

## 深度学习的亚波长窄带陷波滤光片设计

张帅帅<sup>1</sup>, 郭俊华<sup>1</sup>, 刘华东<sup>1</sup>, 张颖莉<sup>1</sup>, 肖相国<sup>2</sup>, 梁海锋<sup>1\*</sup>

1. 西安工业大学光电工程学院, 陕西 西安 710021
2. 西安应用光学研究所, 陕西 西安 710065

**摘要** 亚波长光栅结构表现出优异的陷波滤光特性, 其经典设计是设定亚波长的几何结构参数, 求解麦克斯韦方程组, 设定优化算法求解出最优解, 需要消耗大量的时间和计算资源。提出一种基于深度学习的逆向设计方法, 搭建了可以同时实现正向模拟与逆向设计的串联神经网络。基于 python 语言的 Tensorflow 库进行网络搭建; 优化均匀波导层的高度、亚波长光栅的高度、折射率、周期以及占空比; 研究亚波长光栅在  $0.45\sim 0.7\ \mu\text{m}$  的陷波滤光特性。采用严格耦合波分析(RCWA)数值模拟生成 23 100 组数据集, 在生成的数据集中随机选择 18 480 组数据作为训练集, 4 620 组作为测试集, 并对不同的网络层数, 网络节点以及 Batch\_size 进行了研究。结果显示经过 1 000 次的迭代后, 当网络的模型结构为  $5\times 50\times 200\times 500\times 200\times 26$ , Batch\_size 大小为 128 时, 网络性能最佳。相比独立的网络模型, 串联神经网络的正向模拟测试集损失率从 0.033 63 降到了 0.004 5, 逆向设计的测试集损失率从 0.702 98 降到了 0.052 98, 同时解决了由数据的非唯一性导致的网络逆向设计过程中无法收敛的问题。在完成训练的网络中输入任意的光谱曲线, 网络平均在 1.35 s 内给出亚波长光栅的几何结构参数; 并与 RCWA 数值模拟曲线进行相关性分析, 曲线相似系数均大于 0.658 1, 属于强相关。另外, 设计红、绿、蓝三种颜色的陷波滤光片, 其峰值反射率分别可以达到 98.91%, 99.98% 和 99.88%, 与传统方法相比, 该方法可以快速、精确的求解出光栅的几何参数, 为亚波长光栅设计提供了新方法。

**关键词** 神经网络; 亚波长结构; 深度学习; 陷波滤光片

**中图分类号:** O433.5 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2022)05-1393-07

### 引言

亚波长光栅的周期比入射光小的多, 只有零级衍射光波, 其他级次的光都为倏逝波。在光学、光电子学等许多领域被广泛应用, 成为现代仪器仪表中不可或缺的重要功能器件。亚波长光栅体积小、偏振性能好、允许入射角度大、易于集成等优点, 被广泛的应用在传感<sup>[1]</sup>, 滤光片光谱调控<sup>[2]</sup>和太阳能光伏吸收薄膜等领域<sup>[3]</sup>。通过调节光栅的几何结构, 可以对入射光谱进行调控。亚波长光栅的这种特性被深入研究, 并广泛应用于陷波滤光片。

传统的设计方法是, 通过给定光栅几何结构, 模拟出其光谱特性, 经过优化迭代给出光谱响应曲线, 称为正向模拟设计; 设计中主要使用严格耦合波分析法(RCWA), 时域有限差分法(FDTD), 有限元建模(FEM)等求解麦克斯韦方程

组来计算光谱曲线。但是根据需求的光谱曲线逆向设计并求解光栅的几何参数, 经典设计理论遇到瓶颈。相比于正向模拟设计, 真正的逆向模拟设计方法和理论鲜有报道。

在光栅设计优化方面, 传统数值优化方法, 如遗传算法, 最速下降法, Newton法和共轭梯度法等, 有相对较快的收敛速度, 计算精度高, 但求得的是局部最优解。对全局优化问题, 下降轨线法, 隧道法等收敛快, 计算效率高, 但算法复杂, 求得全局极值的概率不大<sup>[4]</sup>。Monte-Carlo 随机试验法, 模拟退火算法<sup>[5]</sup>等容易实现, 但收敛较慢, 效率低。上述算法都需要大量的计算, 消耗大量的时间才能找到一种合适的设计, 且每次针对新的光栅设计时所需的资源消耗是叠加的, 且只保留最优结果, 在优化过程中先前生成的数据都会被摒弃, 耗费大量的资源。原因是, 上述优化算法依然取决于求解麦克斯韦方程组, 根据设计目标, 引入评价函数和优化策略, 其本质仍然属于正向设计的一部分。

