

## Micro-LED backlight module by deep reinforcement learning and micro-macro-hybrid environment control agent

CHE-HSUAN HUANG,<sup>1,2</sup> YU-TANG CHENG,<sup>2</sup> YUNG-CHI TSAO,<sup>3</sup> XINKE LIU,<sup>1,5</sup> AND HAO-CHUNG KUO<sup>2,4,6</sup>

<sup>1</sup>College of Materials Science and Engineering, Guangdong Research Center for Interfacial Engineering of Functional Materials, Institute of Microelectronics (IME), Shenzhen University, Shenzhen 518060, China

<sup>2</sup>Department of Photonics & Institute of Electro-Optical Engineering, College of Electrical and Computer Engineering, National Yang Ming Chiao Tung University & National Chiao Tung University, Hsinchu 30010, Taiwan

<sup>3</sup>Department of Computer Science, University of Liverpool, Liverpool, UK

<sup>4</sup>Honhai Semiconductor Research Center, Taipei, Taiwan

<sup>5</sup>e-mail: xkliu@szu.edu.cn

<sup>6</sup>e-mail: hckuo@faculty.nctu.edu.tw

Received 25 August 2021; revised 21 September 2021; accepted 23 September 2021; posted 29 September 2021 (Doc. ID 441188); published 7 January 2022

**This paper proposes a micro-LED backlight module with a distributed Bragg reflector (DBR) structure to achieve excellent micro-LED backlight module quality and uses deep reinforcement learning (DRL) architecture for optical design. In the DRL architecture, to solve the computing environment problems of the two extreme structures of micro-scale and macro-scale, this paper proposes an environment control agent and virtual-realistic workflow to ensure that the design environment parameters are highly correlated with experimental results. This paper successfully designed a micro-LED backlight module with a DBR structure by the abovementioned methods. The micro-LED backlight module with a DBR structure improves the uniformity performance by 32% compared with the micro-LED backlight module without DBR, and the design calculation time required by the DRL method is only 17.9% of the traditional optical simulation. © 2022 Chinese Laser Press**

<https://doi.org/10.1364/PRJ.441188>

### 1. INTRODUCTION

With the evolution of semiconductor process technology in recent decades, III-V compound semiconductors with excellent direct bandgap features have been extensively employed in various fields such as high-power electronics, lighting, and display technology [1–5]. Light-emitting diodes (LEDs) are the most widely utilized among them in display technology [6,7]. The backlight display technology with LED as the display light source brings superior energy application efficiency and better color quality than traditional display solutions and guides the entire display industry to a thinner design [8–10]. In the last ten years, LED backlight display technology, which was introduced in 2010, has been widely embraced in the global marketplace. There is a more precise blueprint for the future: higher energy application efficiency, higher color saturation, and the micro-LED backlight module with ultrahigh contrast control capability [11–13]. The micro-LED backlight module offers many advantages to foresight displays, but it also brings many novel challenges and innovative opportunities. For example, the quantum-confined Stark effect generated by the thin epitaxial design in the process stage is alleviated using the

investigation of the nanosphere photomask etching at NCTU SCLAB [14]. Because of the chip size ratio, the lighting angle design area is limited, posing novel low-light-loss optical design challenges [15]. Quite a few research results have also measured the indirect physical parameters of the overall backlight design [16]. Designing a micro-LED module with both high uniformity and low power consumption remains a critical challenge today. In addition, since quantum dot materials can significantly enhance the color quality of displays, quantum dot color conversion technology is also an extremely important and widely discussed technology in forward-looking display technology [17,18].

Many mature approaches and physical concepts of optical design have been established and utilized in various research efforts during the last half-century. In the field of LED, there is a great deal of research and analysis on physical models of different scales [19–27]. However, since the specific manifestations of macro-scale and micro-scale exhibit particle and wave, it is often impossible to design both macro-scale and micro-scale at the same time when performing optical design. This phenomenon has become a challenging issue

for micro-LED modules that pursue ultrathin design since the micro-scale of the crystal grain structure and the macro-scale of the backlight module are both present in the micro-LED module [28,29]. This paper constructs a set of the optical programming process and introduces an environmental control agent technology to control the macro- and micro-scale. Furthermore, due to the traditional optimization calculation methods such as differential evolution and genetic algorithms, all are rule-based algorithms, which are characterized by interpretability and the use of fewer hyperparameters. Although this method is convenient for developers to use, it is more difficult to apply to topics with greater variability. On the contrary, a data-driven algorithm driven by a large amount of data has many hyperparameters to solve various changing problems. In order to achieve a highly adaptable solution, and to avoid overly sensitive search processes that lead to the inability to find the global maximum or minimum value, this study provides artificial intelligence deep reinforcement learning technology and employs Google DeepMind's double DQN (DDQN) architecture as the core network to create artificial intelligence optical design agents for optimum optical design [30–33]. Deep reinforcement learning (DRL) and environment control agent technology are fully implemented in this research [34–40]. In addition, based on the inference results of the artificial intelligence model, it successfully produces ultrathin, high-efficiency, high-uniformity micro-LED modules.

## 2. EXPERIMENT AND ALGORITHM DESIGN

### A. Introduction of the Micro-LED Module and Optical Events

In the application of foresight displays, the micro-LED backlight module is the current mainstream developing technology. A complete micro-LED backlight module is shown in Fig. 1. The main structure contains the substrate, the receiver plane; the LED array, which has a wavelength spectrum with Gaussian distribution in which the peak of the spectrum is at 445 nm and the width of spectrum is 18 nm; and the structure design on the LED. In this

paper, a distributed Bragg reflector (DBR) structure is selected as the structure to control the light-emitting angle of the LED as shown in Fig. 1(b), covering each LED under the LED array. In this module structure, there are four main types of optical events when the light is emitted from the LED to the receiver plane.

(1) Transmitting through the DBR structure into the receiver plane: The angle of the light emitted by the LED and incident into the DBR structure does not reach the total internal reflection angle. Therefore, it follows Snell's law and enters the receiver plane through refraction.

(2) Direct emitting into the receiver plane: The light generated by the LED is incident directly into the receiver plane without going through any objects.

(3) Reflect light from the DBR structure: The reflected light is generated when the angle of the light incident to DBR reaches the total internal reflection angle (TIR) and enters the receiver plane.

(4) Reflect light from the receiver plane: Reflected light is generated when light enters the receiver plane.

Due to the four abovementioned optical events, the light generates many reflection events between the receiver plane, the DBR, and the substrate. If a reflection is not incident on a perfectly reflective surface, it causes energy loss in the optical system. To reduce the loss, in the module design, a high-reflectivity coating is used for coating on the substrate surface as shown in Fig. 1(c), and the diffuser plate with an etched pattern is used as the receiver plane as shown in Fig. 1(d). To reduce the energy loss, the substrate is coated by a thick  $\text{TiO}_2$  layer.

