamental data set generation and feedback strategy.

is to divide the entire parameter space into subspaces. For different subspaces of the SP_Φ, their corresponding subspaces of the SP_Ψ differ. The subspace of the SP_Ψ cannot be randomly selected either; otherwise, the above issues will also appear. To address this issue, we propose a method for determining the subspace by using a reference system with good imaging
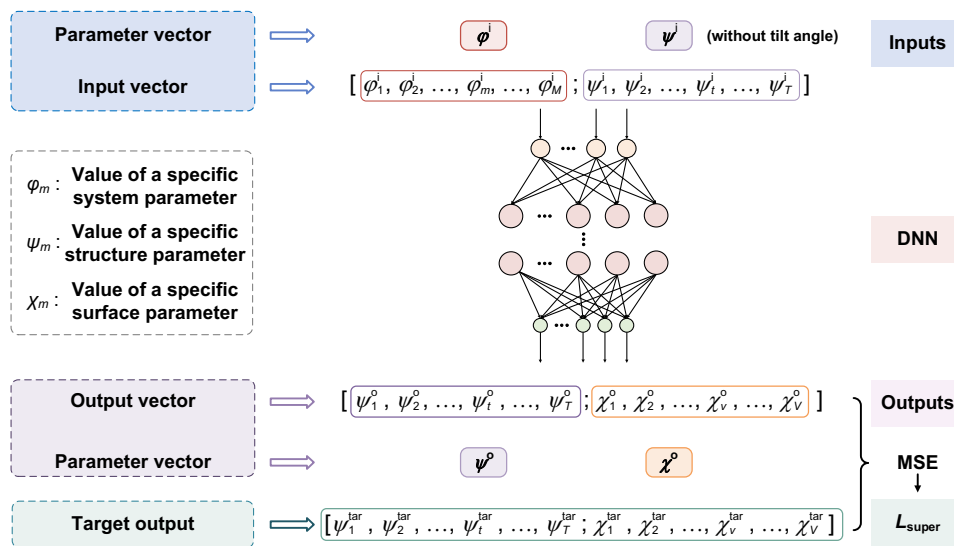

Fig. 3. Illustration of input and output parameters of DNN. The superscripts i, o, and tar denote input, output, and target, respectively. The output and target surface and structure parameters values are used to construct the mean square error (MSE) and then construct the supervised loss function L_{super}.
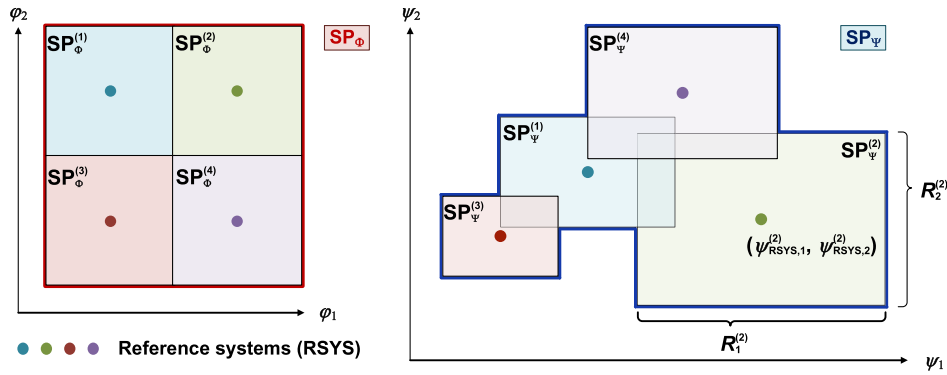
Fig. 4. Sketch of the parameter space and subspace pair $SP_{\Phi}^{(i)}$ and $SP_{\Psi}^{(i)}$. A 2D space (only two system parameters and two structure parameters are considered, respectively) is plotted here for clarity, but actual parameter spaces should be high-dimensional spaces. Four subspaces are plotted here as an example. The subspaces plotted in same color form a subspace pair $SP_{\Phi}^{(i)}$-$SP_{\Psi}^{(i)}$. φ and ψ of reference system $RSYS^{(i)}$ are at the center of $SP_{\Phi}^{(i)}$ and $SP_{\Psi}^{(i)}$.

performance, whose φ and ψ are at the center of the system and structure parameter subspace, as shown in Fig. 4.

After the range of φ_m ($1 \leq m \leq M$) is determined, this range is divided into I_m segments, and the length of each segment is L_m. Thus, SP_{Φ} is divided into $I = \Pi_{m=1}^{M} I_m$ subspaces $SP_{\Phi}^{(i)}$ ($1 \leq i \leq I$) and each $SP_{\Phi}^{(i)}$ has a corresponding $SP_{\Psi}^{(i)}$. A subspace pair is denoted by $SP_{\Phi}^{(i)}$-$SP_{\Psi}^{(i)}$, and all subspace pairs are combined to form the entire parameter space. Next, a reference system $RSYS^{(i)}$ whose system parameter is the center of the $SP_{\Phi}^{(i)}$ is generated for each subspace pair, and all the reference systems are automatically generated by a special system evolution method. The system with the lowest system specifications (easiest to be designed) is optimized first. During evolution, if the system whose system parameter value is φ^* has been generated at a certain step, the weighted distances between φ^* and the system parameter values of all the systems, which have not been optimized, are calculated. Here, the distance $D = ||\mathbf{W} \otimes (\varphi^* - \varphi^{**})||_2$, where $|| \cdot ||_2$ represents the L2-norm, $\otimes$ represents elementwise multiplication, and φ^{**} represent the system parameter value of one system, which has not been optimized. $\mathbf{W} = [W_1, W_2, ..., W_m, ..., W_M]$ is the weight vector, which balances the influence of the different parameters. Among all the remaining systems, the system corresponding to the smallest D is determined to be the next system to be optimized, and it will be evolved from the system with φ^*. The above process is repeated until all the reference systems have been generated. During sequential system evolution, if an abnormal system occurs (fatal ray tracing error, obstruction, or poor imaging performance), the current system is evolved from the second-nearest existing system. During optimization, the image quality, distortion, light obstruction, system structure, and some other constraints should be controlled. When $RSYS^{(i)}$ is generated, it is necessary to obtain its structure parameters $\psi_{RSYS}^{(i)}$. The length of the interval for each structure parameter values' range in this subspace pair is given and represented by the vector $\mathbf{R}^{(i)} = [R_1^{(i)}, R_2^{(i)}, ..., R_t^{(i)}, ..., R_T^{(i)}]$. Consider $\psi_{RSYS}^{(i)}$ as the structure parameter center of the $SP_{\Psi}^{(i)}$. Then, the value range of $\psi_t^{(i)}$ can be determined by $[\psi_{RSYS,t}^{(i)} - 0.5R_t^{(i)}, \psi_{RSYS,t}^{(i)} + 0.5R_t^{(i)}]$; that is, $SP_{\Psi}^{(i)}$ is

determined by $\psi_{RSYS}^{(i)}$ and $\mathbf{R}^{(i)}$. $\mathbf{R}^{(i)}$ should be determined before generating the reference system, and the structure of the reference system can be constrained to avoid the existence of structure with obstructions in the subspace, thus ensuring the reasonability of the structure parameter subspace division. Because of the difference in system size for different system parameters, $\mathbf{R}$ should differ for different subspace pairs. The larger the system parameters, the larger the structure parameters ranges are. The design form of the systems in the subspace is partially determined by the reference system, but the values of structure parameters vary within the subspace.

Next, a random optimization generation method is used to generate the fundamental data set. In each $SP_{\Phi}^{(i)}$-$SP_{\Psi}^{(i)}$, systems with random φ are generated using $RSYS^{(i)}$ as the initial system. During optimization, all the surface parameters of $RSYS^{(i)}$ are set as variables to obtain good imaging performance. To increase diversity of the structure parameter values in the data set, all the structure parameters may be changed to random values within the $SP_{\Psi}^{(i)}$ with probability P_f. Then, the changed structure parameters are frozen with probability P_{f1} during optimization, whereas the unchanged structure parameters are frozen with probability P_{f2}. After optimization, the system parameters, structure parameters, and surface parameters of all the systems are obtained to form the fundamental data set. Because the system generation processes in different subspace pairs are independent, the above process can be performed in parallel to improve the efficiency of data set generation. Using the above method, systems with various system and structure parameters within large ranges, in addition to good imaging performance, are obtained and the required constraint is satisfied. The entire data set generation process is fully automatic; hence, human effort is minimal.

