基于神经网络的高角色散宽带介质超光栅的快速优化

李润泽[1]，董希谱[1]，程洁嵘[1,2**]，常胜江[1,3*]

[1]南开大学现代光学研究所，天津 300350；

[2]天津市微尺度光学信息技术科学重点实验室，天津 300350；

[3]天津市光电传感器与传感网络技术重点实验室，天津 300350

摘要 由于超材料和超表面的亚波长结构单元的形状和尺寸具有很大的设计自由度，可对电磁波的振幅、相位、波前和方向等进行复杂而精确的调控，同时随着结构参数数量的增加，结构设计的时间往往呈指数增长。提出了一种基于反向传播(BP)神经网络快速优化超表面结构的方法，实现了兼具高衍射效率、宽带宽和高角色散等优势的太赫兹介质超光栅。利用有限次数的严格耦合波分析建立的数据集来训练 BP 神经网络，可准确预测任意结构参数的超光栅衍射光谱，并通过遍历所有结构参数快速筛选出具有最高衍射效率且宽带宽的超光栅，相比传统的遍历计算方法速度提高了一万倍，证明了基于 BP 神经网络的超表面优化方法的高效性以及精准性，同时为太赫兹波段提供了一种性能优异的衍射元件。

关键词 光栅；深度学习；BP 神经网络；超光栅；优化

中图分类号 O436 **文献标志码** A **doi**：10.3788/AOS202040.2305001

Fast Optimization of High-Angular-Dispersion Wideband Dielectric Metagratings Based on Neural Networks

Li Runze[1], Dong Xipu[1], Cheng Jierong[1,2**], Chang Shengjiang[1,3*]

[1] *Institute of Modern Optics, Nankai University, Tianjin* 300350, *China*;

[2] *Tianjin Key Laboratory of Micro-Scale Optical Information Science and Technology, Tianjin* 300350, *China*;

[3] *Tianjin Key Laboratory of Optoelectronic Sensor and Sensing Network Technology, Tianjin* 300350, *China*

Abstract Metamaterials and metasurfaces show great potentials to adjust the amplitude, phase, wavefront and direction of electromagnetic waves in a complex and precise manner, since the shape and size of the subwavelength unit can be designed with large degree of freedom. At the same time, with the increase of the number of structural parameters involved, the structural design time increases in an exponential way. This paper proposes a method for the fast optimization of metasurface structures based on the back-propagation (BP) neural network, and a terahertz dielectric metagrating with the merits of high diffraction efficiency, wide bandwidth, and high angular dispersion is achieved. A dataset established via a limited number of rigorous coupled wave analyses is used to train the BP neural network. It can accurately predict the diffraction spectrum of the metagrating with an arbitrary geometry. Simultaneously, the metagrating with the highest diffraction efficiency and wide bandwidth is fast selected by quickly traversing all structural parameters. The designed speed is increased by 10,000 times compared with that of the traditional traversing calculation method, which proves the high efficiency and accuracy of the metasurface optimization method based on the BP neural network. The study provides a diffractive element with excellent performance for terahertz applications.

Key words gratings; deep learning; BP neural network; metagrating; optimization

OCIS codes 050.1950; 160.1245; 160.3918; 160.4670

收稿日期：2020-08-18；**修回日期**：2020-09-03；**录用日期**：2020-09-08

基金项目：国家重点基础研究发展计划(2017YFA0701000)、国家自然科学基金(61831012，61805123)、天津市自然科学基金(18JCQNJC02200)

* **E-mail**：sjchang@nankai.edu.cn；** **E-mail**：chengjr@nankai.edu.cn

1 引　言

随着计算机计算力的发展，神经网络的计算速度有了显著的提升，人工智能再次迎来研究热潮。深度学习反向传播（BP）神经网络具有强大的数据处理能力，可通过学习数据之间的逻辑关系来完成复杂的任务，极大推动了计算机视觉、语音识别和图像处理等领域的发展[1-5]。同时，神经网络也逐步应用于光学元件的优化设计[6]。

超表面是一类紧凑型新颖光学元件，可通过对电磁波振幅、相位和偏振的精确控制来实现波前操控、偏振变换和全息成像等功能[7-8]。超表面的亚波长结构单元有很多结构参数，这表明结构的自由度很高，可为波束操控提供很大的空间，同时也增加了结构设计的复杂度。将神经网络用于超表面结构的优化设计可以提高优化速度，为多变量多目标的复杂问题寻找全局最优解提供一种高效的方法[9]。传统数值优化方法如梯度下降等容易陷入局域最优解，常常需要采用不同的初始解进行多次优化来寻找全局最优解[10]，而另一些数值优化方法如遗传算法和粒子群优化算法[11]等虽然不依赖于初始解，但其收敛速度慢，迭代次数多。在优化过程中，每一步迭代都要采用时域有限差分法（FDTD）、有限元法（FEM）和严格耦合波分析法（RCWA）等方法进行电磁仿真计算，当超表面面积较大且不具备周期性时，每一次电磁仿真都需要一定的时间和计算资源，这大大限制了优化的速度。而神经网络只需要一定数量的电磁仿真数据，无需求解麦克斯韦方程组即可预测出器件的光学性能或反向预测器件的结构，从而将优化设计的时间缩短多个数量级。