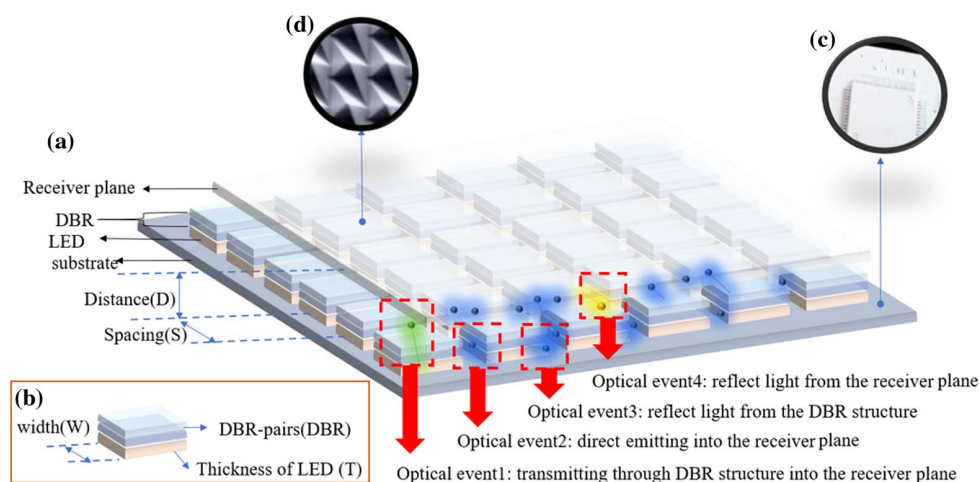
After completing the frame design of the module, the important variables that should be taken into consideration in designing the micro-LED backlight module are indicated in Figs. 1(a) and 1(b).

(1) Distance ( $D$ ): distance between substrate and receiver plane

(2) Space ( $S$ ): spacing between LED chips

(3) Width ( $W$ ): width of the LED

(4) DBR pairs (DBR): the number of DBR pairs



**Fig. 1.** (a) Schematic diagram of micro-LED backlight module; (b) schematic diagram of LED with DBR structure; (c) highly-reflective surface substrate; (d) etching structure of the receiver.

This study will develop an optimization approach for the architecture's four-variable design.

**B. Overall Design Workflow**

As shown in Fig. 2, we provide a comprehensive set of design methodologies for achieving perfect simulation and optimization results. This design process is divided into two major modules. One is the optimization engine for deep reinforcement learning as shown in Fig. 2(a). The task of the module is to optimize the design of artificial intelligence design agents in a controlled environment. The other core module is simulation verification, such as Fig. 2(b). This verification aims to ensure that the environment of the design model is consistent with the actual field. Thus, the gap is small enough to have a sufficient degree of reliability between the analog design and the implementation.

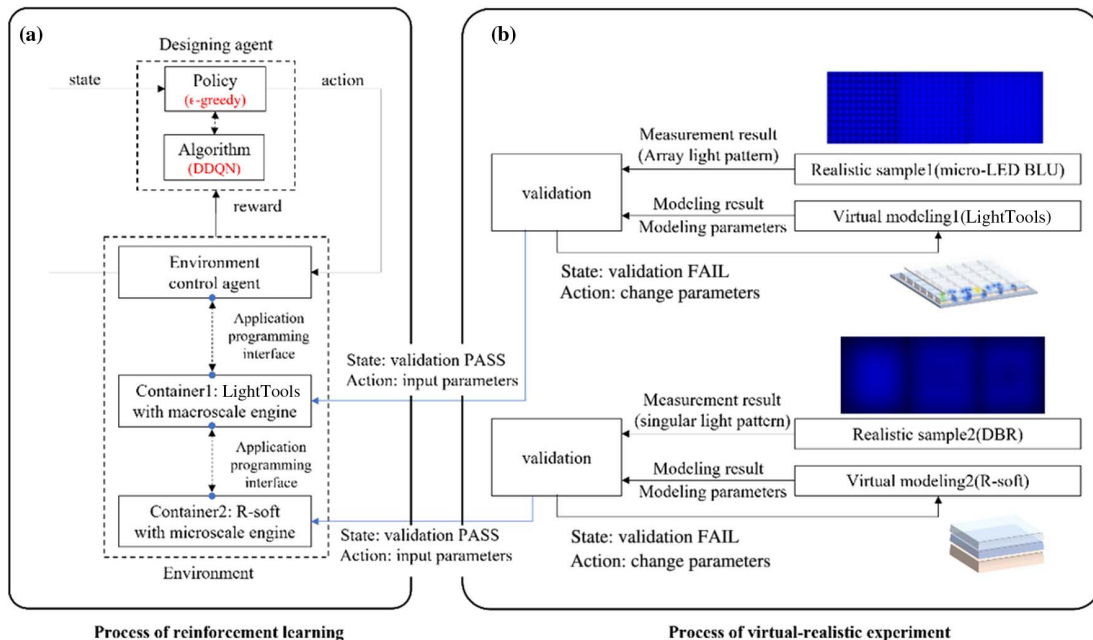
Figure 2(a) shows that two core modules make up the optimization engine, namely, the designing agent and the environment. After the designing agent issues a specified action to the environment for execution each time, the environment reports the environment state and reward to the designing agent. The algorithm network in the designing agent learns deep neural network (DNN) and guides the policy function to give the following action. After multiple iterations when the reward is saturated, optical design optimization can be obtained. The neural network model used in this paper is the double deep Q-learning neural (DDQN) network model published by Google in 2015, with  $\epsilon$ -greedy as the core of the optimization task.

Another part that has a significant impact on the design results is the environment in the reinforcement learning architecture. Since the micro-LED backlight module has both a nano-level structure and a millimeter-level module structure, macro-scale and micro-scale operations must be taken into