C. Supervised Learning

After the fundamental data set is obtained, the DNN can be trained using supervised training. Because the values and units differ significantly between different parameters, to improve the efficiency and convergence of training, all the input data with the same type are normalized to [−1, 1] using a linear preprocessing method similar to minimum–maximum normalization.

The output data of the network need reverse processing to obtain the actual structure and surface parameters of the system. The loss function L_{super} can be defined as the mean square error (MSE) between the target and actual predicted output (values of surface and structure parameters, χ and ψ), as shown in Fig. 3. After the loss is obtained using a feedforward calculation, the gradient of the L_{super} over all parameters in the neural network is calculated using the backpropagation algorithm. Then, the gradient is fed back to the optimizer and used to update the weights and biases to minimize the loss function, which makes the predicted values closer to the target values.

At this time, the DNN can predict the corresponding systems for specific inputs. However, for a large DNN model, a large data set is required for supervised training. In the combined sequential and random freeform system evolution method mentioned above, all the systems in each subspace pair $SP_\Phi^{(i)}$-$SP_\Psi^{(i)}$ are evolved from a single initial system $RSYS^{(i)}$. Additionally, the number of systems in the data set is small, and the values of different system and structure parameters are limited. A feedback training strategy (as shown in Fig. 2) is proposed to diversify the data set and improve the performance of the DNN. Random φ and ψ are selected in different subspace pairs as the DNN input. Then the systems predicted by the current network are directly input into the optical design software (e.g., CODE V and Zemax) for quick optimization. The structure parameters are frozen with probability P_s during optimization. If the optimized system has no ray tracing error but good image quality, the system parameters are obtained and added to the training data set. Further training is conducted, and system generation feedback is executed again. The above process is repeated, and the number and diversity of the systems in the data set are extended.

D. Unsupervised Learning

Through supervised learning, the DNN can converge quickly using the obtained data set and be used to output systems based on the given inputs. However, the network generalization ability may still be weak. Unsupervised learning is then introduced, and supervised and unsupervised learning are combined for subsequent DNN training.

Unsupervised learning does not require a data set with "label information." When random system and structure parameter combinations are input into the DNN, the output systems can be obtained, and the unsupervised loss function $L_{unsuper}$, which contains two parts, can be constructed. One part, $L_{performance}$, is related to imaging performance. The other part, $L_{constraint}$, is related to the design constraints that need to be satisfied, which are weighted penalty functions based on, e.g., quadratic, reciprocal, logarithmic, or higher-order power functions. Unsupervised learning can be regarded as training the DNN to "optimize" the systems generated by the DNN [27]. A differential ray tracing module of the freeform imaging system is essential to connect $L_{unsuper}$ with the surface and structure parameters of the freeform system as well as the parameters of the DNN [31]. This ensures that the entire prediction and computation processes are fully differentiable; thus, the loss gradient can be computed and backpropagated to update the parameters of the DNN and improve DNN performance. The overall computing efficiency can be improved using GPU computing power and parallel computing. The core of ray tracing is to find the point where the ray intersects with the surface and then obtain the outgoing ray direction. The expressions of common freeform surfaces (XY polynomials surface, Zernike polynomials surface, etc.) in local coordinate systems can be written as follows:

$$z = h(x,y) = \frac{c(x^2 + y^2)}{1 + \sqrt{1 - (1+k)c^2(x^2 + y^2)}} + \sum_{i=0}^{p} A_i \varphi_i(x,y), \tag{1}$$

where c represents the curvature, k is the conic constant, $\varphi_i(x,y)$ is the freeform surface term, and A_i is its coefficient. For a ray with a starting point coordinate $\boldsymbol{\mu} = [\mu_x, \mu_y, \mu_z]$ and a unit propagation direction vector $\boldsymbol{\omega} = [\omega_x, \omega_y, \omega_z]$, it can be represented by $(\boldsymbol{\mu}, \boldsymbol{\omega})$. The coordinates of the ray $(\boldsymbol{\mu}, \boldsymbol{\omega})$ after propagating for λ units of length can be written as $\boldsymbol{\mu} + \lambda\boldsymbol{\omega}$. The intersection of this ray with the surface $z = h(x,y)$ shall meet

$$f(x,y,z) = h(x,y) - z = f(\boldsymbol{\mu} + \lambda\boldsymbol{\omega}) = 0. \tag{2}$$

λ can be solved iteratively using Newton's method:

$$\lambda^{(n)} = \lambda^{(n-1)} - \frac{f(\boldsymbol{\mu} + \lambda^{(n-1)}\boldsymbol{\omega})}{f'(\boldsymbol{\mu} + \lambda^{(n-1)}\boldsymbol{\omega})} = \lambda^{(n-1)} - \frac{f(\boldsymbol{\mu} + \lambda^{(n-1)}\boldsymbol{\omega})}{\nabla f \cdot \boldsymbol{\omega}}. \tag{3}$$

The above process is conducted iteratively until the change of λ is smaller than the allowable tolerance. After λ is obtained, we obtain the coordinate of the intersection $\boldsymbol{\mu} + \lambda\boldsymbol{\omega}$, which is also the start point $\boldsymbol{\mu}'$ of the ray reflected or refracted by the surface. The propagation direction $\boldsymbol{\omega}'$ after this surface can be calculated based on Snell's law. Then, by repeating the above process, the ray tracing process can be conducted surface-by-surface sequentially until the ray reaches the image plane.

The actually used rays of each field in the system are determined based on the location of the aperture stop and the chief ray of each field. Once the chief ray is found, other rays can be sampled according to the entrance pupil size and used for imaging performance analysis and optimization. As the freeform system is nonsymmetric, the traditional method of locating the chief ray by obtaining the paraxial entrance pupil position may be not suitable. Therefore, an iterative search method of the chief ray is adopted. The chief ray of each field should intersect with the aperture at its center. Trace a ray $(\boldsymbol{\mu}^b, \boldsymbol{\omega}^b)$ starting from the object space of the optical system with a field point $\boldsymbol{\omega}^b$ and obtain the coordinates $\boldsymbol{\tau}^b$ of its intersection with the aperture stop in its local coordinate system. Here, the point where the ray intersects with a virtual plane in the object space can be taken as $\boldsymbol{\mu}^b$. This process of tracing and finding the intersection point can be expressed by $\boldsymbol{\tau}^b = \varepsilon(\boldsymbol{\mu}^b)$, and its derivative $\varepsilon'(\boldsymbol{\mu}_c^{b,(n)})$ can be calculated by automatic differentiation. The iteration process follows

$$\boldsymbol{\mu}_c^{b,(n+1)} = \boldsymbol{\mu}_c^{b,(n)} - \frac{\boldsymbol{\tau}_c^{b,(n)}}{\varepsilon'(\boldsymbol{\mu}_c^{b,(n)})}, \tag{4}$$

where subscript c represents the chief ray, n is the number of iterations, and iteration can be stopped when $\boldsymbol{\tau}_c^{b,(n)}$ is less than the allowable value. The x and y coordinates of the initial iteration ray's starting point can be the same as the vertex coordinates of the first surface of the system.

Loss functions can be constructed based on the result of ray tracing. $L_{\text{performance}}$ includes the loss related to the root mean square (RMS) spot size radius l_{spot} and the loss related to distortion l_{dist}. l_{spot} is the average value of the RMS spot radius among the full FOV and different wavelengths:

$$l_{\text{spot}} = \frac{1}{H \times W} \sum_{h,w} \left\{ \frac{1}{P} \sum_p [(x_p^{h,w} - x_c^{h,w})^2 + (y_p^{h,w} - y_c^{h,w})^2] \right\}^{\frac{1}{2}},$$

(5)

where x and y represent the local coordinates of the image point of a specific ray; the subscript c represents the chief ray; h, w, and p represent the index of the sampled field point, sampled wavelength, and sampled pupil ray, respectively; and H, W, and P are the total numbers of h, w, and p, respectively.

The relative distortion of off-axis sampled field points is used to construct l_{dist}:

$$l_{\text{dist}} = \frac{1}{H \times W} \sum_{h,w} \frac{[(x_c^{h,w} - x_{\text{ideal}}^h)^2 + (y_c^{h,w} - y_{\text{ideal}}^h)^2]^{\frac{1}{2}}}{[(x_{\text{ideal}}^h)^2 + (y_{\text{ideal}}^h)^2]^{\frac{1}{2}}},$$

(6)

where the subscript "ideal" represents the ideal image point.