收稿日期: 2021-03-05, 修订日期: 2021-06-18

基金项目: 国防科技重点实验室基金项目(61424120506162412002), 陕西省重点研发计划项目(2020GY-045)资助

作者简介: 张帅帅, 1996年生, 西安工业大学硕士研究生 e-mail: 863711514@qq.com

\* 通讯作者 e-mail: lianghaifeng@xatu.edu.cn

最近,深度学习已经被引入到光学领域的诸多方面。在将光学器件的设计优化与深度学习相结合的探索中,国外的研究者率先展开了研究。2020年 Han<sup>[6]</sup>等基于卷积神经网络,在可见光范围内对不规则的二维亚波长光栅结构进行了研究,将随机生成的光栅形状与周期作为输入变量,分别设计了正向模拟网络与逆向设计网络,实现了快速高效的高自由度的逆设计,其中正向模拟网络完成训练之后,消耗时间仅为 RCWA 数值模拟的 1/20 000。2018年 Liu<sup>[7]</sup>在 0.4~1.2  $\mu\text{m}$  范围内,对于 16 层膜系结构进行了网络结构的设计,同时提出了一种串联的神经网络结构,首先训练一个正向设计网络,将设计映射到光学响应。然后,将该训练好的前向网络连接到逆向设计网络的输出,并将前向预测误差作为监视信号。通过在这种串联网中进行间接的训练,解决了数据不一致性导致的深度神经网络训练无法拟合的问题。John Peurifoy<sup>[8]</sup>等使用神经网络模拟计算多层纳米粒子的光散射,从模型中获得的分析梯度用于结构优化,给出了光谱中的特定要求;在深度学习的分析指导下,核壳纳米粒子的单带高散射效应可以得到有效优化。并分别对深度神经网络和内点法,在相同的误差阈值下运行时间进行了比较,结果表明神经网络的运行时间比内点法短两个数量级。Takashi Asano<sup>[9]</sup>等使用卷积神经网络分析光子晶体的 Q 因子,使用反向传播优化纳米腔的位置,使得 Q 因子从  $3.8 \times 10^{-3}$  提高至  $1.6 \times 10^{-9}$ ,由于共振的各种模式,光谱和光信号在时域中会有各种线形。而 RNN 适用于对时域中具有特定线形的光信号或光谱建模。与 CNN 结合使用时,RNN 也用于改善图像中表示的纳米结构的光学响应的近似值。Christopher 等<sup>[10]</sup>2020年在 4~12  $\mu\text{m}$  的中红外波段内,将二维光栅的结构设置为十字结构,通过调整结构的大小与位置得到不同的吸收光谱,以此为训练数据设计了串联神经网络。Jiang<sup>[11]</sup>等在 2019 年探索了将生成对抗网络与伴随优化算法框架耦合在一起的方法。已经证明可以在标记拓扑优化的亚光栅设计上有效训练条件 GAN,以快速生成大量的高效器件设计,与传统的拓扑优化相比,GAN 辅助优化的效率高出 5 倍。

上述工作证实了深度学习在设计特定光谱应用中潜力巨

大,由于受深度学习方法的限制,对某一类应用,需要特定学习集,经过训练后实现特定的功能。陷波滤光片是一类特殊的光学应用,受最近工作的启发,提出了一种基于深度学习的一维周期性的亚波长光栅设计,如图 1 所示。所设计的亚波长光栅滤光片由亚波长光栅层,以及玻璃基板上的均匀波导层组成。其中  $H_1$  表示光栅深度, $H_2$  表示均匀波导层的厚度, $F$  为光栅占空比, $\Lambda$  为光栅周期, $N$  为光栅以及波导层折射率。

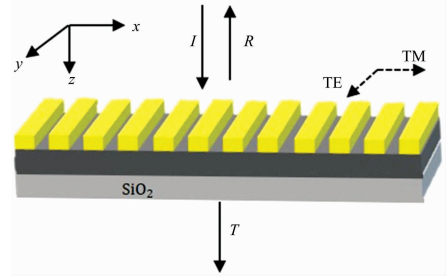


图 1 一维亚波长光栅结构图

Fig. 1 One via wavelength grating structure diagram

为了克服传统亚波长光栅设计优化所面临的核心问题,分别设计了正向神经网络(不需要计算 Maxwell 方程组),逆向设计网络以及串联神经网络,以实现高精度,高效率,响应时间快的特性。其中正向神经网络模拟光谱响应,逆向神经网络设计光栅参数。设计完成的神经网络通过学习 RCWA 数值模拟的仿真数据,不需要再次求解麦克斯韦方程组便能预测出亚波长光栅的光谱响应,并且能精确的逆向设计光栅结构,无需迭代优化,可将优化设计时间缩短多个数量级。