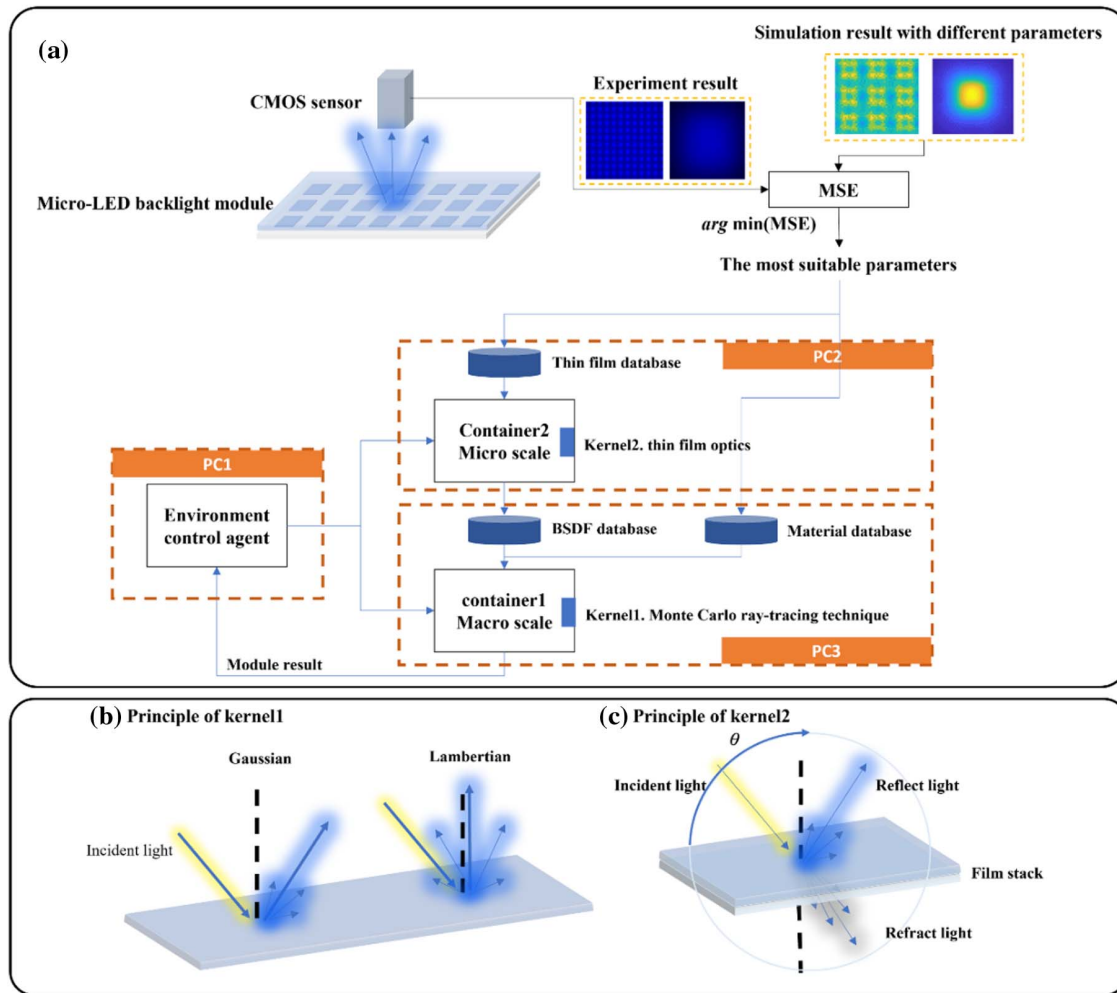
account. Wave optics, in general, dominate the calculation of the micro-scale; this paper uses Synopsys' R-soft software for calculations. On the other hand, macro-scale calculation generally uses geometric optics for optical calculation; this paper adopts Synopsys' LightTools software for calculations. Although there are sophisticated computer-aided engineering (CAE) tools that can handle micro- and macro-scale problems separately, the micro-LED backlight module must deal with two extreme dimensions at once, and thus we add the notion of environment control agent and container into the environment as shown in the environment in Fig. 2(a). We set the environment and database required for the operation of LightTools and R-soft separately in the environment design, and then we utilize the environment agent program to perform collaborative scheduling of the two container tasks. Through the environment control agent and independent container settings, we can effectively solve extreme-scale computing problems.

After completing the environment architecture based on the environment control agent and containers, as shown in Fig. 2(b), to accomplish high-reliability simulation, we undertake two sets of extreme environment individual parameter verifications between the virtual environment and the actual scene, which we term a virtual-realistic experiment. In the virtual-realistic experiment, we compare the completed micro-LED array sample image with the LightTools simulation image and quantify the difference through validation to obtain the best combination of realistic parameters. Similarly, we also use the completed micro-LED with DBR structure to compare with R-soft's simulation result image and quantify the difference through validation to obtain the best realistic parameter combination.

This paper will use the optimized architecture of Fig. 2(a) and the immersive environment setting of Fig. 2(b) to design a highly reliable micro-LED backlight module.



**Fig. 2.** Workflow for optimizing micro-LED backlight module: (a) the process of deep reinforcement learning and (b) the process of the virtual-realistic experiment.



**Fig. 3.** (a) Workflow of environment control agent and schematic diagram of the virtual-realistic experiment; (b) principle of kernel1: Gaussian and Lambertian reflection; (c) principle of kernel2: BSDF properties.

### C. Establishing an Environment Control Agent

The workflow of establishing environment is as shown in Fig. 3(a). To begin, we implement the first batch of micro-LED backlight module prototypes, capture images using a CMOS sensor, and use single chip and chip array images as a baseline for optimizing simulation parameters. On the other hand, the optical simulation models of the DBR and micro-LED backlight module were built using R-soft and LightTools, respectively. There are six groups of experimental combinations:

- (1) single GaN without DBR in the backlight module,
- (2) single GaN with 2.5-pair DBR in the backlight module,
- (3) single GaN with 5.5-pair DBR in the backlight module,
- (4) GaN array without DBR in the backlight module.

In the simulation, the backlight module's bottom reflectivity and receiver reflectivity are the essential systematic variables. To establish close-to-accurate environmental parameters, the optical reflection characteristics of actual samples generally have scattering characteristics. The scattering mode of reflected light is roughly divided into Gaussian type (the angle between the main direction of the reflected light and the normal is

consistent with the incident angle) and Lambertian type (the reflected light at any angle of incidence is scattered along the normal direction) as shown in Fig. 3(b). We perform simulation calculations on the bottom reflectivity and receiver reflectivity for two scattering modes and five variables (Gaussian type 0°, Gaussian type 5°, Gaussian type 10°, Gaussian type 15°, and Lambertian type). The schematic diagram of the scattering model is as shown in Fig. 3(b). The variable intervals of bottom reflectivity and receiver reflectivity are 93%–99% and 20%–40%, respectively; after completing the simulation calculation, the result matrix output of the simulation calculation and the image collected by CMOS are calculated for validation as shown in Fig. 2(b). The calculation is calculated by mean-square error as Eq. (1):

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (\text{ImgS}_i - \text{ImgC}_i)^2, \quad (1)$$

where  $N$  is the total number of pixels in the image,  $i$  is the pixel index,  $\text{ImgS}$  is the image generated by simulation, and  $\text{ImgC}$  is the image collected through the CMOS sensor.