神经网络与超表面元器件设计的结合已取得了一定进展[12]。神经网络对超表面的优化可分为正向预测和反向设计[13]。正向预测是输入结构的几何参数，输出预测的效率和电磁场分布等光学响应。Qian 等[14]利用神经网络提出一种隐形斗篷，可对直流偏置电压进行正向预测，进而快速地调节超表面内部每个元素的反射特性，实现了快速高效的自适应隐身。神经网络也被用于预测不同纳米结构的光谱[15]以及纳米波导的有效折射率[16]。

反向设计分为反向参数设计以及反向拓扑优化。反向参数设计是针对一个大致确定的结构，输入性能参数，输出结构的几何参数；反向拓扑优化是根据性能的需要，对结构进行拓扑生成。常见的反向设计包括利用对抗神经网络反向预测折射率分布[17]，也可以通过训练深层神经网络进行纳米光子结构逆设计[18-20]，利用卷积逆向生成等离子体超表面[21]，或按照所需的圆二向色性谱，通过生成式对抗网络（GAN）对抗神经网络对手性器件进行拓扑优化设计，逆向生成器件拓扑结构[22-24]。针对一个含有有限可变参数的超表面结构，正向设计就可以对光学性能进行较好的预测。而当超表面需要实现多目标的复杂功能时，基于神经网络的反向拓扑优化可提高设计自由度，增强模型的多样性。

本文构建了一种深度学习 BP 神经网络，采用正向预测的方法实现了太赫兹波段宽带宽、高效率和高角色散的衍射光栅的快速优化设计，并利用全连接 BP 神经网络，对衍射物理过程进行了泛化学习，以神经网络表征复杂的数值计算过程。该神经网络可准确地预测各种光栅结构参数下的衍射光谱，并最终得到了衍射效率大于 80%、带宽为 19.2%的太赫兹超光栅，其在 161 GHz 处的峰值衍射效率为 95.7%，角色散为 43.61(°)/mm。相比于基于电磁仿真的遍历搜索方法，该方法的优化速度提高了一万倍。

2 基于神经网络的超光栅衍射光谱预测

2.1 物理问题及数据集的制作

衍射光栅可以控制波束的传播方向，其衍射方向决定于光栅方程：

$$\sin(\theta_{k'}) - \sin(\theta_{\mathrm{i}}) = k'\frac{\lambda}{\Lambda}, \tag{1}$$

式中：λ 为波长；Λ 为光栅周期；k'为衍射级次；θ_{i} 为入射角度；$\theta_{k'}$为第 k'级衍射角度。光栅的角色散率表征衍射角随波长的变化率，可表示为

$$D = \frac{\mathrm{d}\theta}{\mathrm{d}\lambda} = \frac{k'}{\cos(\theta_{k'})\Lambda} = \frac{k'}{\sqrt{\Lambda^2 - \lambda^2}}。 \tag{2}$$

由于其角色散特性，衍射光栅被广泛应用于光谱分析和脉冲整形等。对于刻线密度为 300 lp/mm、闪耀角为 17.5°的可见光透射式光栅（THORLABS GT13-03），其衍射效率峰值约为 70%，带宽（效率大于 70%）约为 22.2%，角色散约为 0.0174 (°)/nm。为了得到更大的角色散，往往需要减小光栅周期，牺牲衍射效率和带宽。关于太赫兹波段衍射光栅的研究较少，发展太赫兹衍射光栅对于太赫兹波束方向的控制以及太赫兹光谱分析具有重要意义。因此本文以实现太赫兹波段的高效率、宽频带、高角色散的衍射光栅为目标，利用全连接 BP 神经网络实现该器件结构的快速优化。

衍射光栅的结构及功能示意图如图 1(a)所示，该光栅可将垂直入射的太赫兹波高效率地偏折至+1 衍射级。该光栅由上下两层光栅及中间层构成，光栅的构成材料为聚乳酸(PLA)，其在 0.1～0.3 THz 处的折射率为 1.57[25]。光栅沿 y 方向无限延伸，沿 z 方向的周期固定为 2.28 mm，当入射频率为 161 GHz 时，衍射角度为 54.81°，角色散为 43.61 (°)/mm。可变结构参数分别为图 1(b)所示的上下层光栅宽度 W_1、W_2，以及光栅高度 H_1、H_2 和中间层高度 H_{sub}。入射光沿 x 轴负方向传播，电场方向与光栅垂直。