## 1 基于深度学习的串联神经网络构建

### 1.1 仿真学习网络建立

为了设计亚波长光栅结构的窄带陷波滤光片,设计了图 2 所示神经网络。正向模拟的网络结构如图 2(a),包括输入层,多个隐藏层以及输出层。

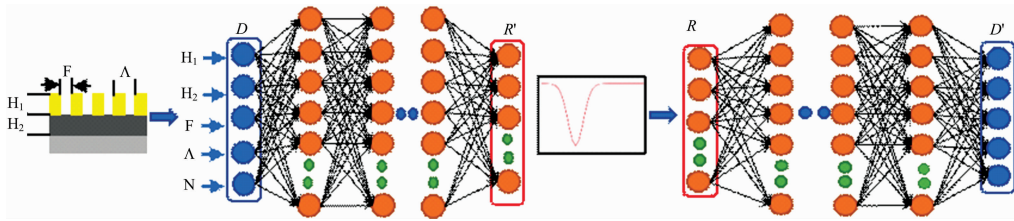


图 2 神经网络结构

(a): 正向模拟网络; (b): 逆向设计网络

Fig. 2 Neural network structure

(a): Forward simulation network; (b): Reverse-design network

输入层 5 个节点分别为,光栅高度  $H_1$  (0.01~0.09  $\mu\text{m}$ ),波导层厚度  $H_2$  (0.09~0.19  $\mu\text{m}$ ),光栅周期  $\Lambda$  (0.2~0.4  $\mu\text{m}$ ),光栅折射率  $N$  (1.5~2.4),占空比  $F$  (0.4~0.75)。对不同个数的隐藏层以及隐藏层中包含的不同个数的神经节点进行测试以及比对。层数为 5 层时,网络无法收敛,均方

误差较大;层数为 3 层时,虽然消耗时间较少,但均方误差为 4 层时的十倍。最终设置网络的隐藏层个数为 4,4 个隐藏层的神经元个数分别为 (50, 200, 500, 200),详细评价指标列于表 1,表 2。输出层 26 个节点,对应可见光范围 0.45~0.7  $\mu\text{m}$  内的 26 个点,步长为 0.01  $\mu\text{m}$ ,作为光谱响应特征

曲线。Batch\_size 为网络每一次训练所选取的样本数，影响着模型的优化程度和速度。Batch\_size 设置过小训练精度高，但消耗时间越长需要的迭代次数越高，可能造成内存爆炸。Batch\_size 过大消耗时间少，但是过大可能会导致迭代无法收敛。因此需要选取合适的 Batch\_size 使得网络达到最优。合适的 Batch\_size 选取可以使得梯度下降更加准确，达到良好的训练效果，通过对不同大小的测试，最终选取 Batch\_size 大小为 128，详细评价列于表 3。

表 1 不同隐藏层层数的评价指标  
Table 1 Evaluation indexes of network with different hidden layers

	网络层数	消耗时间/s	均方误差
Model_A	5	无法收敛	0.513 905
Model_B	4	465	0.014
Model_C	3	382	0.10

表 2 不同网络结构的评价指标  
Table 2 Evaluation indexes of network with different network structures

	网络结构	消耗时间/s	均方误差
Model_1	50, 200, 200, 50	91.64	0.068 427
Model_2	50, 200, 200, 200	121.89	0.082 540
Model_3	50, 200, 500, 200	190.29	0.031 505
Model_4	50, 200, 500, 500	373.53	0.033 413

表 3 不同 Batch\_size 的评价指标  
Table 3 Evaluation indexes of network with different Batch sizes

	样本数	消耗时间/s	均方误差
Batch_size_1	32	3 483	0.025 68
Batch_size_2	64	574	0.034 15
Batch_size_3	128	381	0.028 23
Batch_size_4	256	282	0.031 52
Batch_size_5	512	268	0.040 17

正向模拟的作用是根据输入的光栅几何结构参数来迅速，精确的得到光谱响应曲线。而逆向设计的作用是根据期望光谱响应快速设计出光栅的几何结构。逆向设计的网络结构如图 2(b)，网络结构与正向模拟的网络结构相反，输入层为期望光谱曲线量化的 26 个点，输出为光栅的 5 个结构参数。

## 1.2 训练数据生成以及网络优化

为了使网络能够良好的运行并达到预期的效果，基于 RCWA 模拟仿真生成数据集，数据集的建立包括光栅的几何结构参数，其对应的光谱响应曲线，以及数据标签。光栅的几何结构如图 1，以 Si<sub>2</sub>O 为光栅基底，入射光为 TE 偏振光，垂直入射光栅表面。通过改变光栅层以及波导层的折射率、高度、周期、占空比，得到不同的数据。取可见光波段 0.45~0.7 μm 的波长范围，每 0.01 μm 取一个点，总共取 26 个点来量化光栅的反射光谱响应  $T(t_1, t_2, t_3, \dots, t_{26})$ 。