After completing the MSE calculation of the virtual-realistic experiment, we take the parameter combination when the MSE is minimum as the basic parameters of the environment model. The equation is expressed as Eq. (2):

$$\text{Environment parameters} = \arg \min(\text{MSE}(\text{unput\_parameters})). \quad (2)$$

After completing the value of the environment parameters, we set up the environments of the micro-scale and macro-scale in PC2 and PC3 computers, respectively. PC2 constructs a computing core based on the micro-scale, and its environment is called container2. The computing kernel in container2 is thin-film optics, and the main database is the thin-film database, which contains various film material characteristic values such as refractive index  $n$  and extinction coefficient  $\alpha$ . On the other hand, PC3 builds a macro-scale-based computing core, and its environment is called container1. The computing kernel in container1 is the Monte Carlo ray-tracing technique, and the main databases are the material database and bidirectional scattering distribution function (BSDF) database. The calculation of container2 generates this database as shown in Fig. 3(c). When incident light enters the film surface, reflection and refraction occur. Since the film surface is not perfectly flat, it will produce scattered energy distribution no matter whether it is reflection or refraction. The BSDF database records the intensity and angular distribution of reflection scattering and refraction scattering created by rays of various incident angles incident on the film stack. The optical problems of micro-scale and macro-scale, respectively, can be handled by the calculations of container1 and container2. After completing the PC2 and PC3 environment construction, the two computing engines cannot transfer sensitively. This paper introduces a program of environment control agents to PC1 such as Fig. 3(a). This program transfers the instructions from the reinforcement learning to PC2 and PC3 work scheduling; this method can simultaneously assign the action from reinforcement learning to PC2 and PC3 to perform parallel processing and optimization processing or directly organize the state from the micro-scale and macro-scale into a unified format and output to the reinforcement learning to perform operations.

#### D. Definition of Action Function, State Function, and Reward Function

It can be seen from Fig. 2(a) that in reinforcement learning, there are three functions, action, state, and reward, that should be defined between the environment and the designing agent. Referring to the design parameters in the schematic diagrams of Figs. 1(a) and 1(b), we define the action function as Table 1. For the five kinds of variations to adjust up and down, we defined the action functions of  $a_1$ – $a_{10}$ . In addition, we also define the state function of  $S_1$  to  $S_5$  for the five variation states as shown in Table 2. Finally, referring to the definition of uniformity of the backlight module, we have set three different reward functions as Eqs. (3)–(5):

$$\text{reward function1} = (\text{Uniformity new} - \text{Uniformity old})/100, \quad (3)$$

**Table 1. Definition List of Action**

Action No.	Action Definition	Variation Symbol
$a_1$	Add distance	$D$
$a_2$	Reduce distance	$D$
$a_3$	Add width	$W$
$a_4$	Reduce width	$W$
$a_5$	Add spacing	$S$
$a_6$	Reduce spacing	$S$
$a_7$	Add thickness	$T$
$a_8$	Reduce thickness	$T$
$a_9$	Add DBR pairs	DBR
$a_{10}$	Reduce DBR pairs	DBR

**Table 2. Definition List of State**

State No.	State Definition	Variation symbol
$S_1$	Value of distance	$D$
$S_2$	Value of spacing	$W$
$S_3$	Value of DBR pairs	$S$
$S_4$	Value of thickness	$T$
$S_5$	Value of width	DBR

$$\text{reward function2} = (\text{Uniformity} - 75)^3/1000, \quad (4)$$

$$\text{reward function3} = (\text{Uniformity} - 79)^3/1000. \quad (5)$$

Among them, Eq. (3) adopts the floating reward computer system, while Eqs. (4) and (5) adopt the computer system of fixed uniformity comparison method. First, we need to define the environment's boundary value before executing the reinforcement learning calculation. The environment parameters are as Table 3.

#### E. Double Deep Q-learning Neural Network

As shown in Fig. 2(a), the double DQN (DDQN) network is used as the optimized algorithm in this paper. The detailed calculation process is shown in Fig. 4. The action ( $a_t$ ) is issued to the environment control agent through the  $\epsilon$ -greedy policy calculation. After the control and calculation of the environment control agent, the  $t$ th and the  $(t + 1)$ th states,  $S_t$ ,  $S_{t+1}$ , action and reward values are reported. The state, action, and reward values of each iteration are recorded through the replay memory. Then through two neural networks of  $Q$  network ( $\theta_t$ ) and target  $Q$  network ( $\theta'_t$ ), we calculate the loss function to determine the new parameter update. Target  $Q$  and loss function are defined as in Eqs. (6) and (7):

$$\text{Target } Q = R_{t+1} + \gamma Q(S_{t+1}, \text{Max } Q(S_{t+1}, a; \theta_t), \theta'_t), \quad (6)$$

where  $R_{t+1}$  is the reward value of  $(t + 1)$ th state, and  $\gamma$  is the model hyperparameters, and

**Table 3. Definition List for Range of Parameters**

No.	Parameters	Variation Symbol	Range
1	Distance from receiver to LED	$D$	100–180 $\mu\text{m}$
2	LED width	$W$	10–100 $\mu\text{m}$
3	LED spacing	$S$	300–700 $\mu\text{m}$
4	Thickness of LED	$T$	5–55 $\mu\text{m}$
5	DBR-BSDF pairs	DBR	4.5–9.5 pairs



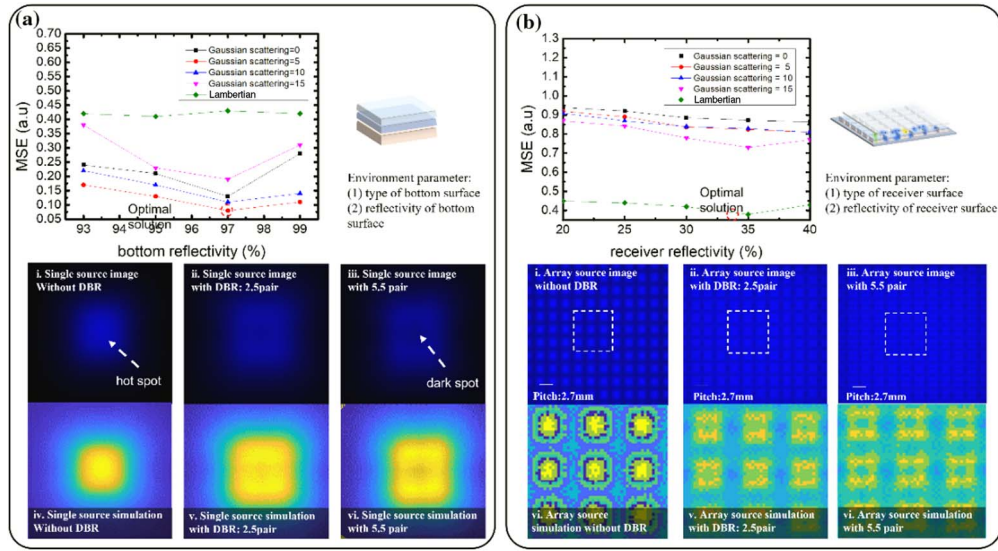


Fig. 5. Virtual-realistic experiment: (a) single light pattern analysis; (b) module pattern analysis.

functions and parameter optimization boundaries in reinforcement learning, such as Tables 3 and 4. The design of reward function1 is defined as “strict decreasing,” which means it is required that each loss must be lower than the previous one. Besides, reward function2 and reward function3 are defined as the reward function compared with a fixed reference value. Since reward functions2 and 3 do not need to follow strict incremental rules, they can avoid overfitting and have a more stable convergence result.