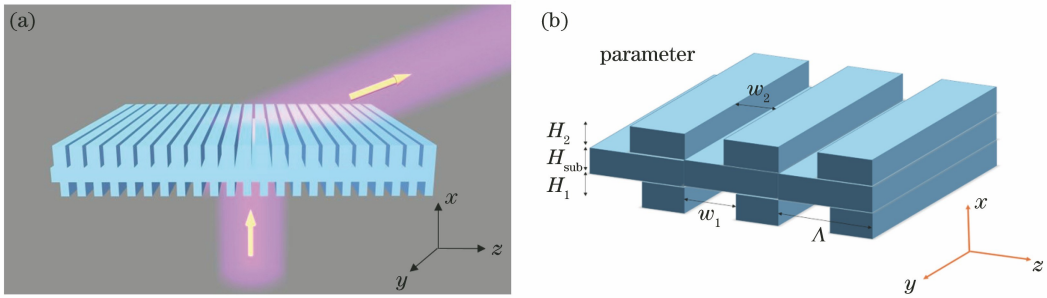

图 1　超光栅结构示意图。(a)结构及功能；(b)结构参数

Fig. 1　Structural diagram of metagrating. (a) Structure and function; (b) structural parameters

为了找到一阶衍射效率最高、带宽最宽、角色散较大的光栅结构，优化流程主要包含图 2 所示的 4 个部分，其中 MSE 为均方误差，Adam 为自适应矩估计梯度下降算法。1)数据集的制作：制作共 155 列 5000 组数据(前 5 列为结构参数，后 150 列为衍射效率点)，并且按照 1∶2的抽样比例将数据集划分为训练集与测试集；2)神经网络的训练；3)利用神经网络进行结构参数区间内的高效遍历预测；4)筛选具有宽带宽的高效光栅结构。

在光栅的 5 个结构参数中，考虑到加工精度，宽度 W_1、W_2 的变化范围设置为 0.30～2.28 mm，高度 H_1、H_2、H_{sub} 的变化范围为 0.5～6.0 mm。如果每个参数以 0.1 mm 的步长进行遍历，共有 70246400 种可能的情况。如果直接利用 RCWA 进行计算，时间过长。因此采用可取范围内参数随机生成的方法，随机生成 5000 组数据，并且利用 RCWA 计算其在 130～180 GHz 区间内的 150 个频点的一阶衍射效率。生成了一个 5000×155 的二维数据集，数据中每一行代表 5 个结构参数以及 150 个对应频点的衍射效率，使用 RCWA 计算 5000 组数据的时间为 2201 s。以这 5000 组数据为原始数据集，对其进行训练集以及测试集的分割，按照比例抽样的准则，每隔两个数据抽取一个数据加入测试集，其他加入训练集。

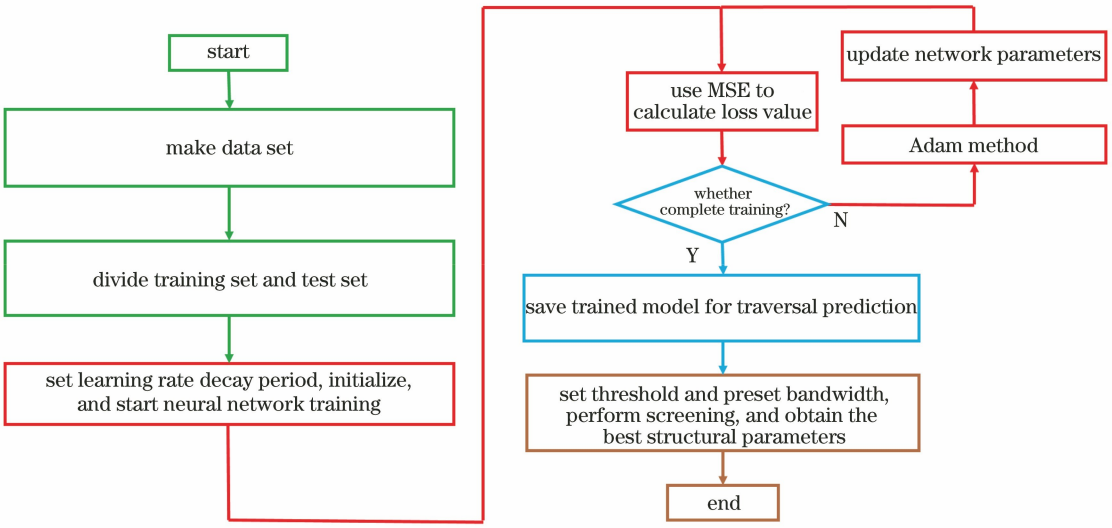

图 2　基于深度学习神经网络的光栅优化流程图

Fig. 2　Flow chart of grating optimization based on deep learning neural network

2.2　训练神经网络，保存网络模型

通过使用全连接的神经网络对衍射强度进行预测，输入 5 个结构参数，分别标为 x_1，x_2，x_3，x_4，x_5，经过 7 个隐藏层，第 k 个隐藏层的第 i 个神经元的数值 z_{ki} 由前层元素的权重乘积与偏置之和决定：

$$z_{ki}=\sum_{i=1}^{m}[w_{(k-1)i}\times z_{(k-1)i}]+b_{(k-1)i},\tag{3}$$

式中：$w_{(k-1)i}$ 为第$(k-1)$个隐藏层的第 i 个神经元的权重；$b_{(k-1)i}$ 为第$(k-1)$个隐藏层的第 i 个神经元偏置系数；m 为神经元个数。最终输出层输出的是 130～180 GHz 波段范围内 150 个等间隔频点的衍射效率。实验环境为型号为 i9-9900k 的中央处理器(CPU)，型号为 NVIDIA-RTX2080TI 的图形处理器(GPU)，内部存储器为 48 GB。图 3 为神经网络训练过程示意图。