将训练好的数据集 80% 作为训练集，20% 作为测试集，设置网络的损失函数为

$$\text{Loss} = \text{MSE}(\hat{T}, T) = \frac{1}{2} \sum_{i=1}^N (\hat{t}_i - t_i)^2 \quad (1)$$

式(1)中， $\hat{t}_i$  为神经网络预测的光谱曲线， $t_i$  为 RCWA 数值模拟光谱曲线， $N$  为输出的节点数，此处为 26。损失函数是预测光谱  $\hat{t}_i$  与 RCWA 数值模拟光谱  $t_i$  的均方误差。对于逆向设计来说，损失函数是预测结构参数与真实结构参数之间的均方误差。神经网络每一层的输出为：

$$\hat{t}_i = \sigma(Z^l) \quad (2)$$

$$Z^l = W^l a^{l-1} + b^l \quad (3)$$

$$a^{l-1} = \sigma(Z^{l-1}) \quad (4)$$

$$\text{Sigmoid} \sigma = 1/(1 + e^{-z}) \quad (5)$$

$$\text{ReLU} \sigma = \max(0, Z) \quad (6)$$

其中  $\sigma$  为激活函数，激活函数的作用是加入非线性因素，提高神经网络对模型的表达能力，解决线性模型不能解决的问题。隐藏层的激活函数设置为 ReLU 函数，具有收敛速度快的优点。输出层设置为 Sigmoid 激活函数，适用于线性回归层。 $Z$  为网络的神经节点；输入层的神经节点为 5，隐藏层的输出节点分别为 (50, 200, 500, 200)。 $W$  为网络的权值， $b$  为网络的偏置， $l$  为网络层数，包含输入层，输出层以及四个隐藏层，输入层为第 0 层。

训练过程中使用 Adam 梯度下降算法，进行参数优化

$$\nu_{dW}[L] = \beta_1 \nu_d W[L] + (1 - \beta_1) \frac{\partial J}{\partial W^l[L]} \quad (7)$$

$$\nu_{dW}^{\text{corrected}}[L] = \frac{\nu_{dW}[L]}{1 - (\beta_1)^l} \quad (8)$$

$$S_{dW}[L] = \beta_2 S_{dW}[L] + (1 - \beta_2) \left( \frac{\partial J}{\partial W^l[L]} \right)^2 \quad (9)$$

$$S_{dW}^{\text{corrected}} = \frac{S_{dW}[L]}{1 - (\beta_2)^l} \quad (10)$$

$$W^{[L]} = W^{[L]} - \alpha \frac{\nu_{dW}^{\text{corrected}}}{\sqrt{S_{dW}^{\text{corrected}}}} + \epsilon \quad (11)$$

同理

$$b^{[L]} = b^{[L]} - \alpha \frac{\nu_{db}^{\text{corrected}}}{\sqrt{S_{db}^{\text{corrected}}}} + \epsilon \quad (12)$$

其中  $l$  为当前迭代次数， $\beta_1$  和  $\beta_2$  为控制两个指数加权平均值的超参数，一般设置为 0.9, 0.999。 $\alpha$  为学习率，其初始值设置为 0.002。 $\epsilon$  用于避免分母为零，一般为  $1 \times 10^{-8}$ 。通过计算之前梯度的指数加权平均值，和之前梯度平方的指数加权平均值。分别将其存储在变量  $\nu$  (偏差校正前)， $\nu^{\text{corrected}}$  (偏差校正后) 和  $s$  (偏差校正前)， $s^{\text{corrected}}$  (偏差校正后) 之中。通过 Adam 梯度下降算法不断的更新  $W$ ,  $b$ , 得到最佳值。

## 1.3 网络训练和测试

为了使网络运行达到预期的目标，建立了 23 100 组数据。在训练过程中发现，正向模拟训练效果良好，收敛速度快，如图 3(a) 所示。在逆向设计的时候，由于数据集中数据量庞大，有相似的光谱响应对应不同的光栅结构参数，造成单输入对应多输出的问题，使得网络无法找到最优解，迭代无法收敛，如图 3(b) 所示。例如两个不同的光栅结构  $M_n$  和

$M_b$ , 有相似的光谱响应曲线, 当神经网络在训练的过程中有两个数据 ( $M_a, T$ ) 与 ( $M_b, T$ ) 时训练将难以收敛, 因为这两个数据为网络提供了两个不同的答案, 使得网络无法达到训练的目的。神经网络的调控是通过最小化 Loss 函数来完成训练目的的, 在逆向设计中 Loss 函数为

$$\text{Inverse Loss} = \text{MSE}(M, \hat{M}) = \frac{1}{2} \sum_{i=1}^N (M_i - \hat{M}_i)^2 \quad (13)$$