The experimental results are shown in Figs. 6(a), 6(b), and 6(c), respectively, representing uniformity for every iteration with reward functions 1, 2, and 3. The iteration results using reward function1 are volatile initially, and results with a uniformity of 0 often appear. Although it does not reach the more stable operation optimization until after 250 iterations,

the value can still discover a significant attenuation instability result in the stable area. The best result appears in the 289th iteration. The result is shown in Fig. 6(d). According to the value of the average  $Y = 0$  mm cutoff line, the uniformity calculation is written as Eq. (8):

$$\text{Uniformity} = \min(L[25\%, 75\%]) / \max(L[25\%, 75\%]), \quad (8)$$

where  $L$  is the cutoff line when  $Y = 0$  mm and  $L[25\%, 75\%]$  means to take the value of the 25th to 75th interval of the  $L$  line.

Figure 6(d) shows that the state of the best solution when calculating reward function1 is  $S_{b1}(D = 0.18$  mm,  $W = 0.04$  mm,  $S = 0.4$  mm,  $T = 0.02$  mm, DBR = 5.5 pairs), and the uniformity at this time is 83.97%. In addition, in Figs. 6(b) and 6(c) when using reward function2 and reward

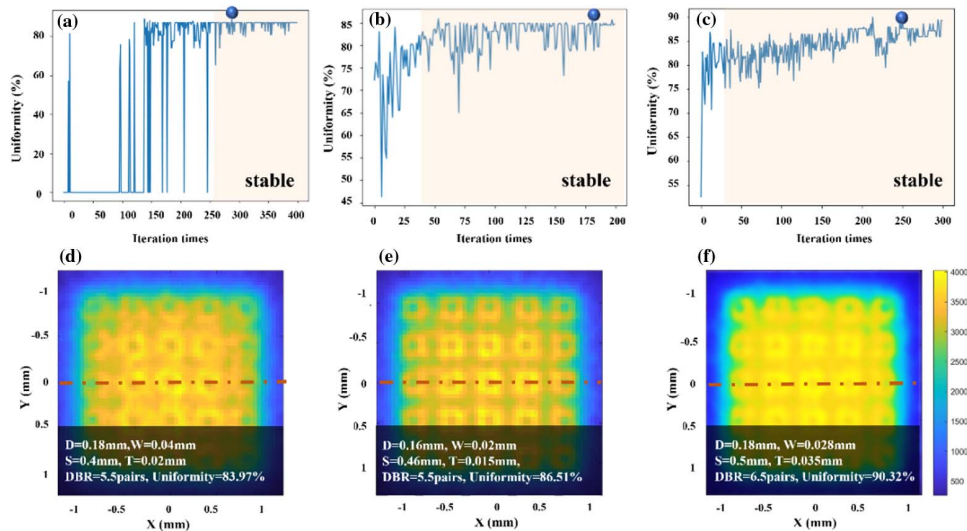
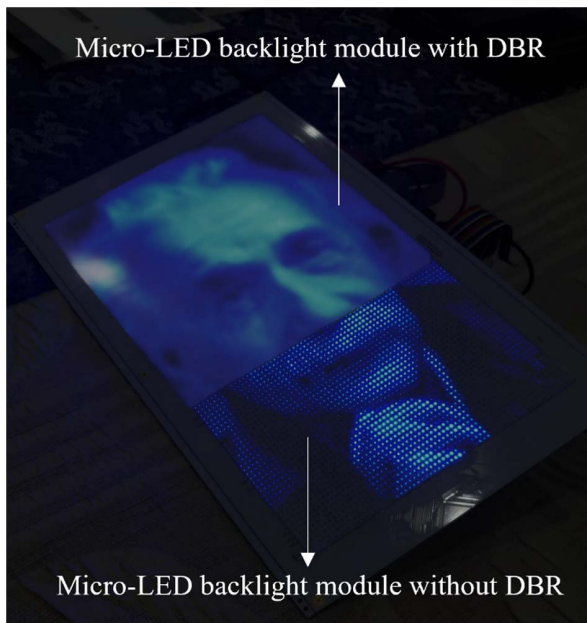


Fig. 6. Result of reinforcement learning: (a) uniformity for every iteration with reward function1; (b) uniformity for every iteration with reward function2; (c) uniformity for every iteration with reward function3; (d) the best result by reward function1; (e) the best result by reward function2; (f) the best result by reward function3.



**Fig. 7.** Demonstration of the micro-LED backlight module.

function3, they can reach a stable state with a small number of iterations and are less prone to computing crashes. In Figs. 6(b) and 6(c), the best solutions can be found when the numbers of iterations are 183 and 249. The results of the best solutions are shown in Figs. 6(e) and 6(f). The state is  $S_{b2}(D = 0.16 \text{ mm}, W = 0.02 \text{ mm}, S = 0.46 \text{ mm}, T = 0.015 \text{ mm}, \text{DBR} = 5.5 \text{ pairs})$  and  $S_{b3}(D = 0.18 \text{ mm}, W = 0.028 \text{ mm}, S = 0.5 \text{ mm}, T = 0.035 \text{ mm}, \text{DBR} = 6.5 \text{ pairs})$ , while the uniformity is 86.51% and 90.32%, respectively.

### C. Result of Real Sample

After completing the optimized model design above, we made a prototype demonstration according to the  $S_{b3}$  design. Using the DBR structure can reduce the overall graininess as shown in Fig. 7.

## 4. DISCUSSION

According to the above experimental results, the designing agent designed by the method proposed in this paper can

effectively optimize the optical design, but more physical characteristics still need to be summarized. We will focus on the following points and discuss them.