$L_{\text{constraint}}$ includes constraints related to, for example, the system specifications (e.g., EFL and F#), light obstruction, system volume, and the chief ray of the central field. The difference between the output and input structure parameters should also be constrained to make the DNN work normally.

For the calculation of EFL, the traditional method used for a rotational symmetric system is not feasible; hence, we use a method that uses real ray tracing. For the focal length EFL_x in the x direction, the chief ray of a field with small angle $\Delta\theta$ relative to the central field in the x direction is traced, and the image height h_x in the x direction relative to the central field can be obtained. Based on $h_x = \text{EFL}_x \times \tan \Delta\theta$, EFL_x can be calculated. The focal length EFL_y in the y direction can also be calculated using the same method. The loss of EFL l_{EFL} can be calculated using the MSE loss between the actual focal lengths and the required focal length EFL^*:

$$l_{\text{EFL}} = \frac{1}{2}[(\text{EFL}_x - \text{EFL}^*)^2 + (\text{EFL}_y - \text{EFL}^*)^2].$$

(7)

Similar methods can be used to calculate other system parameters, such as the F# and magnification.

Light obstruction should be eliminated in freeform imaging optical systems, particularly for reflective systems. It can be controlled by constraining the distances between the edge points of surfaces and the edge rays of light beams. The distances can be calculated using real ray trace data. When the loss l_{obs} related to light obstruction is calculated, all the J key distances d_j (a negative distance means that obstructions exist) are determined. If d_j is greater than given tolerance $d_{\min j}$, the residual space is considered to be sufficient, and the contribution to the loss function should be zero; otherwise, a penalty value is added. The loss functions can be written as

$$l_{\text{obs}} = \frac{1}{J} \sum_{j=1}^{J} l_{\text{obs},j}, \qquad \text{where}$$

$$l_{\text{obs},j} = -\min(d_j - d_{\min j}, 0).$$

(8)

Different from aspherical surfaces, using the off-axis section of the freeform surface will not improve the correction ability of nonsymmetric aberrations, as the on-axis section of the surface is also nonrotationally symmetric. The off-axis section of a freeform surface can be characterized by the on-axis section of another freeform surface. In addition, if the used area of the freeform surface is different from the mathematical vertex (the origin of the local coordinate system) of the surface, unnecessary troubles may appear during optomechanical design and system assembly. Therefore, for the design of freeform imaging systems, it is generally required that the chief ray of the central field intersects with each freeform surface at its vertex (the local coordinates should be zero). The local coordinates $(x_c^{1,s}, y_c^{1,s})$ of the chief ray of the central field with the sth surface ($1 \leq s \leq S$) can be obtained using ray tracing. The loss function $l_{c\text{-ray}}$ is defined as

$$l_{c\text{-ray}} = \left\{ \frac{1}{S} \sum_s [(x_c^{1,s})^2 + (y_c^{1,s})^2] \right\}^{\frac{1}{2}}.$$

(9)

As the DNN should output systems whose structure parameters are the same as or similar to the input, the difference between the input and output structure parameters should be controlled. The loss l_{str} is calculated using the root mean square error loss:

$$l_{\text{str}} = \left[\frac{1}{T} \sum_t (\text{STP}_t^o - \text{STP}_t^i)^2 \right]^{\frac{1}{2}},$$

(10)

where ψ_t^i and ψ_t^o are the tth input and output structure parameter, respectively.

L_{unsuper} can be constructed using the weighted sums of the above individual losses for all output systems, where w_{spot}, w_{dist}, w_{EFL}, w_{obs}, $w_{c\text{-ray}}$, and w_{str} are the weights. Additionally, for systems with more advanced system specifications, more aberrations should be tolerated. An adjustment factor ρ_w is added to balance the loss contribution of different systems to $L_{\text{performance}}$, and ρ_w decreases as the system specifications increase. The losses of different systems can be summed:

$$\begin{aligned} L_{\text{unsuper}} &= \sum (\rho_w L_{\text{performance}} + L_{\text{constraint}}) \\ &= \sum [\rho_w(w_{\text{spot}} l_{\text{spot}} + w_{\text{dist}} l_{\text{dist}}) \\ &\quad + (w_{\text{EFL}} l_{\text{EFL}} + w_{\text{obs}} l_{\text{obs}} + w_{c\text{-ray}} l_{c\text{-ray}} + w_{\text{str}} l_{\text{str}})]. \end{aligned}$$

(11)

The total loss when supervised and unsupervised trainings are combined is

$$L_{\text{total}} = L_{\text{super}} + w_{\text{unsuper}} L_{\text{unsuper}}.$$

(12)

By regulating weight w_{unsuper}, the contribution of unsupervised learning to overall network training can be modulated. Network training can be conducted based on L_{total} to obtain the final DNN called FreeformNet with good performance. The overall framework of combined supervised and unsupervised training is shown in Fig. 5.

E. System Generation

After FreeformNet is obtained, as shown in Fig. 6, for a specific design task, the values of all input system parameters and structure parameters can be provided according to the design requirements. In this case, the single design result tailored to
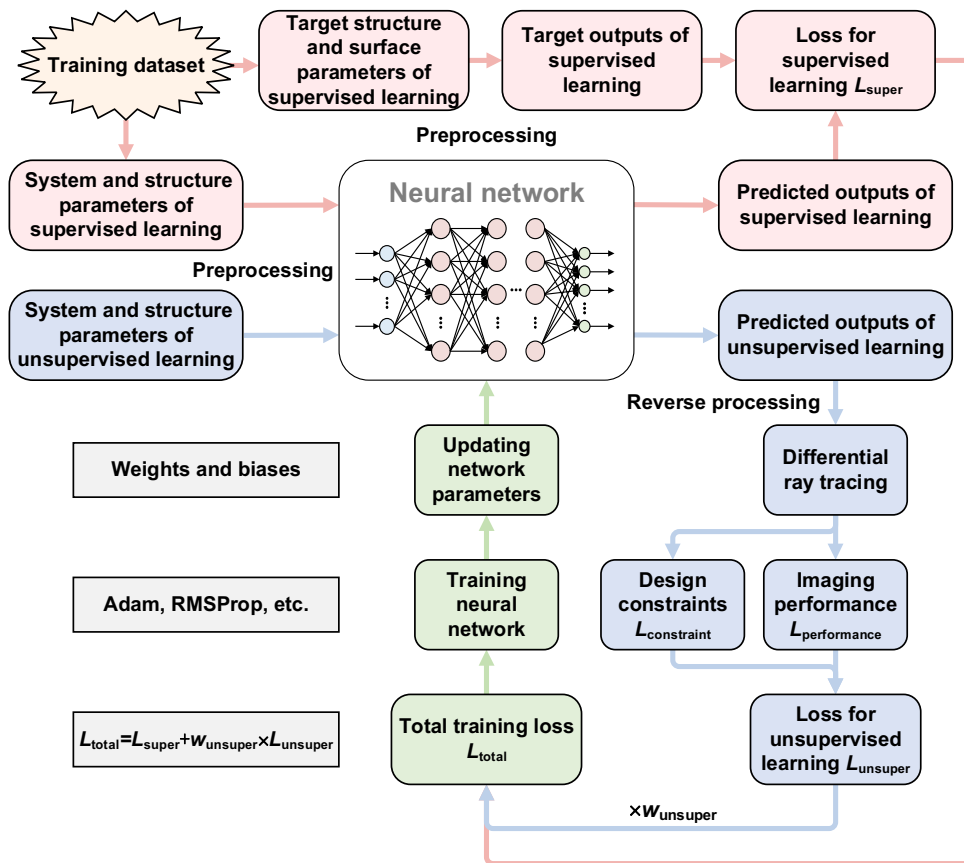
Fig. 5. Illustration of the training mode of combined supervised and unsupervised training.
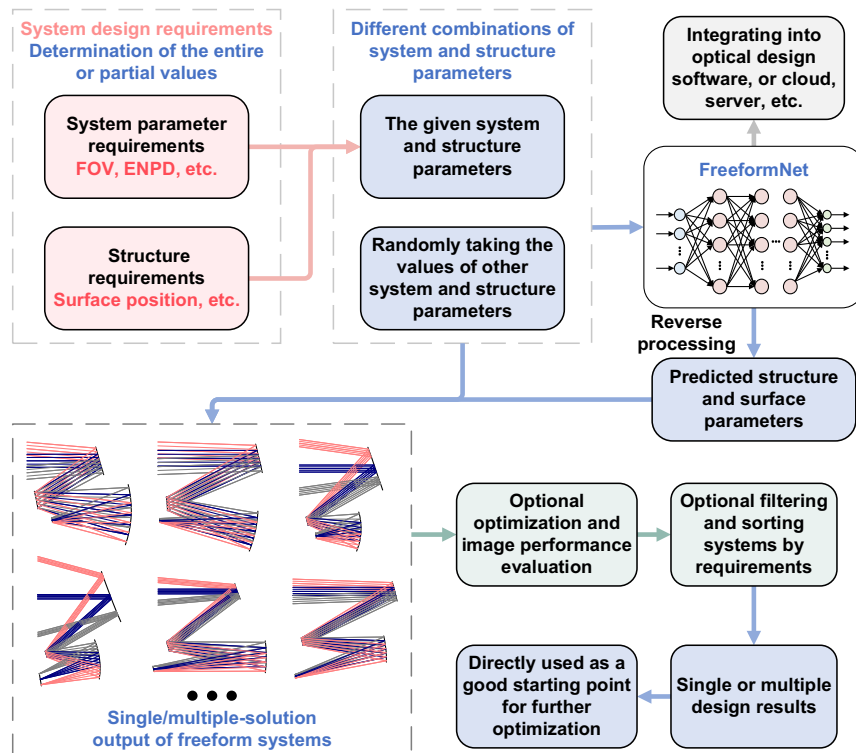