式(13)中,  $M$  为真实数据,  $\hat{M}$  为在给定输入光谱  $T$  的情况下神经网络设计的结构参数。  $N$  为输出节点, 此处为 5。由于数据的非唯一性, 造成迭代无法收敛。

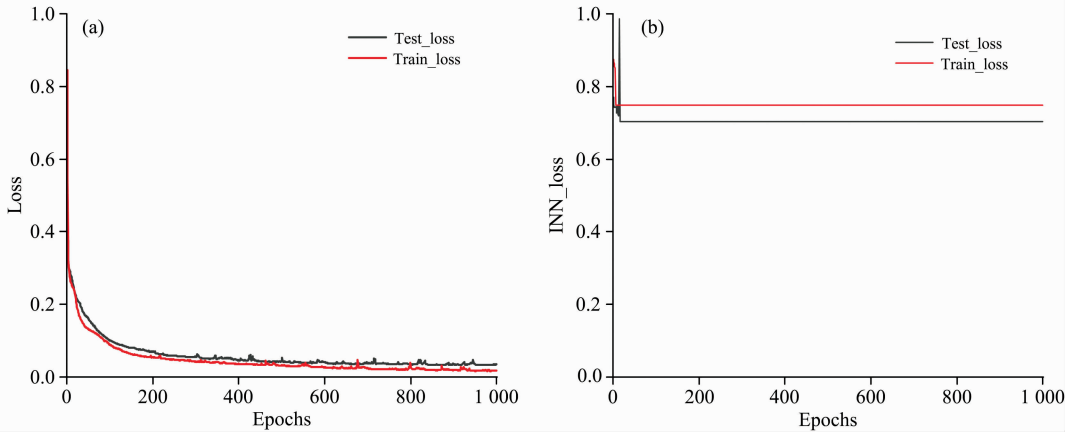


图 3 (a) 正向模拟损失函数曲线; (b) 逆向设计损失函数曲线

Fig. 3 (a) Forward simulation Loss function curve; (b) Inverse design Loss function curve

为了克服数据的非唯一性导致的迭代无法收敛的问题, 使用一种串联的神经网络(TNN)<sup>[9]</sup>如图 4, 将先前建立的正向模拟神经网络(红框部分)与逆向设计网络连接在一起, 中间层为光栅的结构参数  $D$ (蓝色部分), 整个串联网络的输入为目标光谱  $R$ , 串联网络的输出是根据逆向设计得到的光栅

结构参数  $D'$ , 输入已经训练完成的正向模拟网络所得到的预测光谱响应  $R'$ 。在串联网运行之前, 先将训练好的正向网络的权重  $W_f$  与偏置  $b_f$  设定。然后通过 Loss 函数调整逆向设计网络的权重  $W_i$  与偏置  $b_i$ 。与先前两种网络不同的是, 串联网络的损失函数计算的是输入光谱  $R$  与输出光谱  $R'$  之间的均方误差

$$\text{Loss} = \text{MSE}(R, R') = \frac{1}{2} (R - R')^2 \quad (14)$$

该 Loss 函数定义为预测光谱响应和目标光谱响应之间的均方误差。这种网络结构克服了逆向设计网络的非唯一性问题, 因为网络的设计结果不需要与训练样本中的实际设计相同。相反, 只要生成的几何结构参数和实际设计的几何结构参数具有相似的响应, Loss 函数就会很低。

为了测试串联网的逆向设计能力, 使用 RCWA 数值模拟任意三个光栅的光谱响应曲线, 并将光谱曲线输入串联网, 得到光栅结构参数, 将串联网给出的结构参数与 RCWA 数值模拟的结构参数列于表 4, 并作出串联网的 Loss 函数图, 如图 5 所示。

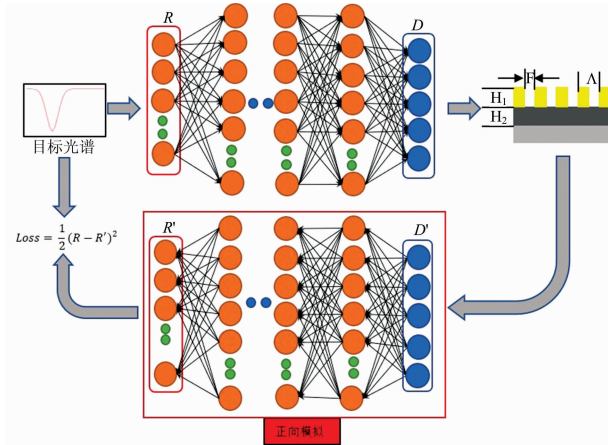


图 4 串联的神经网络

$R$ : 期望光谱;  $R'$ : 正向模拟预测光谱;  $D$ : 原始训练集的样本结构;  $D'$ : 逆向设计预测结构; 红色框线为正向模拟网络, 通过修改损失函数来解决数据的不唯一性导致的网络无法拟合问题, 中间层为逆向神经网络的输出, 正向模拟的输入