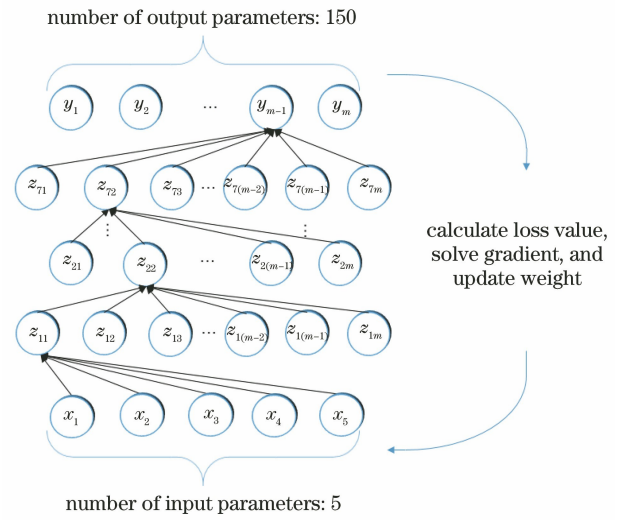

图 3　神经网络训练过程示意图

Fig. 3　Schematic of neural network training process

目标函数 L_{loss} 为各频点衍射效率与 RCWA 计算的衍射效率的均方误差(MSE)总和，即

$$L_{\text{loss}}=\frac{1}{M}\times\frac{1}{N}\times\sum_{m=1}^{M}\sum_{i'=1}^{N}[E_{\text{pred}(i')}^{(m)}-E_{\text{real}(i')}^{(m)}]^2,\tag{4}$$

式中：m 为当前光栅结构的标号；M 为预测的结构总数；N 为频点总数；$E_{\text{real}(i')}^{(m)}$ 为第 m 个结构的第 i' 个频点的真实衍射效率，由 RCWA 计算获得；$E_{\text{pred}(i')}^{(m)}$ 为第 m 个结构的第 i' 个频点的预测衍射效率，表达式为

$$[E_{\text{pred}(i')}^{(m)}]_{n=1}^{N=150}=G_{\varphi}[x_1^{(m)},x_2^{(m)},x_3^{(m)},x_4^{(m)},x_5^{(m)}],\tag{5}$$

式中：n 为当前预测频点的标号；G_{φ} 代表整个神经网络的计算过程。训练过程使用 Adam 梯度下降算法[26]：

$$\boldsymbol{g}_w\leftarrow\nabla_w L_{\text{loss}},\tag{6}$$

式中：$\boldsymbol{g}_w$ 为 L_{loss} 的梯度；w 为神经元权重。适当设置学习率 α，然后不断更新神经元权重 w：

$$w\leftarrow\text{Adam}(w,\alpha,\boldsymbol{g}_w,\beta_1,\beta_2),\tag{7}$$

式中：Adam(·)为 Adam 自适应梯度下降函数；β_1，β_2 为 Adam 梯度下降算法中的一阶、二阶梯度指数衰减率。优化目标经过反向传播[27]使 L_{loss} 的计算结果最小，即预测值趋近 RCWA 计算值，获得最佳权重参数 w^*：

$$w^*\leftarrow\text{argmax}(-L_{\text{loss}}),\tag{8}$$

学习率 α 的初始值设置为 0.001，β_1 和 β_2 用于控制一阶梯度和二阶梯度，为了使梯度方向不变时维度上的更新速度变快及梯度方向有所改变时维度上的更新速度变慢，将 β_1 和 β_2 分别设置为 0.9 和 0.99[26]。设置相应的批次大小(b_{a})以及训练周期(e_{p})。为了提高训练速度，增大训练的每批次数据量，选取 b_{a} 为 2500；为了使训练更加充分，选取较大的训练周期，e_{p} 设置为 1000。训练过程大约为 147 s，训练时间主要由训练周期以及批次大小决定，e_{p} 越大，b_{a} 越小，训练时间越长，反之则用时更短。

同时采用指数衰减调整学习率的方法[28]来获得更小的 L_{loss} 值。具体设置 A 为梯度下降学习率的变化周期，衰减因子底数设置为 0.1，每隔 A 个周期学习率变为原来的 1/10。A 对 L_{loss} 值收敛效果的影响如图 4(a)所示，随着 A 的增加，L_{loss} 值逐渐减小并趋于收敛。过小的 A 会使学习率下降过快，更新步长趋近于 0，使模型收敛到较大的 L_{loss} 值。合适的 A 可以使模型在最优点波动的时候更新步长减小，模型逐渐收敛到最优点。