Fig. 6. Fast generation process of multiple-solution freeform imaging systems using FreeformNet.

the input parameters can be output immediately by FreeformNet. Another case is when the values of only some of the system parameters and structure parameters have been provided based on actual design needs. In this case, for parameters without designated values, the values can be randomly selected within the corresponding parameter range. Combined with the determined parameters, a certain number of different parameter value combinations can be obtained and input into FreeformNet. A series of systems that meet the basic design requirements can be generated, and a fast multisolution output with different structure and system parameters is realized. For a traditional or existing freeform system generation method using DNN, the design input is only limited to specific system parameters combinations, and only one system can be generated fulfilling the design requirements. Compared with the existing single-solution method, in our design framework, for given specific system parameters (or partial system and structure parameters), multiple solutions can be generated that fulfill the given fixed input parameters (basic design requirements) but having different values for other parameters. Therefore, the proposed method can be seen as a "multisolution" method. Based on the preset evaluation indicators (e.g., various imaging performance metrics and the system volume) and other constraints, the solutions can be filtered or sorted and ready for user selection. The systems generated by FreeformNet can be used directly as good starting points for further optimization. The systems output directly from the network can also be quickly optimized in parallel to generate output systems with substantially improved imaging performance, thus greatly improving the efficiency of freeform optical design. For systems with small or moderate FOV, the optimization can be done in several seconds for a single system in general. In this way, it only takes several seconds in total to obtain a freeform system with good performance after the design requirements are input into the FreeformNet. For systems with large FOV, the time cost for generating the starting point using FreeformNet is the same. However, the optimization difficulty is much larger if high imaging performance across the full FOV is required. Some trial-and-error and minor adjustments may be required during the optimization process in order to balance the aberrations. Therefore, the optimization time may be longer.

3. RESULTS

The freeform off-axis three-mirror imaging system design was used to verify the feasibility and effect of the proposed design

framework and FreeformNet. The selected freeform off-axis three-mirror imaging system had a central field of (0°, 0°) and symmetry about the YOZ plane, with traditional zig-zag folding geometry, as shown in Fig. 7(a). M1, M2, M3, and IMG denote the primary mirror, secondary mirror, tertiary mirror, and image plane, respectively. The aperture stop was located at M2. For this system, the FOV in the x direction (XFOV), FOV in the y direction (YFOV), and ENPD were chosen to describe the system parameters; that is, $\boldsymbol{\Phi} =$ [XFOV, YFOV, ENPD]. The focal length was set to a fixed value of 1 mm, and systems with other focal lengths were obtained by scaling. System structure parameters included the tilt angles and the y and z vertex coordinates of all the surfaces in the system. Because the origin of the global coordinate system was set to the vertex of M2, the y and z coordinates of the vertex of M2 were not included in $\boldsymbol{\Psi}$. Therefore, $\boldsymbol{\Psi} = [M1_y,$ $M1_z, M3_y, M3_z, IMG_y, IMG_z, M1_{tilt}, M2_{tilt}, M3_{tilt}, IMG_{tilt}]$ (note that the input structure parameter did not contain the tilt angles). The freeform surface type was an XY polynomial freeform surface up to the sixth order with no base conic for simplicity, and the odd-order terms of x were not used. Thus, **X** had 42 individual parameters. Overall, system parameters and partial structure parameters were considered as network inputs, and there were nine in total. All structure and surface parameters were considered as network outputs, and there were 52 in total.

The next step was to determine the appropriate $SP_{\boldsymbol{\Phi}}$. XFOV and YFOV were specified in the range of 4°–40°, and ENPD in the range 1/6–2/3 mm ($F\#$ in the range 1.5–6). However, because it was difficult to achieve a large etendue (for example, a large FOV and small $F\#$ simultaneously), while satisfying the design requirements and achieving good imaging performance, only system parameters that satisfied the following condition were selected:

$$0.5 \times \text{XFOV} + 0.5 \times \text{YFOV} + 20 \times \text{ENPD} \leq 33.34. \quad \textbf{(13)}$$

In this way, the etendue will be effectively limited. During the training of the DNN, the training result for the systems at the edge of the system parameter space may be bad, as the number of systems in the data set whose system parameters are close to these systems at the edges is smaller, compared with the systems in the inner part of the parameter space. Therefore, to achieve a good overall training effect for the system parameters within the specified range, the entire system parameter space for network training was moderately larger than the system parameter space used for system generation. This can be done by
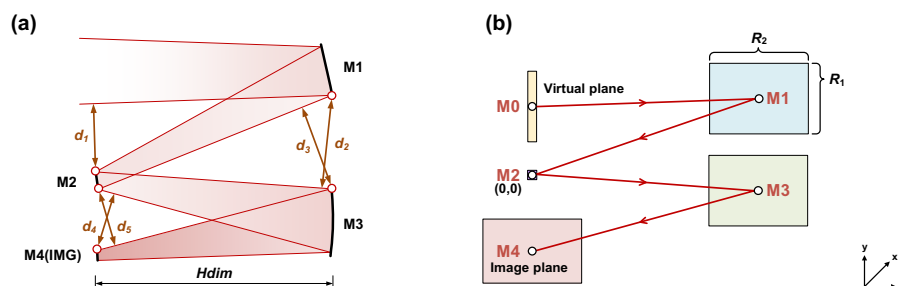

Fig. 7. (a) The selected folding geometry of freeform off-axis three-mirror imaging system and its structure constraints. (b) Sketch of the concept of structure parameters range.

extending the upper and lower limits of mth system parameter Φ_m by $0.5L_m$ (here, L_m is the length of Φ_m for each subspace, which is defined in Section 2.B). For example, for Φ_m, if the range of its value for system generation is $a \leq \varphi_m \leq b$, then, during the DNN training process, the range is enlarged to $a - 0.5L_m \leq \varphi_m \leq b + 0.5L_m$. In this way, good imaging performance of the systems generated by the DNN can be obtained. Then, the SP_{Φ} was divided into smaller subspaces. The ranges of XFOV, YFOV, and ENPD are segmented according to the lengths $L_1 = 2°$, $L_2 = 2°$, and $L_3 = 0.05$, respectively. Therefore, $I_1 = 19$, $I_2 = 19$, and $I_3 = 11$ were calculated, and the total number of subspace pairs was $I = \Pi_{m=1}^{3} I_m = 19 \times 19 \times 11 = 3971$. After the subspace pairs whose central system parameters did not meet Eq. (13) were removed, 2598 subspace pairs remained.