Fig. 4 Series neural network

$R$ : Expected spectral response;  $R'$ : Forward simulation prediction spectrum;  $D$ : Sample structure of the original training set;  $D'$ : Reverse design forecast structure. The red frame is the forward simulation network. The Loss function is modified to solve the problem that the network cannot be fitted due to the non-uniqueness of the data. The middle layer is the output of reverse design and the input of forward simulation

表 4 结构参数对比

Table 4 Comparison of structural parameters

Method	$H_1$ /nm	$H_2$ / nm	$F$	$\Lambda$ /nm	$N$	
1	RCWA	90	150	0.5	360	2.2
	Network	88.441 5	150.476	0.485 823	362.866 6	2.193 37
2	RCWA	70	110	0.7	360	2.1
	Network	67.619 93	120.378 0	0.676 11	360.900 5	2.069 56
3	RCWA	90	90	0.65	360	2.3
	Network	86.095 81	104.919 36	0.620 73	350.569	2.302 507

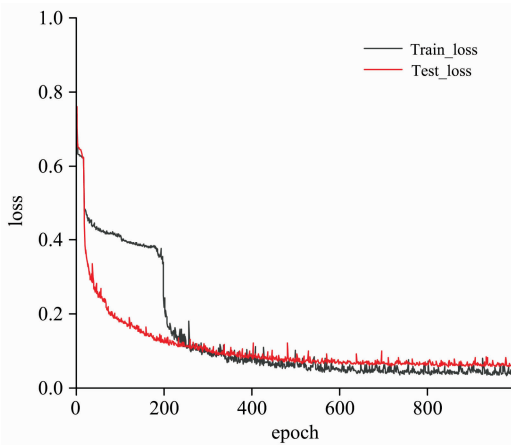


图 5 串联网络 Loss 函数曲线

Fig. 5 Series network loss function curve

## 2 实验验证

### 2.1 正向模拟验证

为了验证神经网络正向模拟的准确性以及可行性，查阅参考文献[12]，将文献中的设计参数以及光谱响应曲线与正

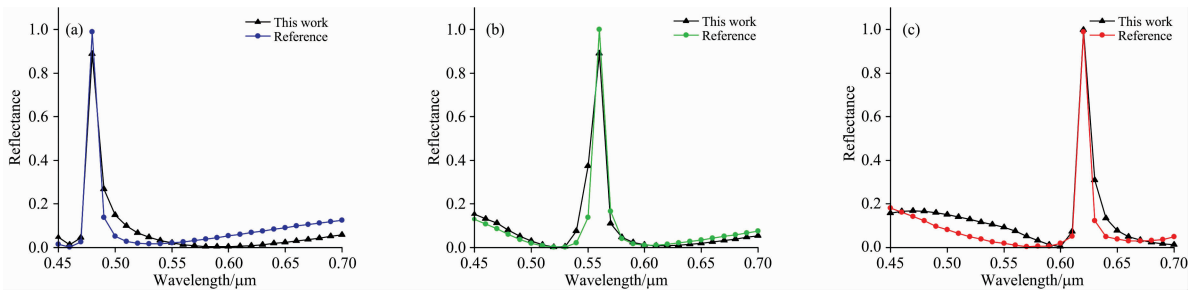


图 6 红、绿、蓝为参考文献的光谱响应曲线，黑色曲线为相同的结构参数 TNN 所生成的光谱响应曲线

Fig. 6 Red, green and blue are the spectral response curves reported by references, and black curves are the spectral response curves of TNN with the same structural parameter

表 6 相关系数评价指标<sup>[13]</sup>

Table 6 Evaluation Index of correlation coefficient

相关性	无相关	弱相关	中度相关	强相关
$r$	$r < 0.1$	$0.1 < r < 0.3$	$0.3 < r < 0.5$	$0.5 < r < 1$

### 2.2 逆向设计验证

验证串联网络的逆向设计的可行性，将未经训练的光谱曲线数据输入串联神经网络，输入的目标响应光谱曲线为 0.45~0.7 μm 范围内的高斯曲线

$$f = 0.98 \exp \left[ -\frac{(x - x_0)^2}{2\sigma^2} \right] + 0.02 \quad (16)$$

分别设置  $x_0$  为 0.48, 0.55, 0.61,  $\sigma$  设置为 0.005。将三条中心波长分别在 0.48, 0.55 与 0.61 μm, 且峰值反射率达到 100% 的红, 绿, 蓝光谱响应曲线输入串联网络, 并将逆向设计输出的几何结构参数(表 7)使用 RCWA 数值模拟, 得出