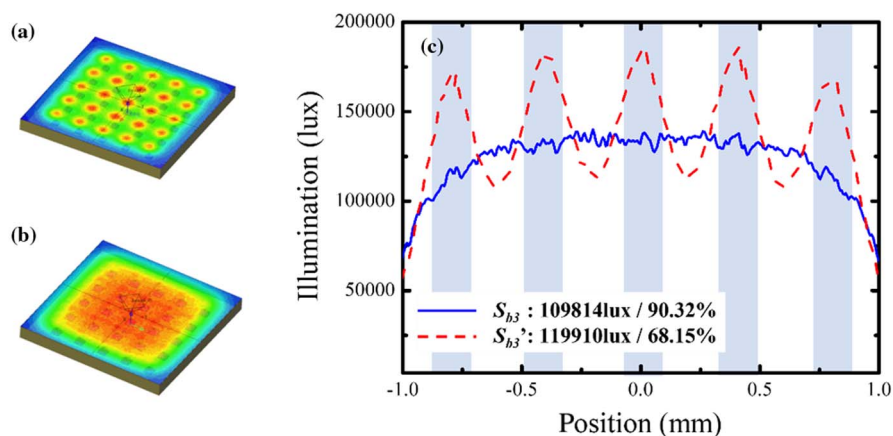
- (1) Inference of DBR structure
- (2) Energy loss mechanism of the micro-LED backlight module
- (3) Designing rule of the micro-LED backlight module
- (4) Work efficiency of reinforcement learning

### A. Inference of DBR Structure

First, we set the state  $S_{b3}(D = 0.18 \text{ mm}, W = 0.028 \text{ mm}, S = 0.5 \text{ mm}, T = 0.035 \text{ mm}, \text{DBR} = 6.5 \text{ pairs})$  of the best calculation result as the experimental group with DBR, and we set the DBR in  $S_{b3}$  to 0 as the control group,  $S'_{b3}(D = 0.18 \text{ mm}, W = 0.028 \text{ mm}, S = 0.5 \text{ mm}, T = 0.035 \text{ mm}, \text{DBR} = 0 \text{ pairs})$ ; the comparison result is shown in Fig. 8. Figures 8(a) and 8(b) are the optical class distributions of  $S'_{b3}$  and  $S_{b3}$ , respectively. It is obvious that  $S'_{b3}$  without DBR structure has apparent hot spot distribution, while  $S_{b3}$  with DBR structure has no apparent hot spot distribution. Figure 8(c) shows the energy distribution results of  $S_{b3}$  and  $S'_{b3}$  at  $Y = 0 \text{ mm}$ . From this result, it can be calculated that the illuminance of  $S_{b3}$  is 109,814 lux, the uniformity is 90.32%, and the illuminance of  $S'_{b3}$  is 119,910 lux, the uniformity is 68.15%. The DBR structure will reflect a large amount of light emitted by the chip to the backplane, resulting in a certain degree of energy loss. The energy loss in this design is only 8.5%. In addition, the reflected light also improves the energy loss between the chips. The energy of the space enables the light energy originally distributed above the wafer to be guided to the dark band between the wafers. Overall, the DBR structure improves the uniformity by 32%.

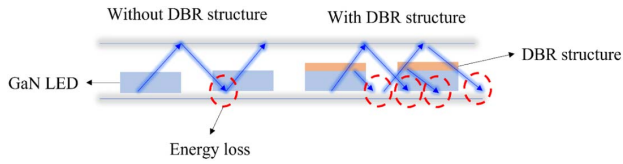
### B. Energy Loss Mechanism of the Micro-LED Backlight Module

We can observe from the above results that DBR can optimize the light energy of the dark band by reflecting the light of the chip. In the process, each reflection is often accompanied by energy loss as shown in Fig. 9. The light loss generated by the light reflected by the DBR structure to the backplane is the main energy loss mechanism of the micro-LED backlight module, and the other is the light loss caused by the light reflected by the receiver to the backplane as shown in Fig. 1.



**Fig. 8.** Influence of DBR structure: (a) light pattern with state  $S'_{b3}$ , (b) light pattern with state  $S_{b3}$ , and (c) uniformity of  $S_{b3}$  and  $S'_{b3}$ .





**Fig. 9.** Schematic of resonant loss for the micro-LED backlight module.

The structure of the micro-LED backlight module uses a high-thickness and highly-reflective filler surface reflective paint structure from the beginning of the architecture due to the consideration of optical event3 and optical event4. This structure is concluded to have a reflectivity of up to 97% in the results of the virtual-realistic experiment, allowing the optical loss of the entire system to reduce to within 10%.

**C. Designing Rule of the Micro-LED Backlight Module**

Table 5 provides the optimization results under different reward function designs.  $S_{b1}$  ( $D = 0.18$  mm,  $W = 0.04$  mm,  $S = 0.4$  mm,  $T = 0.02$  mm, DBR = 5.5 pairs),  $S_{b2}$  ( $D = 0.16$  mm,  $W = 0.02$  mm,  $S = 0.46$  mm,  $T = 0.015$  mm, DBR = 5.5 pairs), and  $S_{b3}$  ( $D = 0.18$  mm,  $W = 0.028$  mm,  $S = 0.5$  mm,  $T = 0.035$  mm, DBR = 6.5 pairs). Further analysis of the results of these three sets of the state, except for  $S_{b2}$ , all gets the conclusion that optical distance ( $D$ ) is 0.18 mm; this value is the maximum value in the optimization parameter range we set, and this result is consistent with the designing rule of the traditional LCD backlight module. On the other hand, the main factor that affects the result of  $S_{b1}$ ,  $S_{b2}$ , and  $S_{b3}$  is the space energy between the chips. As shown in Fig. 8, the ability to transfer the energy at the center of the chip to a large angle will help suppress the hot spot on the chip and the dark spot phenomenon between chips. We take out  $W$  and  $T$  from  $S_{b1}$ ,  $S_{b2}$ , and  $S_{b3}$  and calculate the geometrical factor  $f$  as the Eq. (9) based on the lateral and forward light-emitting surface area:

$$f = \frac{4T}{W}, \tag{9}$$

where  $T$  is the thickness of the LED and  $W$  is the width of the LED. The symbols of parameters are as shown in Fig. 1.

The  $f$  values of  $S_{b1}$ ,  $S_{b2}$ , and  $S_{b3}$  are respectively 2, 3, and 5. According to previous research, a larger  $f$  value can provide a larger light-emitting angle, increasing the occurrence ratio of optical event2 at the chip pitch. The DBR parameter of  $S_{b3}$

**Table 5. Best Uniformity for Different Reward Functions**

Best State	Reward Formula	Parameters	Number of Iterations	Best Uniformity
$S_{b1}$	(Uniformity new - Uniformity old)/100	$D = 0.18$ mm, $W = 0.04$ mm, $S = 0.4$ mm, $T = 0.02$ mm, DBR = 5.5 pairs	289	83.97%
$S_{b2}$	(Uniformity - 75) <sup>3</sup> /1000	$D = 0.16$ mm, $W = 0.02$ mm, $S = 0.46$ mm, $T = 0.015$ mm, DBR = 5.5 pairs	183	86.51%
$S_{b3}$	(Uniformity - 79) <sup>3</sup> /1000	$D = 0.18$ mm, $W = 0.028$ mm, $S = 0.5$ mm, $T = 0.035$ mm, DBR = 6.5 pairs	249	90.32%

**Table 6. Work Efficiency of Designing Agent**

	Number of Iterations	Time (h)	Optimal Result
Entire loop	106,920	297	Uniformity: 89.61%
Designing agent	16,000	53.3	Uniformity: 90.32%

is 6.5 pairs, while the DBR parameter of  $S_{b1}$  and  $S_{b2}$  is 5.5 pairs. It has higher reflectivity, which can reduce the incidence of optical event1, reduce optical event1 and increase optical event2 to achieve significant energy transfer so that  $S_{b3}$  has the best optical uniformity design result.