图 4 (b)为 $A=300$ 的训练过程中 L_{loss} 值的变化曲线图，实线为训练过程的 L_{loss} 值，虚线为测试过程的 L_{loss} 值。模型在经过 200 个周期后 L_{loss} 值表现为上下波动；在 300 个周期后，减小学习率，L_{loss} 值进一步减小，模型逐渐收敛到最优点，最终保存训练 L_{loss} 值为 0.00095，测试 L_{loss} 值为 0.00160 的预测模型。

3　神经网络预测及最优结果

对可变参数范围内以 0.1 mm 为步长的所有可能的光栅结构参数组合进行遍历，加载训练好的神经网络模型，神经网络经过 3037 s 的计算后得到 70246400 组 150 个频点的衍射光谱预测值。通过

同一台计算机的单线程运算方式统计两种方法的耗时。RCWA 方法完成同样的计算量需要耗时 423 d。在 Matlab 中利用此神经网络预测可以快速获取光栅的衍射光谱，比 RCWA 的计算速度提高了约一万倍。选取部分预测结果与 RCWA 计算结果进行对比，如图 5 所示。

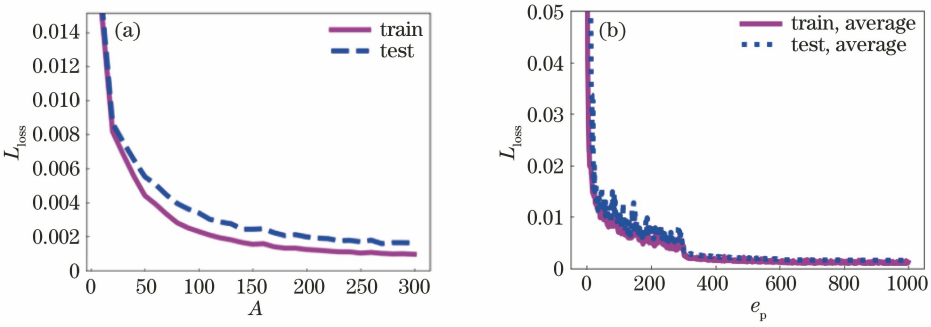

图 4　参数优化过程。(a) $e_p=1000$，$b_a=2500$ 时 A 与 L_{loss} 的关系；(b) $A=300$，$b_a=2500$ 时 e_p 与 L_{loss} 的关系

Fig. 4　Parameter optimization process. (a) Relationship between A and L_{loss} when $e_p=1000$ and $b_a=2500$; (b) relationship between e_p and L_{loss} when $A=300$ and $b_a=2500$

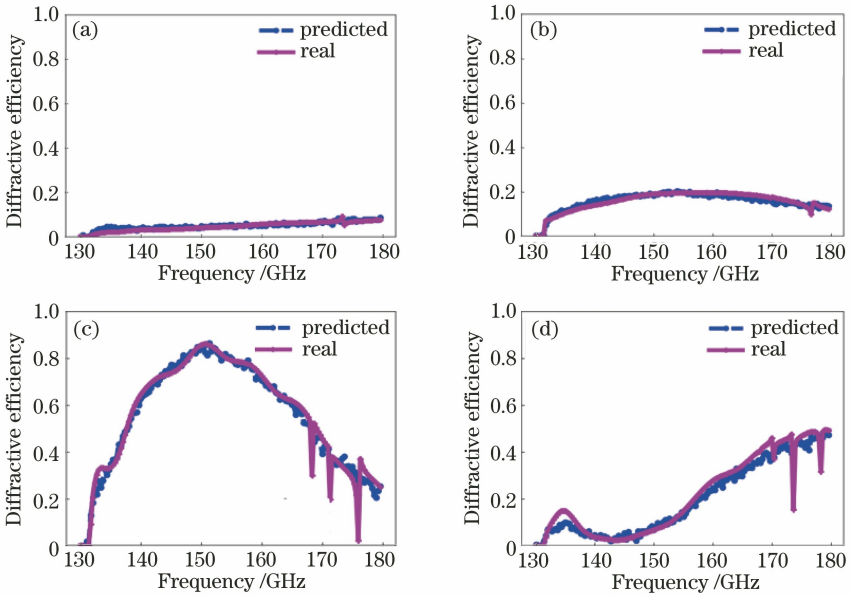

图 5　不同条件下预测的超光栅衍射光谱与实际衍射光谱。(a) $W_1=2.23$ mm，$W_2=1.61$ mm，$H_1=5.21$ mm，$H_2=0.64$ mm，$H_{sub}=1.82$ mm；(b) $W_1=1.94$ mm，$W_2=1.70$ mm，$H_1=2.09$ mm，$H_2=1.20$ mm，$H_{sub}=3.07$ mm；(c) $W_1=1.51$ mm，$W_2=0.80$ mm，$H_1=1.90$ mm，$H_2=4.83$ mm，$H_{sub}=4.79$ mm；(d) $W_1=1.60$ mm，$W_2=0.69$ mm，$H_1=2.38$ mm，$H_2=2.59$ mm，$H_{sub}=5.70$ mm