After the subspace pairs were divided, RSYS was generated using a special system evolution method. The optimization of the reference systems was completed using optical design software CODE V. As the system was symmetrical about the YOZ plane, only half of the full FOV was considered in the design. The six sampled field points $(0, 0)$, $(0, \text{YFOV}/2)$, $(0, -\text{YFOV}/2)$, $(\text{XFOV}/2, 0)$, $(\text{XFOV}/2, \text{YFOV}/2)$, and $(\text{XFOV}/2, -\text{YFOV}/2)$ were selected for each system. During system generation, the focal length, light obstruction, distortion, and intersection coordinates of the chief ray of the central field with the freeform surfaces were controlled and allowed for larger aberrations for systems with larger ENPD and FOV. The elimination of obstructions was achieved by controlling the five distances d shown in Fig. 7(a). The maximum acceptable relative distortion of the off-axis field in the x and y directions for each system is

$$\text{Max Distortion}(\text{FOV}, \text{ENPD}) = \frac{0.15 \times \text{FOV} + 10 \times \text{ENPD}}{100}. \tag{14}$$

Let $\mathbf{R}^{(i)} = \mathbf{A} \times \text{Hdim}_{\text{RSYS}}^{(i)}$, where $\text{Hdim}_{\text{RSYS}}^{(i)}$ was the horizontal dimension (along the z direction) of the $\text{RSYS}^{(i)}$ and $\mathbf{A} = [A_1, A_2, \ldots, A_t, \ldots, A_T]$, which was corresponding to the input $\mathbf{\Psi}$, and the value of $\mathbf{A}$ was the same for all subspace pairs. Before generating the reference systems, the value of A_t was determined as $A_1 = A_2 = \cdots = A_6 = 1/3$. As the reference system was optimized, Hdim was constantly changing, resulting in $\mathbf{R}$ also changing until optimization was finished and the final value of $\mathbf{R}$ was determined. To minimize the structure with obstructions in the subspace pair, when generating the reference system, the distance between surfaces could be controlled according to the value of $\mathbf{R}$ to ensure, as much as possible, no overlap between color blocks in Fig. 7(b) (circles indicate each surface of the reference system; color blocks indicate the range of position for each surface). The evolution direction of the reference systems was determined by weight $\mathbf{W}$, where the value was $\mathbf{W} = [1, 1, \sqrt{12}]$. The reference systems were generated automatically, which took 0.83 h. In the process of evolution, the structure parameter values of all the reference systems were obtained, and all the subspace pairs were determined.

Then, the full fundamental data set was generated using a random optimization generation method. The values of P_f, P_{f1}, and P_{f2} were 0.6, 0.5, and 0.2, respectively. Twelve

systems were generated in each subspace pair, and a total of 31,176 systems were obtained for 2598 subspace pairs, which took 10.2 h. By obtaining the system parameters, structure parameters, and surface parameters of these systems, the fundamental data set was formed.

Next, the DNN was pretrained with the fundamental data set. The DNN used in this example had 20 hidden layers, and the largest number of nodes in one layer was 300. The activation function, optimizer, loss function, learning rate, and batch number were the tanh function, Adam, MSE, 10^{-4}, and 50, respectively. All the data in the training data set were fed into the network. A total of 15,000 epochs were performed in pretraining, which took 2 h. The final loss function value was 2.67×10^{-3}. Training was performed on a computer with an Intel Core i9-12900K CPU at 3.2 GHz, 64 GB of internal memory, and NVIDIA GeForce RTX 3090 Ti GPU. After pretraining, the feedback training strategy was further used. Each time the system generation feedback was executed, 1000 random parameter combinations were selected from different subspace pairs as the DNN input. The output systems were optimized, and the P_s value was 0.5. The good systems were then added to the data set for further training. After feedback training is completed, supervised training was continued for a period of time, and the learning rate decreased gradually as training progressed. A total of 147,000 epochs were trained, the number of systems in the data set grew to 109,703, which took 73.3 h, and the final loss function value achieved was 4.68×10^{-4}. The performance of the model was tested using 2,000 random inputs. A total of 115 systems had ray tracing errors, 26 systems had obstructions, and the average RMS spot diameter of the remaining 1859 normal systems was 0.0044 mm.

Next, unsupervised learning was introduced. As the structures with obstructions are avoided as much as possible when determining the subspace pairs, the loss function related to obstruction was not used during unsupervised training in this example. If the structure parameter range is relatively large, and there are inevitably many system structures with obstructions in other design cases, adding this loss function can be considered. The weights of each part of training loss were set as follows: $w_{\text{unsuper}} = 5$, $w_{\text{spot}} = 4$, $w_{\text{dist}} = 0.05$, $w_{\text{EFL}} = 0.05$, $w_{\text{c-ray}} = 0.1$, $w_{\text{str}} = 5$, and $\rho_w = -0.01 \times (\text{XFOV} + \text{YFOV}) - 0.01 \times \text{ENPD} + 1.1$, and the batch number was adjusted to 5. Input data used for unsupervised training were not fixed. In each epoch, we randomly selected 500 subspace pairs and randomly generated a combination of input parameter values in each subspace pair as an input and then used these 500 inputs for unsupervised training. A total of 1210 epochs were performed in the combined supervised and unsupervised training, which took 81.5 h. The performance of the model was tested again and the same 2000 systems were predicted. The number of systems with a ray tracing error decreased to 61 and the average RMS spot diameter of the normal systems decreased to 0.0025 mm. Thus, unsupervised training significantly improved the performance of the DNN.

The obtained network FreeformNet can quickly generate multiple-solution freeform imaging systems based on the design requirements (DNN input). To help the designer
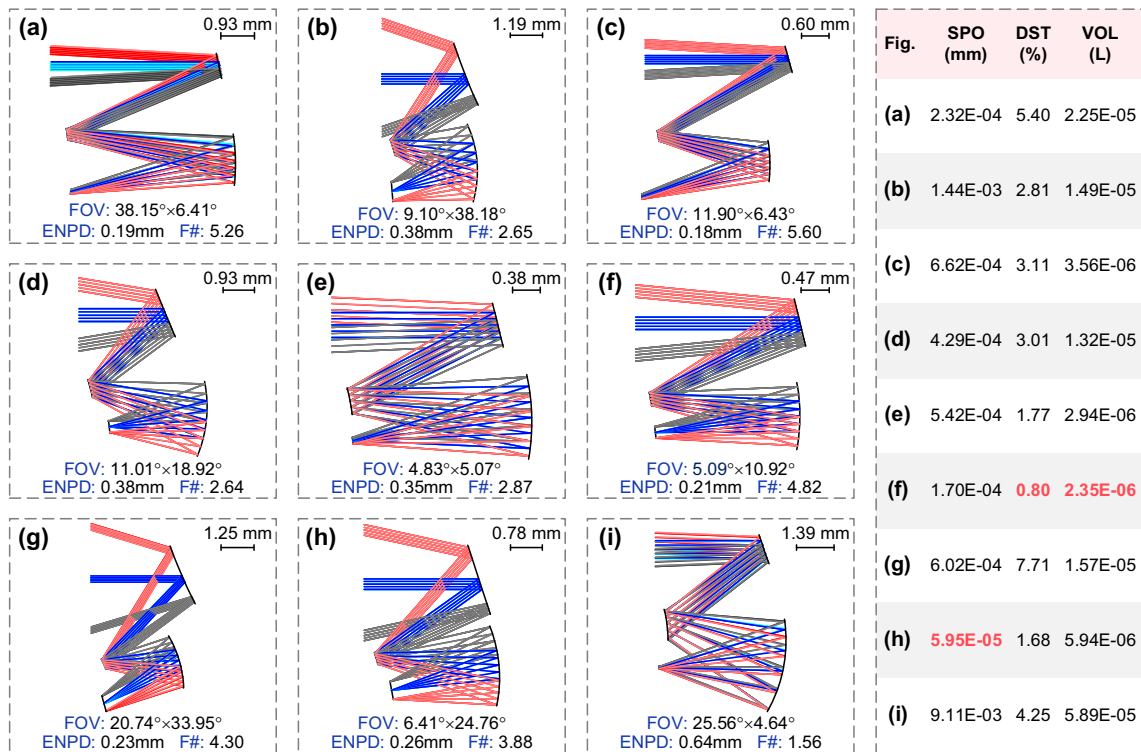
Fig.	SPO (mm)	DST (%)	VOL (L)
(a)	2.32E-04	5.40	2.25E-05
(b)	1.44E-03	2.81	1.49E-05
(c)	6.62E-04	3.11	3.56E-06
(d)	4.29E-04	3.01	1.32E-05
(e)	5.42E-04	1.77	2.94E-06
(f)	1.70E-04	0.80	2.35E-06
(g)	6.02E-04	7.71	1.57E-05
(h)	5.95E-05	1.68	5.94E-06
(i)	9.11E-03	4.25	5.89E-05

Fig. 8. Typical predicted systems when system parameters and structure parameters were all provided.

complete the above design task using FreeformNet, a corresponding program has been written and used. Designers can input the system's entire or partial system and structure parameters based on the design requirements, and single or multiple systems can be generated. It is possible to choose whether to evaluate the system quality and whether to perform further software optimization. The output systems can be filtered or sorted based on a selected metric, such as the average RMS spot diameter, the maximum relative distortion of the sampled fields, volume of the system, and modulation transfer function, and are then ready for user selection.