**D. Work Efficiency of Reinforcement Learning**

In this study, a high-efficiency designing agent is established through reinforcement learning and environment control agent. To ensure that the solution obtained by the designing agent within the parameter limit interval is the best, step is added to the lower limit interval defined in the  $D$ ,  $W$ ,  $S$ ,  $T$ , and DBR in Table 3 to perform a full-variable loop operation, entire loop. The steps of the parameters  $D$ ,  $W$ ,  $S$ ,  $T$ , and DBR are 10, 10, 50, 10, and 0.5, respectively. Thus, we get that  $D$ ,  $W$ ,  $S$ ,  $T$ , and DBR have 18, 10, 9, 6, and 11 variable values, respectively. There are 106,920 combinations in total. The total calculation time is 297 h. The best result obtained is uniformity 89.61%. On the other hand, the design of reinforcement learning is adopted. The agent performs a total of 16,000 calculations and takes 53.3 h. The best result obtained is uniformity 90.32% as shown in Table 6. Thus, the designing agent only needs 17.9% of the entire loop in optimized solution time.

According to Table 6, although the entire loop seems to describe the complete solution plane, since the steps of each parameter are fixed values, the solution plane depicted by the entire loop is a plane depiction of a discrete solution. On the other hand, it adopts the designing agent of reinforcement learning that has the characteristics of a deep learning solution, which is to narrow the search range to find the best solution. Therefore, it can provide more efficient and superior optimization solutions.

**5. CONCLUSION**

Traditional optical design methods generally emphasize the accuracy of model parameters. There are many studies on measuring detailed parameters; however, considering the influence of extreme scales at the same time can be challenging. The environment control agent technology introduced in this research

efficiently integrated micro-scale and macro-scale models into a single set of optimization models, eliminating all calculation barriers imposed by different calculation scales. A high-efficiency model optimization design method for micro-LED backlight modules is also provided by the reinforcement learning model.

In comparison to the entire loop mode, a higher resolution optimal position on the solution plane can be determined; hence, the designing agent's optimization result is somewhat better than that of the entire loop. Besides, compared with the multi-parameter discrete loop solution, the designing agent can find a more accurate solution, and it only needs 17.9% of the entire loop time to find a solution. This research successfully developed a micro-LED backlight module with a DBR structure as the backlight uniformity optimization element using a novel multi-environment control approach and reinforcement learning framework. Its uniformity has improved by 32% when compared to a micro-LED backlight module without a DBR structure.

**Funding.** Honhai Research Institute.

**Acknowledgment.** The authors thank LEDA-technology Co., Ltd. for technical support in advising of programing and deep learning.

**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

- I. C. Kizilyalli, A. P. Edwards, H. Nie, D. Disney, and D. Bour, "High voltage vertical GaN p-n diodes with avalanche capability," *IEEE Trans. Electron Devices* **60**, 3067–3070 (2013).
- S. Han, S. Yang, and K. Sheng, "High-voltage and high- $I_{ON}/I_{OFF}$  vertical GaN-on-GaN Schottky barrier diode with nitridation-based termination," *IEEE Electron Device Lett.* **39**, 572–575 (2018).
- J. Wang, L. Cao, J. Xie, E. Beam, R. McCarthy, C. Youtsey, and P. Fay, "High voltage, high current GaN-on-GaN p-n diodes with partially compensated edge termination," *Appl. Phys. Lett.* **113**, 023502 (2018).
- S. Pimpitkar, J. Speck, S. DenBaars, and S. Nakamura, "Prospects for LED lighting," *Nat. Photonics* **3**, 180–182 (2009).
- C.-H. Huang, K.-J. Chen, M.-T. Tsai, M.-H. Shih, C.-W. Sun, W.-I. Lee, C.-C. Lin, and H.-C. Kuo, "High-efficiency and low assembly-dependent chip-scale package for white light-emitting diodes," *J. Photon. Energy* **5**, 057606 (2015).
- T. Wu, C.-W. Sher, Y. Lin, C.-F. Lee, S. Liang, Y. Lu, S.-W. H. Chen, W. Guo, H.-C. Kuo, and Z. Chen, "Mini-LED and micro-LED: promising candidates for the next generation display technology," *Appl. Sci.* **8**, 1557 (2018).
- R. S. West, H. Konijn, W. Sillevs-Smith, S. Kuppens, N. Pfeffer, Y. Martynov, Y. Takaaki, S. Eberle, G. Harbers, T. W. Tan, and C. E. Chan, "43.4: high brightness direct LED backlight for LCD-TV," *SID Symp. Dig. Tech. Pap.* **34**, 1262–1265 (2003).
- Z. Luo, Y. Chen, and S.-T. Wu, "Wide color gamut LCD with a quantum dot backlight," *Opt. Express* **21**, 26269–26284 (2013).
- C. H. Huang, C. Y. Kang, S. H. Chang, C. H. Lin, C. Y. Lin, T. Wu, C. W. Sher, C.-C. Lin, P.-T. Lee, and H.-C. Kuo, "Ultra-high light extraction efficiency and ultra-thin mini-LED solution by freeform surface chip scale package array," *Crystals* **9**, 202 (2019).
- S. Kikuchi, Y. Shibata, T. Ishinabe, and H. Fujikake, "Thin mini-LED backlight using reflective mirror dots with high luminance uniformity for mobile LCDs," *Opt. Express* **29**, 26724–26735 (2021).
- F.-C. Lin, Y.-P. Huang, L.-Y. Liao, C.-Y. Liao, H.-P. D. Shieh, T.-M. Wang, and S.-C. Yeh, "Dynamic backlight gamma on high dynamic range LCD TVs," *J. Display Technol.* **4**, 139–146 (2008).
- G. Tan, Y. Huang, M.-C. Li, S.-L. Lee, and S.-T. Wu, "High dynamic range liquid crystal displays with a mini-LED backlight," *Opt. Express* **26**, 16572–16584 (2018).
- H. Chen, R. Zhu, M.-C. Li, S.-L. Lee, and S.-T. Wu, "Pixel-by-pixel local dimming for high-dynamic-range liquid crystal displays," *Opt. Express* **25**, 1973–1984 (2017).
- S.-W. H. Chen, Y.-M. Huang, K. J. Singh, Y.-C. Hsu, F.-J. Liou, J. Song, J. Choi, P.-T. Lee, C.-C. Lin, Z. Chen, J. Han, T. Wu, and H.-C. Kuo, "Full-color micro-LED display with high color stability using semipolar (20-21) InGaN LEDs and quantum-dot photoresist," *Photon. Res.* **8**, 630–636 (2020).
- E.-L. Hsiang, Z. He, Y. Hung, F. Gou, Y.-F. Lan, and S.-T. Wu, "Optimal chip size for reducing the power consumption of micro-LED displays," *Proc. SPIE* **11708**, 117080M (2021).
- C.-H. Huang, S. H. Chang, B.-Y. Liaw, C.-Y. Chou, C.-C. Lin, H.-C. Kuo, L.-J. Song, F. Li, and X. Liu, "Research on a novel GaN-based converted mini-LED backlight module via a spectrum-decouple system," *IEEE Access* **8**, 138823–138833 (2020).
- E. Chen, J. Guo, Z. Jiang, Q. Shen, Y. Ye, S. Xu, J. Sun, Q. Yan, and T. Guo, "Edge/direct-lit hybrid mini-LED backlight with U-grooved light guiding plates for local dimming," *Opt. Express* **29**, 12179–12194 (2021).
- E. Chen, H. Xie, J. Huang, H. Miu, G. Shao, Y. Li, T. Guo, S. Xu, and Y. Ye, "Flexible/curved backlight module with quantum-dots micro-structure array for liquid crystal displays," *Opt. Express* **26**, 3466–3482 (2018).
- H.-R. Moon, M.-H. Shin, J.-Y. Lee, K.-J. Jang, Y.-O. Chung, and Y.-J. Kim, "Design of integrated light guide plate with functional structure of enhanced diffusion length for ultra-slim LED backlight unit," *J. Display Technol.* **11**, 768–775 (2015).
- S. Wang, K. Wang, F. Chen, and S. Liu, "Design of primary optics for LED chip array in road lighting application," *Opt. Express* **19**, A716–A724 (2011).
- W.-S. Oh, D. Cho, K.-M. Cho, G.-W. Moon, B. Yang, and T. Jang, "A novel two-dimensional adaptive dimming technique of X-Y channel drivers for LED backlight system in LCD TVs," *J. Disp. Technol.* **5**, 20–26 (2009).
- J. H. Lee, S. G. Han, and M. J. Jin, "Minimum achievable height of a single-module LED low-beam projection headlamp," *Appl. Opt.* **60**, E8–E16 (2021).
- R. J. Lin, M.-S. Tsai, and C.-C. Sun, "Novel optical lens design with a light scattering freeform inner surface for LED down light illumination," *Opt. Express* **23**, 16715–16722 (2015).
- B.-Y. Joo and J.-H. Ko, "Analysis of color uniformity of white LED lens packages for direct-lit LCD backlight applications," *J. Opt. Soc. Korea* **17**, 506–512 (2013).
- S.-L. Hsiao, N.-C. Hu, and H. Cornelissen, "Phosphor-converted LED modeling using near-field chromatic luminance data," *Opt. Express* **21**, A250–A261 (2013).
- S.-S. Schad, B. Neubert, C. Eichler, M. Scherer, F. Habel, M. Seyboth, F. Scholz, D. Hofstetter, P. Unger, W. Schmid, C. Karnutsch, and K. Streubel, "Absorption and light scattering in InGaN-on-sapphire- and AlGaInP-based light-emitting diodes," *J. Lightwave Technol.* **22**, 2323–2332 (2004).
- N. Pendum and C. P. Vardhani, "Design and simulation of asymmetrical Y-branch optical power splitter on SOI platform and study the propagation loss with branching angle, index difference, and slab height," in *13th International Conference on Fiber Optics and Photonics* (Optical Society of America, 2016), paper W3A.44.
- A. D. Griffiths, J. Herrnsdorf, R. K. Henderson, M. J. Strain, and M. D. Dawson, "High-sensitivity inter-satellite optical communications using chip-scale LED and single-photon detector hardware," *Opt. Express* **29**, 10749–10768 (2021).
- X. Li, L. Wu, Z. Liu, B. Hussain, W. C. Chong, K. M. Lau, and C. P. Yue, "Design and characterization of active matrix LED microdisplays