Fig. 5 Predicted diffraction spectra and actual diffraction spectra of metagratings under different conditions. (a) $W_1=2.23$ mm, $W_2=1.61$ mm, $H_1=5.21$ mm, $H_2=0.64$ mm, $H_{sub}=1.82$ mm; (b) $W_1=1.94$ mm, $W_2=1.70$ mm, $H_1=2.09$ mm, $H_2=1.20$ mm, $H_{sub}=3.07$ mm; (c) $W_1=1.51$ mm, $W_2=0.80$ mm, $H_1=1.90$ mm, $H_2=4.83$ mm, $H_{sub}=4.79$ mm; (d) $W_1=1.60$ mm, $W_2=0.69$ mm, $H_1=2.38$ mm, $H_2=2.59$ mm, $H_{sub}=5.70$ mm

由图 5 可见，神经网络对各种衍射光谱均能进行比较准确的预测，验证了神经网络预测的有效性和准确性。对于平滑的曲线，神经网络预测的效果十分理想，但是对于波动较多的频段，神经网络预测的是大致的趋势以及整体的衍射效率水平，因此其对窄谐振峰的预测并不是十分准确。这是由于窄谐振峰在数据样本中的比例过小，为了使 L_{loss} 值更小，预测值会偏向于无谐振预测，即对回归曲线进行了平滑。

为了选取衍射效率最高、带宽最宽的光栅，对衍

射效率高于80%、宽度大于60个频点(20 GHz)的波段进行筛选，得到的最佳光栅结构参数 W_1，W_2，H_1，H_{sub}，H_2 依次为 1.03 mm，0.41 mm，1.92 mm，2.79 mm，2.9 mm。图6为基于神经网络筛选出的最优光栅的衍射光谱。可以明显看出，神经网路预测的结果可以较好地反映实际的衍射光谱。RCWA 的计算结果显示，在 139.7～169.3 GHz 频段中，筛选的最佳结构均具有大于80%的衍射效率，最高衍射效率达95.7%，兼具高效与宽带宽特性。

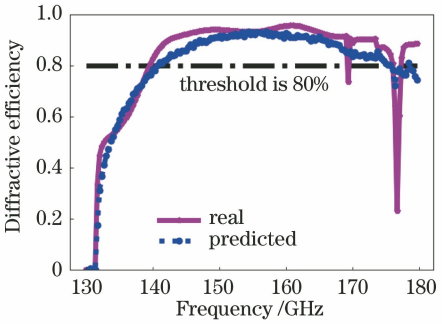

图6　基于神经网络筛选出的最优光栅的衍射光谱

Fig. 6　Diffraction spectrum of optimal metagrating screened based on neural network

在衍射效率大于80%的频率范围内，对应的衍射角变化范围为51°至70.4°，如图7(a)所示，对应的角色散为43.61 (°)/mm。图7(b)为 RCWA 计算的不同衍射角度对应的衍射效率。表1所示为不同频率太赫兹波的远场强度，其是采用二维 FDTD 仿真得到的。采用线监视器获得近场(与出射面距离约为波长的十分之一)的电场分布之后，利用近远场变换获得距离结构中心30 cm 处的场分布。由于光源面积有限(束腰半径为25 mm)，因此强度的角分布有一定程度的展宽。利用160.6 GHz 处的强度峰值对不同频点的强度均进行了归一化。不同频率太赫兹波的－1级和0级衍射强度很弱，而＋1级的衍射强度很强。

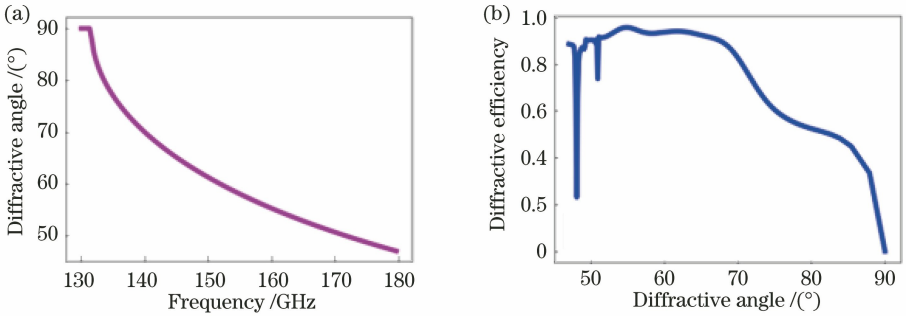

图7　光栅的衍射特性。(a)频率与衍射角度的关系；(b)衍射角与衍射效率的关系

Fig. 7　Diffraction characteristics of gratings. (a) Relationship between frequency and diffraction angle; (b) relationship between diffraction angle and diffraction efficiency