We used three different kinds of parameter inputs to evaluate the effect of the FreeformNet obtained above. In each case, 5000 inputs were tested. The predicted system focal length was expected to be 1 mm, and the predicted systems were not further optimized.

In the first case, all system parameters and structure parameters were provided (all the parameter combinations were random values in the parameter space). A single system would be output corresponding to one input. Among the 5000 predicted systems, 138 systems had ray tracing errors, 43 systems had obstructions, and the average RMS spot diameter of the other 4819 normal systems was 0.0026 mm. Except for a few systems with relatively larger aberrations, the output systems can be taken as good starting points for further optimization. Figure 8 shows nine typical predicted systems, where SPO, DST, and VOL represent the average RMS spot diameter, the maximum relative distortion of the sampled fields, and system volume, respectively. The average RMS spot diameter and the maximum relative distortion of all the normal systems are

shown in Fig. 9. If further optimization is conducted, it takes about several seconds in general to obtain good imaging performance for a system with narrow or moderate FOV.

The second case was to provide all system parameters and partial structure parameters. The provided parameters were XFOV = 24°, YFOV = 16°, ENPD = 1/3 mm ($F\#$ = 3.00), and M1$_y$ = 2.38 mm (note that M2$_y$ = 0). Other structure parameters were set as random values within corresponding subspaces. Among the predicted systems, three systems had ray tracing errors, and no system had obstructions. The average RMS spot diameter of the other 4997 normal systems was 0.0010 mm. Among them, nine typical systems with various structure parameters are shown in Fig. 10. The systems with
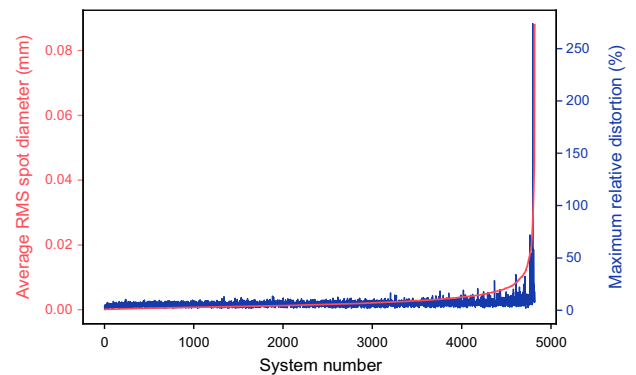

Fig. 9. Average RMS spot diameter and maximum relative distortion of normal systems in case one (system numbers were arranged in ascending order according to the average RMS spot diameter).
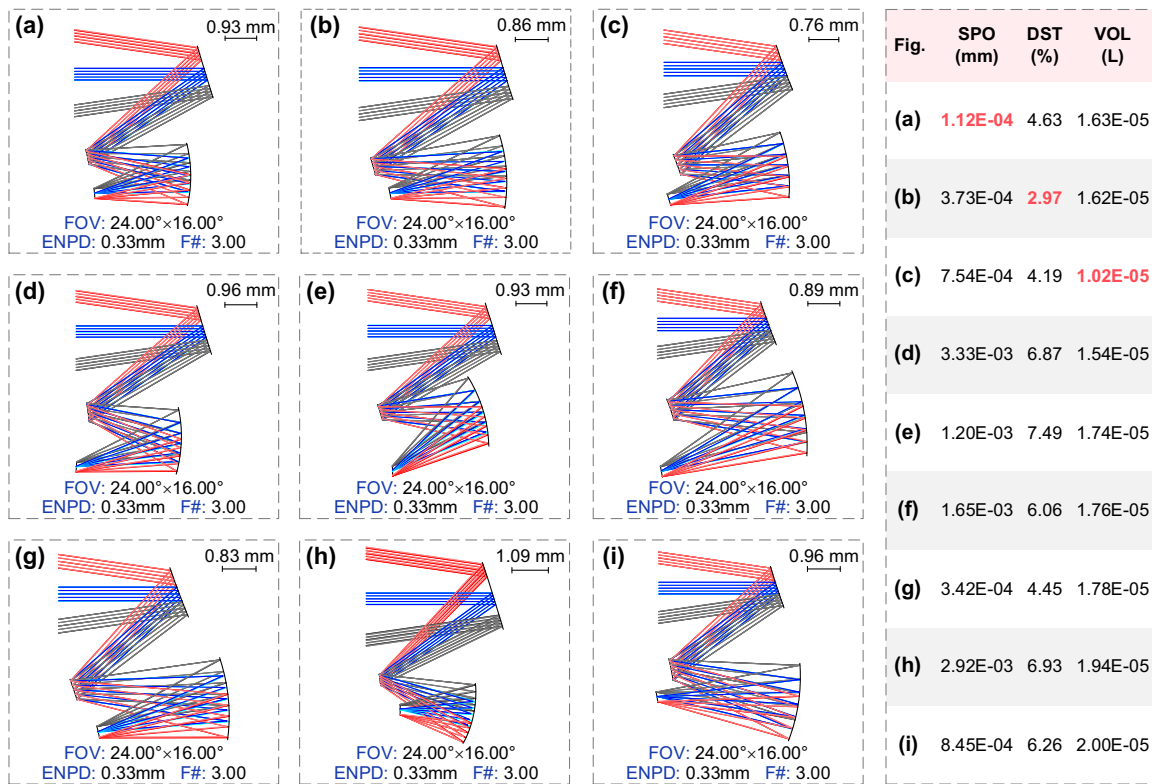
Fig.	SPO (mm)	DST (%)	VOL (L)
(a)	**1.12E-04**	4.63	1.63E-05
(b)	3.73E-04	**2.97**	1.62E-05
(c)	7.54E-04	4.19	**1.02E-05**
(d)	3.33E-03	6.87	1.54E-05
(e)	1.20E-03	7.49	1.74E-05
(f)	1.65E-03	6.06	1.76E-05
(g)	3.42E-04	4.45	1.78E-05
(h)	2.92E-03	6.93	1.94E-05
(i)	8.45E-04	6.26	2.00E-05

Fig. 10. Typical predicted systems when all system parameters and partial structure parameters were provided.

the smallest RMS spot diameter, smallest distortion, and smallest volume are shown in Figs. 10(a), 10(b), and 10(c) respectively. The average RMS spot diameter, the maximum relative distortion of the sampled fields, and system volume of all the 4997 normal predicted systems are given in Fig. 11.

The third case was to provide partial system parameters. The provided parameters were XFOV = 30° and ENPD = 0.2 mm ($F\# = 5.00$). YFOV was a random value within 4°–28.68°, and all the structure parameters were set as random values within corresponding subspaces. Among the predicted systems, 113 systems with ray tracing errors, seven systems had obstructions. The average RMS spot diameter of the other 4880 normal systems was 0.0015 mm. Among them, nine typical systems with various structure parameters are shown in Fig. 12. For all the design examples given in Section 3, the power distribution on different mirrors of the output systems is dependent on the training data set and unsupervised learning. In this paper, during the system optimization included in the data set generation and the unsupervised learning process, the design mainly focuses on achieving high imaging performance, and the power distribution is not considered. For a freeform off-axis three-mirror design with a traditional zig-zag folding geometry and no intermediate image while M2 is taken as the aperture stop, the optical power of M1 is generally smaller than that of M3 (see Refs. [17,19,20,22,23,25]). If specific power distribution is required for the systems generated by the DNN, systems with various or specific power distribution types should be included in the training data set and unsupervised learning by adding specific design constraints during optimization. Furthermore, specific power distribution can be

realized during final optimization on the generated starting points.