- with embedded visible light communication transmitter,” *J. Lightwave Technol.* **34**, 3449–3457 (2016).
30. P. S. Rajendran and A. Geetha, “Optimization of hospital bed occupancy in hospitals using double deep Q network (DDQN),” in *3rd International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)* (2021), pp. 1029–1033.
  31. X. Huang, W. Luo, and J. Liu, “Attitude control of fixed-wing UAV based on DDQN,” in *Chinese Automation Congress (CAC)* (2019), pp. 4722–4726.
  32. M. G. P. de Lacerda, F. B. de Lima Neto, H. de Andrade Amorim Neto, H. Kuchen, and T. B. Ludermir, “On the learning properties of dueling DDQN in parameter control for evolutionary and swarm-based algorithms,” in *IEEE Latin American Conference on Computational Intelligence (LA-CCI)* (2019), pp. 1–6.
  33. B. Wang, Y. Li, W. Ming, and S. Wang, “Deep reinforcement learning method for demand response management of interruptible load,” *IEEE Trans. Smart Grid* **11**, 3146–3155 (2020).
  34. B. C. Sanders, P. Palittapongarpim, and S. S. Vedaie, “Reinforcement learning for adaptive optical quantum-enhanced metrology,” in *Imaging and Applied Optics 2018* (Optical Society of America, 2018), paper STu5H.4.
  35. X. Hu and A. Schülzgen, “Design of negative curvature hollow core fiber based on reinforcement learning,” *J. Lightwave Technol.* **38**, 1959–1965 (2020).
  36. S. P. K. Karri, D. Chakraborti, and J. Chatterjee, “Learning layer-specific edges for segmenting retinal layers with large deformations,” *Biomed. Opt. Express* **7**, 2888–2901 (2016).
  37. R. Zhu, S. Li, P. Wang, L. Li, A. Samuel, and Y. Zhao, “Deep reinforced energy efficient traffic grooming in fog-cloud elastic optical networks,” in *Optical Fiber Communication Conference (OFC)* (Optical Society of America, 2020), paper M1A.4.
  38. Q. Chen, M. Zhu, T. Shen, J. Gu, C. Yan, and P. Gu, “Deep-reinforcement-learning-assisted network orchestration for VNF-SC provisioning in inter-DC elastic optical networks,” in *Asia Communications and Photonics Conference/International Conference on Information Photonics and Optical Communications (ACP/IPOC)* (Optical Society of America, 2020), paper M4A.185.
  39. L. Moro and E. Prati, “Optical manipulation of qubits by deep reinforcement learning,” in *OSA Quantum 2.0 Conference* (Optical Society of America, 2020), paper QM6A.4.
  40. X. Wang, Y. Huang, J. Liu, and S. Yu, “A subcarrier-slot autonomous partition scheme based on deep-reinforcement-learning in elastic optical networks,” in *Asia Communications and Photonics Conference (ACPC)* (Optical Society of America, 2019), paper M4A.211.