向模拟网络进行对比。文献中以 SiO<sub>2</sub> 为基层, 光栅以及波导层选用 Si<sub>3</sub>N<sub>4</sub>。在 0.45~0.7 μm 分别对红、绿、蓝三色陷波滤光片进行验证, 三种陷波滤光片的结构参数列于表 5。将表 5 中光栅的结构参数输入正向模拟网络, 所得到的光谱响应曲线与文献中的曲线比较如图 6 所示。

表 5 红绿蓝结构参数<sup>[12]</sup>

Table 5 Red, green and blue structural parameters

	$H_1$ /nm	$H_2$ /nm	$F$	$\Lambda$ /nm	$N$
Red (607 nm)	60.5	104.5	0.46	363.6	2.04(Si <sub>3</sub> N <sub>4</sub> )
Green(551 nm)	59.2	105.8	0.46	323.5	2.04(Si <sub>3</sub> N <sub>4</sub> )
Blue(479.5 nm)	58.5	106.5	0.46	273.15	2.04(Si <sub>3</sub> N <sub>4</sub> )

分别将文献中的数据与正向模拟所设计的光谱响应进行相关性计算, 计算两者之间的欧氏距离

$$d = \left( \sum_{k=1}^n (x_{1k} - x_{2k})^2 \right)^{1/2} \quad (15)$$

相关系数  $r=1/(1+d)$ , 计算红、绿、蓝三色的相关系数分别为 0.763 82, 0.835 87, 0.890 88, 将 3 个相关系数与表 6 中的评价指标进行比对, 验证其相关性为强相关。

设计结构参数对应的真实光谱响应。将目标响应光谱与真实光谱的结果相比较, 如图 7。

表 7 逆向设计结构参数

Table 7 Reverse design parameters

	$H_1$ /nm	$H_2$ /nm	$F$	$\Lambda$ /nm	$N$	$r$
Red (607 nm)	79.011 7	120.795	0.584	381.602	1.926	0.685 1
Green(551 nm)	53.246 52	107.609	0.411	326.287	2.052	0.813 4
Blue(479.5 nm)	45.680 69	93.037	0.602	283.814	2.014	0.789 6

在训练完成的神经网络中进行各个目标光谱的设计, 仅需 1.35 s。极大的提高了陷波滤光片的设计效率, 且峰值反射率都能达到 99% 以上。

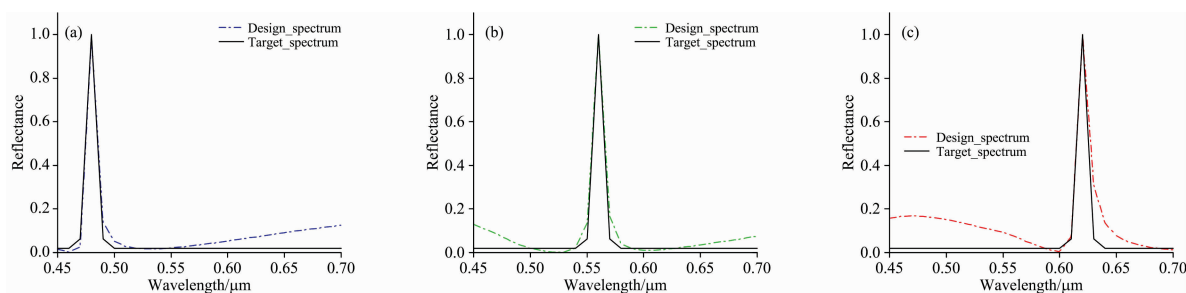


图 7 串联网路逆向设计

黑色曲线为目标光谱, 反射率为 100%, 红绿蓝三条曲线为逆向设计参数的 RCWA 数值模拟曲线, 峰值波长分别为 479.5, 551 和 607 nm, 反射率分别为 98.91%, 99.98% 和 99.88%

Fig. 7 RCWA numerical simulation curves with inverse design of series network

Black curve is target spectrum with a reflectivity of 100%; red-green-blue curves are RCWA simulation curves of reverse design with the reflectivity of 98.91%, 99.98% and 99.88% at the peak wavelenghtes of 479.5, 551.0 and 607.0 nm, respectively

### 3 结 论

基于深度学习搭建了可以同时实现正向模拟与逆向设计的串联神经网络, 并基于 RCWA 数值模拟生成 23100 组数据。将两者相结合应用在亚波长光栅陷波滤光片的结构设计, 可以快速精确的得到最优光栅结构参数, 输入特定的响