The system and structure parameter combinations used for testing (created designs) are generated randomly within the parameter space, and are all different from the system and structure parameters used in the data set for network training. As the above design examples show, good starting points (corresponding to random testing system and structure parameters) can be generated by the network in most cases, which prove that the network generalizes. New systems are "learned" from the data set and unsupervised learning under the design framework and are not "memorized," as they do not exist in the training data set (as well as the input during unsupervised learning).
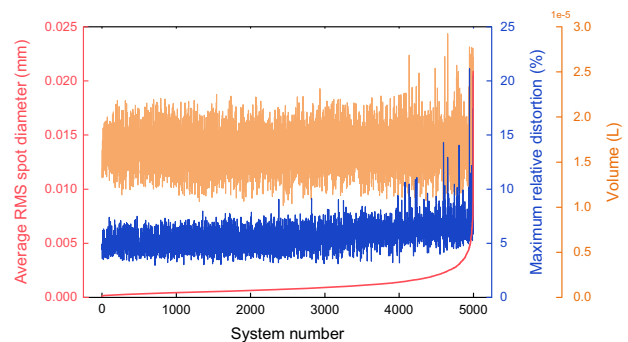

Fig. 11. Average RMS spot diameter, maximum relative distortion, and system volume of normal systems in the second case (the system numbers were arranged in ascending order according to the average RMS spot diameter).
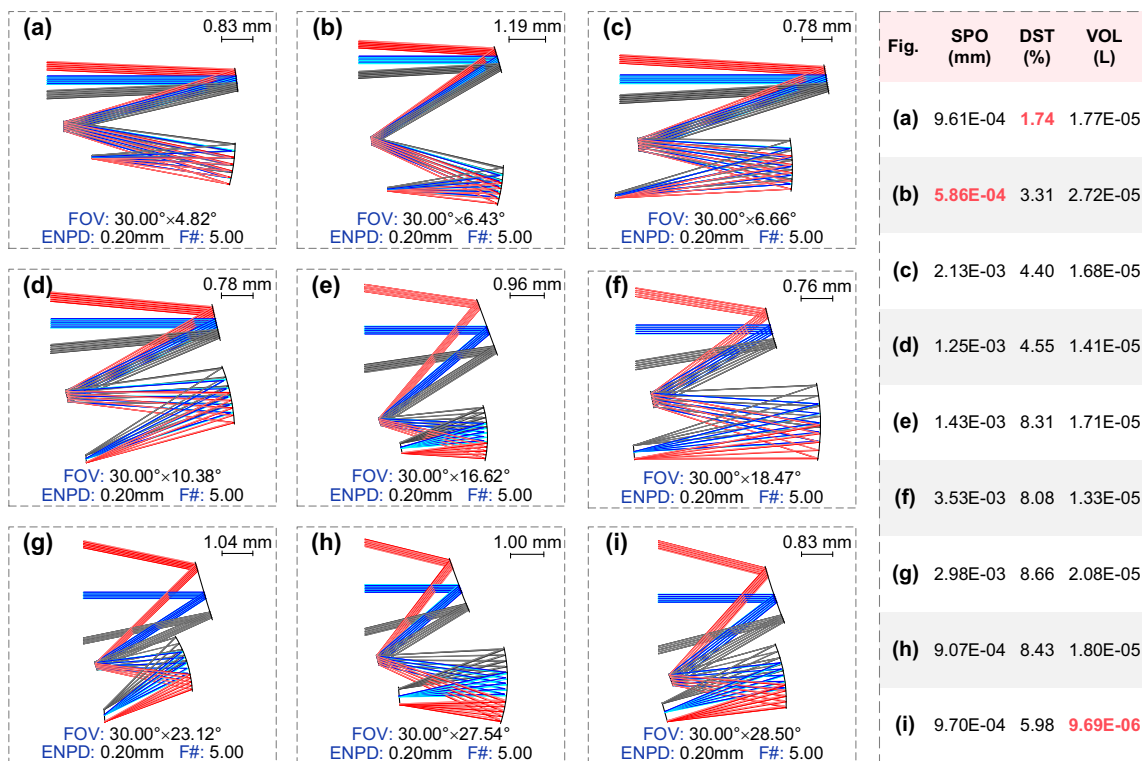
Fig. 12. Typical predicted systems when partial system parameters were provided.

We also compared the proposed method with some typical design methods of freeform imaging systems. All the methods will be focused on the design of freeform off-axis three-mirror systems. The comparison results are listed in Table 1. For the method proposed in the manuscript, each system (starting point) can be generated in less than 0.003 s. The time cost can be further reduced if more different systems are predicted simultaneously. For example, predicting 1000 systems simultaneously takes 0.015 s, 1.5×10^{-5} s per system. Note that our method works for a system with narrow and wide FOV (the range of the system parameters can be wide). The networks can be actually and easily used or integrated into current optical design software and act as a powerful design tool or "database" for ultrafast system starting point generation. For other methods in the table, it is generally not applicable for systems with wide FOV (or not reported). In addition, for these design methods, the knowledge obtained in one design task can barely be transferred and used for other design tasks; thus, it is necessary to start from scratch using these methods when approaching a new design task. If the starting points are not designed directly but are searched and found from literatures, maybe tens of minutes or several hours will be needed, and there is a large probability that a feasible starting point cannot be found. Other report methods based on deep-learning are not applicable for the freeform system design nor applicable for the multiple-solution design.

If further optimization is conducted for the systems generated by the FreeformNet, it takes several seconds in general to obtain good imaging performance for a system with narrow or moderate FOV. Considering the ultrafast generation of the starting point, it only takes several seconds in total to generate a high-performance system after the design requirements are input into the FreeformNet. For example, for each design task corresponding to the first six reference design specifications in Table 1, 20 systems with different structure parameters are generated and optimized. The optimization of each group of 20 systems can be done in between 32 and 37 s, less than 2 s for each system. In conclusion, compared with existing methods, the method proposed in this paper realizes an ultrafast multiple-solution freeform imaging system generation, which works for a wide range of system and structure parameters. Further optimization can be done in several seconds for a system with narrow or moderate FOV. The efficiency of the freeform optical design improved significantly and human effort was minimized.

4. CONCLUSIONS AND DISCUSSION

In this paper, we proposed a design framework enabling fast generation of freeform imaging systems based on deep learning. The training data set is generated automatically using a combined sequential and random freeform system evolution method, and its diversity is further improved by a special feedback strategy. Supervised learning and unsupervised learning based on freeform surface differential ray tracing are combined to obtain a neural network with high performance. Given the required system and structure parameters as input, one or multiple solutions can be output by the trained model FreeformNet almost immediately and can be taken as good starting points for further optimization. If further optimization

Table 1.   Comparison of Time Cost Using Different Design Methods

Method	Time for System Generation with Narrow or Moderate FOV		Time for System Generation with Wide FOV
	Starting Point	With Good Imaging Performance	
[19]	Several minutes for system: FOV = 3° × 3°, EFL = 60 mm, $F\#$ = 2, LWIR	Not reported	Not applicable or not reported
[20]	Several minutes for system: FOV = 8° × 8°, EFL = 95 mm, $F\#$ = 1.8, LWIR	2.44 h for system: FOV = 8° × 8°, EFL = 95 mm, $F\#$ = 1.8, LWIR	
[23]	Not reported. No individual starting point design process	Several minutes for system: FOV = 4° × 4°, EFL = 600 mm, $F\#$ = 3, VIS	
[25]	Not reported. No individual starting point design process	5.9 min for system: FOV = 3° × 3°, EFL = 60 mm, $F\#$ = 1.5, LWIR	
[22]	Not reported. No individual starting point design process	About 30 min for system: FOV = 8° × 6°, EFL = 50 mm, $F\#$ = 1.8, LWIR	
[17]	Not reported. A step-by-step design method based on nodal aberration theory: FOV = 4° × 4°, EFL = 600 mm, $F\#$ = 3, VIS	Not reported	
Finding systems from literatures	Maybe tens of minutes or several hours. A large probability that a feasible starting point cannot be found		
[26–30]	Not applicable for freeform system design or not applicable for multiple-solution design		
Our method	Less than 0.003 s per system for starting point generation for the above systems or systems with much wider FOV. If 1000 systems are simultaneously predicted, 1.5×10^{-5} s per system. If further optimization is conducted, it takes only about several seconds in general to obtain good imaging performance for a system with narrow or moderate FOV (similar with the cases shown in this table). For example, optimizing 20 different systems can be done in between 32 and 37 s, less than 2 s for each system		

is conducted for the systems generated by the FreeformNet, it takes several seconds in general to obtain good imaging performance for a system with narrow or moderate FOV. Considering the ultrafast generation of the starting point, it only takes several seconds in total to generate a system with good performance after the design requirements are input into the FreeformNet. The design framework works for generalized freeform systems with advanced system specifications. Human effort for the complicated freeform system design task can be reduced to a minimum and the design efficiency is dramatically improved. This revolutionary design framework opens up a new pathway for the complicated lens design task and will promote the development of the next-generation optical design software. Future work will focus on the generalized one-to-multidesign framework generating systems with various and different kinds of structures and folding geometries as well as a different number of optical elements.