表1　远场衍射角度和强度分布

Table 1　Far field diffraction angle and intensity distributions

Incident light frequency /GHz	Diffraction angle /(°)	Intensity
139.7	51.16	0.93
145.2	55.28	1.00
151.9	60.22	0.96
160.6	65.22	0.78
169.3	70.20	0.65

图8(a)、(b)分别为139.7 GHz 和169.3 GHz 处的近场分布，对应的衍射角分别为70.4°和51°，其中 H 为归一化磁场复振幅的实部。可以看出，场分布均匀，衍射角度与理论分析结果对应，衍射效率比较理想。为了验证该结构具有高角色散性，分别计算该光栅与可见光闪耀光栅(光栅刻线密度为300 lp/mm，闪耀角为17.5°)在波长变化10%时的衍射角度的改变，结果显示，前者在波长改变10%(从1.86 mm 到2.05 mm)时衍射角度偏转9.21°，后者在波长改变10%(从2.0 μm 到2.2 μm)时衍射角度偏转3.65°。可以明显看出，神经网络筛选的太赫兹光栅的角色散更大。

当然，该光栅结构也可以采用梯度下降等数值优化方法进行优化设计，但是该方法依赖参数的初始位置，容易陷入局部最优。使用深度网络学习进行预测可以避免局部最优，在训练出可靠的模型之后，进行遍历预测，对遍历结果进行筛选，找到全局最优解。同时，当结构更为复杂、涉及参数更多时，直接数值优化的时间会呈指数增加，但是神经网络预测时间变化不大。

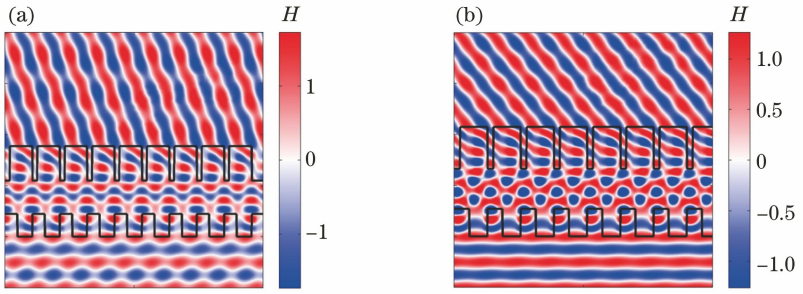

图 8　不同频率处超光栅的近场分布。(a) 139.7 GHz;(b) 169.3 GHz

Fig. 8　Near-field distributions of metagrating at different frequencies. (a) 139.7 GHz; (b) 169.3 GHz

4　结　　论

利用深度学习神经网络与传统数值计算相结合的方法,对太赫兹介质超光栅进行了快速优化设计。通过传统数值计算获得了部分超光栅结构的衍射光谱,建立数据集,训练出神经网络的最佳模型。该方法可以对各种光栅结构的衍射光谱进行准确预测,进而快速遍历并筛选出最佳性能的光栅结构,相比传统的遍历计算方法速度提高了一万倍,证明了深度学习神经网络在超表面和超光栅结构设计中的准确性和高效性。最终筛选的介质光栅峰值衍射效率为 95.7%,在 139.7～169.3 GHz 带宽范围内衍射效率大于 80%,且角色散大,当波长变化 10%时,衍射角度偏转 9.21°。同时该设计可通过 3D 打印快速加工制备,在太赫兹波段的光谱分析等应用中具有重要意义。

参　考　文　献

[1] Graves A, Fernández S, Gomez F, et al. Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks[C]//Proceedings of the 23rd International Conference on Machine Learning, June 25-29, 2006, Pittsburgh, Pennsylvania. New York: ACM Press, 2006: 369-376.

[2] Devlin J, Chang M W, Lee K, et al. Bert: pre-training of deep bidirectional transformers for language understanding[EB/OL]. (2019-05-24)[2020-08-13]. https://arxiv.org/pdf/1412.6980v8.pdf.

[3] He K M, Gkioxari G, Dollar P, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 386-397.

[4] Ma H Q, Ma S P, Xu Y L, et al. Low-light image enhancement based on deep convolutional neural network[J]. Acta Optica Sinica, 2019, 39(2): 0210004.
马红强, 马时平, 许悦雷, 等. 基于深度卷积神经网络的低照度图像增强[J]. 光学学报, 2019, 39(2): 0210004.

[5] Wu L, Lü G Q, Xue Z T, et al. Super-resolution reconstruction of images based on multi-scale recursive network[J]. Acta Optica Sinica, 2019, 39(6): 0610001.
吴磊, 吕国强, 薛治天, 等. 基于多尺度递归网络的图像超分辨率重建[J]. 光学学报, 2019, 39(6): 0610001.

[6] Campbell S D, Sell D, Jenkins R P, et al. Review of numerical optimization techniques for meta-device design[J]. Optical Materials Express, 2019, 9(4): 1842-1863.