应光谱曲线可以在 1.35 s 内给出光栅的几何结构参数。通过计算验证, 所设计曲线与目标曲线的相关系数大于 0.5, 属于强相关。为亚波长光栅的设计提供了新方法和思路, 相比于传统的优化设计方法极大的提高了亚波长光栅的设计效率。同时以光栅滤光片为例提出了一种深度学习法, 不仅仅适用于亚波长光栅的结构设计, 也可用于光学的其他领域, 例如膜系设计, 图像处理, 光学系统设计等。

### References

- [ 1 ] Cunningham B T, Qiu J, Lin B, et al. Proc IEEE Sensors, 2002, 1: 212.
- [ 2 ] Wang J, Fan Q, Zhang S, et al. Applied Physics Letters, 2017, 110(3): 31110.
- [ 3 ] Cui Yanxia, Fung Kung Hin, Xu Jun, et al. Nano Letters, 2012, 12(3): 1443.
- [ 4 ] ZHAO Ming-wang(赵明旺). Control and Decision(控制与决策), 1997, 12(5): 589.
- [ 5 ] Kirkpatrick S, Gelatt C D, Vecchi M P. Science, 1983, 220(4598): 671.
- [ 6 ] Xiao Han, Fan Ziyang, Liu Zeyang. InfoMat. , 2020, 3.10.1002/inf2.12116.
- [ 7 ] Liu D, Tan Y, Khoram E, et al. ACS Photonics, 2018, 5(4): 1365.
- [ 8 ] Peurifoy J, Shen Y, Jing L, et al. Sci. Adv. , 2018, 4: 4206.
- [ 9 ] Asano T, Noda S. Opt. Express, 2018, 26: 32704.
- [ 10 ] Yeung Christopher, Tsai Ju-Ming, King Brian, et al. Designing Multiplexed Supercell Metasurfaces with Tandem Neural Networks, 2020. arXiv: Optics(2020): n. pag.
- [ 11 ] Jiang J, Fan J A. Nano Lett. 2019, 19(8): 5366.
- [ 12 ] Uddin M J, Magnusson R. Opt. Express, 2013, 21(10): 12495.
- [ 13 ] Cohen J. Statistical Power Analysis for the Behavioral Sciences . 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.

# Design of Subwavelength Narrow Band Notch Filter Based on Depth Learning

ZHANG Shuai-shuai<sup>1</sup>, GUO Jun-hua<sup>1</sup>, LIU Hua-dong<sup>1</sup>, ZHANG Ying-li<sup>1</sup>, XIAO Xiang-guo<sup>2</sup>, LIANG Hai-feng<sup>1\*</sup>

1. School of Optoelectronics Engineering, Xi'an Technological University, Xi'an 710021, China

2. Xi'an Institute of Applied Optics, Xi'an 710065, China

**Abstract** Subwavelength grating structures exhibit excellent notch filtering properties. The classical design is to find the optimal solution by setting the geometric structure parameters of the subwavelength, solving Maxwell's equations, and setting an optimization algorithm. It consumes a lot of time and computing resources. This paper presents an inverse design method based on deep learning and constructs a series neural network which can realize both forward simulation and inverse design. The Tensorflow library based on Python language is constructed to optimize the height of uniform waveguide layer, the height of sub-wavelength grating, refractive index, period and duty cycle, and to study the characteristics of sub-wavelength grating notch filtering in the range of  $0.45\sim 0.7\ \mu\text{m}$ . Using rigorous coupled wave analysis (RCWA) numerical simulation to generate 23 100 data sets, 18 480 data sets were randomly selected as training sets, and 4 620 data sets were used as test sets, the network node and Batch were studied. The results show that the network performance is best when the network model structure is  $5\times 50\times 200\times 500\times 200\times 26$ , and the Batch size is 128 after 1 000 iterations. Compared with the independent network model, the loss rate of the forward simulation test set of the series neural network decreased from 0.033 63 to 0.004 5, and that of the reverse design decreased from 0.702 98 to 0.052 98. At the same time, the problem that the network can not converge in the reverse design process caused by the non-uniqueness of data is solved. The geometric structure parameters of the sub-wavelength grating are given in 1.35 s on average by inputting any spectral curve into the trained network, and the correlation between the parameters and the RCWA numerical simulation curve is analyzed, the similarity coefficients of the curves were all greater than 0.658 1, which belonged to strong correlation. In addition, a red, green and blue notch filter is designed, whose peak reflectivity can reach 98.91%, 99.98% and 99.88% respectively. Compared with the traditional method, this method can quickly and accurately calculate the geometric parameters of the grating. It provides a new method for sub-wavelength grating design.

**Keywords** Neural network; Sub-wavelength structure; Deep learning; Notch filter

(Received Mar. 5, 2021; accepted Jun. 18, 2021)

\* Corresponding author