Currently, the freeform surface type of the systems generated by the DNN is the same as the fixed surface type of the systems in the data set. However, the proposed design framework can be further extended to generate systems with different freeform surface types. To achieve this goal, systems using different surface types should be designed during the data set generation process. The surface type can be taken as one of the "parameters" of the system, and different surface types correspond to different values. This value is also taken as one of the inputs to the DNN. The loss function of systems using different surface types during supervised learning should be calculated individually and then summed. Another way is to train different DNNs corresponding to different kinds of surface types. During system generation, systems with different surface types can be generated by changing the surface type input to the DNN or using different DNNs. Then, the effects of different freeform surface types on imaging performance or other metrics can be compared, and the preferred surface type can be selected. Similarly, the proposed design framework can be further extended to the design of transmission optical systems. The refractive indices and Abbe numbers of lens materials can be taken as parameters of the systems and the input of the DNN. Systems using different lens materials (or fictitious glass model) should be designed during the data set generation process. Combining supervised and unsupervised learning, the DNN will be able to output different systems using different inputs of lens materials (or refractive indices and Abbe numbers). Then, for these systems, the imaging performance or other metrics can be compared, and better solutions can be selected. In this way, the replacement of lens materials can be realized to some extent.

In addition, for current design framework, different DNNs are needed for system generation tasks with different folding geometries. For one folding geometry, the design examples given in Section 3 show that the obtained DNN has good

generalization ability for a wide range of system and structure parameter values. This shows the effect and feasibility of the proposed design framework, including automatic data set generation and combined supervised and unsupervised learning. For other folding geometries (or other kinds of systems with different structures, such as systems with a real exit pupil), the same data set generation method can also be used to obtain the corresponding data set with the same basic system structure type and various system and structure parameter values automatically, and the DNN training process can be done in the same way. The corresponding DNNs can be generated easily and used to generate freeform systems. The data set generation and network training process for different kinds of systems can be done automatically and nonstop in computers, work stations, or cloud servers, in order to generate DNNs for these various kinds of systems. Future work will focus on the design framework enabling the exploration and generation of systems with different or best folding geometries, combining more complex training data set and generative adversarial network.

Disclosures. The authors declare no conflicts of interest.

Data Availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

REFERENCES

1. S. Wills, "Freeform optics: notes from the revolution," Opt. Photon. News **28**, 34–41 (2017).
2. J. P. Rolland, M. A. Davies, T. J. Suleski, C. Evans, A. Bauer, J. C. Lambropoulos, and K. Falaggis, "Freeform optics for imaging," Optica **8**, 161–176 (2021).
3. M. Beier, J. Hartung, T. Peschel, C. Damm, A. Gebhardt, S. Scheiding, and A. Tünnermann, "Development, fabrication, and testing of an anamorphic imaging snap-together freeform telescope," Appl. Opt. **54**, 3530–3542 (2015).
4. A. Wilson and H. Hua, "Design and demonstration of a vari-focal optical see-through head-mounted display using freeform Alvarez lenses," Opt. Express **27**, 15627–15637 (2019).
5. P. Benitez, J. C. Miñano, P. Zamora, D. Grabovičkić, M. Buljan, B. Narasimhan, and R. Mohedano, "Advanced freeform optics enabling ultra-compact VR headsets," Proc. SPIE **10335**, 103350I (2017).
6. D. Cheng, J. Duan, H. Chen, H. Wang, D. Li, Q. Wang, and Y. Wang, "Freeform OST-HMD system with large exit pupil diameter and vision correction capability," Photon. Res. **10**, 21–32 (2022).
7. L. Gu, D. Cheng, Y. Liu, J. Ni, T. Yang, and Y. Wang, "Design and fabrication of an off-axis four-mirror system for head-up displays," Appl. Opt. **59**, 4893–4900 (2020).
8. S. Wei, Z. Fan, Z. Zhu, and D. Ma, "Design of a head-up display based on freeform reflective systems for automotive applications," Appl. Opt. **58**, 1675–1681 (2019).
9. Z. Zhuang, J. Parent, P. Roulet, and S. Thibault, "Freeform wide-angle camera lens enabling mitigable distortion," Appl. Opt. **61**, 5449–5456 (2022).
10. A. Toulouse, J. Drozella, P. Motzfeld, N. Fahrbach, V. Aslani, S. Thiele, and A. M. Herkommer, "Ultra-compact 3D-printed wide-angle cameras realized by multi-aperture freeform optical design," Opt. Express **30**, 707–720 (2022).
11. Y. Xie, X. Mao, J. Li, F. Wang, P. Wang, R. Gao, and R. Dong, "Optical design and fabrication of an all-aluminum unobscured two-mirror freeform imaging telescope," Appl. Opt. **59**, 833–840 (2020).
12. J. Zhang, Y. Zheng, C. Lin, Y. Han, and Y. Shi, "Large numerical aperture off-axis reflective telescope design with a freeform mirror based on aperture expansion strategy," Appl. Opt. **62**, 1510–1520 (2023).
13. J. Reimer, A. Bauer, K. P. Thompson, and J. P. Rolland, "Freeform spectrometer enabling increased compactness," Light Sci. Appl. **6**, e17026 (2017).
14. B. Zhang, Y. Tan, G. Jin, and J. Zhu, "Imaging spectrometer with single component of freeform concave grating," Opt. Lett. **46**, 3412–3415 (2021).
15. L. Feng, L. Wei, Y. Nie, M. Huang, L. Yang, X. Fu, and J. Zhou, "Design of a compact spectrometer with large field of view based on freeform surface," Opt. Commun. **444**, 81–86 (2019).
16. K. Fuerschbach, J. P. Rolland, and K. P. Thompson, "A new family of optical systems employing φ-polynomial surfaces," Opt. Express **19**, 21919–21928 (2011).
17. A. Bauer, E. M. Schiesser, and J. P. Rolland, "Starting geometry creation and design method for freeform optics," Nat. Commun. **9**, 1756 (2018).
18. Y. Dai, Y. Liu, F. Shen, C. Kuang, Z. Zheng, and R. Wu, "Calculation of aberration fields for freeform imaging systems using field-dependent footprints on local tangent planes," Appl. Opt. **61**, 9576–9582 (2022).
19. T. Yang, J. Zhu, X. Wu, and G. Jin, "Direct design of freeform surfaces and freeform imaging systems with a point-by-point three-dimensional construction-iteration method," Opt. Express **23**, 10233–10246 (2015).
20. T. Yang, G. F. Jin, and J. Zhu, "Automated design of freeform imaging systems," Light Sci. Appl. **6**, e17081 (2017).
21. Y. Nie, H. Thienpont, and F. Duerr, "Multi-fields direct design approach in 3D: calculating a two-surface freeform lens with an entrance pupil for line imaging systems," Opt. Express **23**, 34042–34054 (2015).
22. B. Zhang, G. Jin, and J. Zhu, "Towards automatic freeform optics design: coarse and fine search of the three-mirror solution space," Light Sci. Appl. **10**, 65 (2021).
23. F. Duerr and H. Thienpont, "Freeform imaging systems: Fermat's principle unlocks 'first time right' design," Light Sci. Appl. **10**, 95 (2021).
24. J. C. Miñano, P. Benítez, W. Lin, J. Infante, F. Muñoz, and A. Santamaría, "An application of the SMS method for imaging designs," Opt. Express **17**, 24036–24044 (2009).
25. W. Wu, H. Wang, G. Jin, and J. Zhu, "Fast automatic design method for freeform imaging systems through system construction and correction," Opt. Lett. **45**, 5140–5143 (2020).
26. G. Côté, J. F. Lalonde, and S. Thibault, "Extrapolating from lens design databases using deep learning," Opt. Express **27**, 28279–28292 (2019).
27. G. Côté, J. F. Lalonde, and S. Thibault, "Deep learning-enabled framework for automatic lens design starting point generation," Opt. Express **29**, 3841–3854 (2021).
28. G. Côté, Y. Zhang, C. Menke, J. F. Lalonde, and S. Thibault, "Inferring the solution space of microscope objective lenses using deep learning," Opt. Express **30**, 6531–6545 (2022).
29. T. Yang, D. Cheng, and Y. Wang, "Direct generation of starting points for freeform off-axis three-mirror imaging system design using neural network based deep-learning," Opt. Express **27**, 17228–17238 (2019).
30. W. Chen, T. Yang, D. Cheng, and Y. Wang, "Generating starting points for designing freeform imaging optical systems based on deep learning," Opt. Express **29**, 27845–27870 (2021).
31. Q. Sun, C. Wang, F. Qiang, D. Xiong, and H. Wolfgang, "End-to-end complex lens design with differentiable ray tracing," ACM Trans. Graph. **40**, 71 (2021).