[7] Yu M X, Zhou W C, Zhou S, et al. Manipulation on amplitude of anomalous refraction in staggered terahertz V-shaped metasurface[J]. Acta Optica Sinica, 2018, 38(3): 0318001.
余慕欣, 周文超, 周松, 等. 太赫兹 V 形错位超表面异常折射振幅调控研究[J]. 光学学报, 2018, 38(3): 0318001.

[8] Liu S F, Yuan S J, Sun Y Q, et al. Focusing properties of cylindrical vector beams through subwavelength grating lenses with negative refractive indices[J]. Acta Optica Sinica, 2019, 39(11): 1105001.
刘山峰, 袁沭娟, 孙钰淇, 等. 亚波长光栅负折射透镜的柱矢量光束聚焦特性[J]. 光学学报, 2019, 39(11): 1105001.

[9] Zhang S. AI empowered metasurfaces[J]. Light: Science & Applications, 2020, 9: 94.

[10] Dong X P, Cheng J R, Fan F, et al. Low-index second-order metagratings for large-angle anomalous reflection[J]. Optics Letters, 2019, 44(4): 939-942.

[11] Zhu C L, Jiao Q B, Tan X, et al. Fast convergent particle swarm optimization algorithm for subwavelength azimuthally polarized metal grating design[J]. Acta Optica Sinica, 2019, 39(7): 0705002.
朱春霖, 焦庆斌, 谭鑫, 等. 应用于亚波长角向偏振

金属光栅设计的快速收敛粒子群算法优化[J]. 光学学报, 2019, 39(7): 0705002.

[12] Yao K, Unni R, Zheng Y B. Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale [J]. Nanophotonics, 2019, 8(3): 339-366.

[13] He J, He C, Zheng C, et al. Plasmonic nanoparticle simulations and inverse design using machine learning [J]. Nanoscale, 2019, 11(37): 17444-17459.

[14] Qian C, Zheng B, Shen Y C, et al. Deep-learning-enabled self-adaptive microwave cloak without human intervention[J]. Nature Photonics, 2020, 14(6): 383-390.

[15] Li X Z, Shu J, Gu W H, et al. Deep neural network for plasmonic sensor modeling[J]. Optical Materials Express, 2019, 9(9): 3857-3862.

[16] Chugh S, Ghosh S, Gulistan A, et al. Machine learning regression approach to the nanophotonic waveguide analyses [J]. Journal of Lightwave Technology, 2019, 37(24): 6080-6089.

[17] Jiang J Q, Fan J A. Global optimization of dielectric metasurfaces using a physics-driven neural network [J]. Nano Letters, 2019, 19(8): 5366-5372.

[18] Peurifoy J, Shen Y C, Yang Y, et al. Nanophotonic inverse design using artificial neural network[C]// Frontiers in Optics 2017, September 18-21, 2017, Washington, D. C., United States. Washington, D.C.: OSA, 2017: FTh4A.4.

[19] Malkiel I, Mrejen M, Nagler A, et al. Deep learning for the design of nano-photonic structures [C] // 2018 IEEE International Conference on Computational Photography, May 4-6, 2018, Pittsburgh, PA, USA. New York: IEEE, 2018: 17806131.

[20] Liu D J, Tan Y X, Khoram E, et al. Training deep neural networks for the inverse design of nanophotonic structures[J]. ACS Photonics, 2018, 5(4): 1365-1369.

[21] Lin R H, Zhai Y F, Xiong C X, et al. Inverse design of plasmonic metasurfaces by convolutional neural network[J]. Optics Letters, 2020, 45(6): 1362-1365.

[22] Ma W, Cheng F, Liu Y M. Deep-learning-enabled on-demand design of chiral metamaterials[J]. ACS Nano, 2018, 12(6): 6326-6334.

[23] Jiang J Q, Sell D, Hoyer S, et al. Free-form diffractive metagrating design based on generative adversarial networks[J]. ACS Nano, 2019, 13(8): 8872-8878.

[24] Liu Z C, Zhu D Y, Rodrigues S P, et al. Generative model for the inverse design of metasurfaces [J]. Nano Letters, 2018, 18(10): 6570-6576.

[25] Dong X P, Cheng J R, Fan F, et al. Extremely large-angle beam deflection based on low-index sparse dielectric metagratings [J]. Journal of Physics D: Applied Physics, 2020, 53(24): 245101.

[26] Kingma D P, Ba J. Adam: a method for stochastic optimization[EB/OL]. (2017-01-30)[2020-08-13]. https://arxiv.org/pdf/1412.6980v8.pdf.

[27] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors [J]. Nature, 1986, 323(6088): 533-536.

[28] Feng Y X, Li Y M. An overview of deep learning optimization methods and learning rate attenuation methods[J]. Hans Journal of Data Mining, 2018, 8(4): 186-200.

冯宇旭, 李裕梅. 深度学习优化器方法及学习率衰减方式综述[J]. 数据挖掘, 2018, 8(4): 186-200.