

基于BP神经网络的面等离激元干涉仪 传感方法研究

郝秋月*, 彭润玲, 胡海峰#

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2025年4月14日; 录用日期: 2025年5月7日; 发布日期: 2025年5月14日

摘 要

本研究提出了一种融合微纳光学仿真与深度学习算法的折射率传感预测方法。针对传统狭缝-凹槽干涉结构, 通过理论建模与数值仿真, 生成不同折射率环境下的干涉条纹图像。利用主成分分析对图像数据进行特征提取与降维, 构建低维表征空间以简化输入复杂度。在此基础上, 设计基于贝叶斯正则化反向传播的BP神经网络模型, 通过引入正则化约束优化网络参数, 有效抑制过拟合现象, 实现了干涉条纹至折射率的高精度映射。仿真结果表明, 该模型预测误差低于 10^{-4} 量级, 展现出优异的稳定性与泛化潜力。然而, 受限于理想化仿真数据的有限样本量, 未来需通过实验数据扩充与噪声嵌入进一步提升模型对实际复杂场景的适应性。本研究为光学传感系统的智能化分析与设计提供了新的参考方法。

关键词

表面等离激元, 微纳光学, 光学干涉仪, BP神经网络, 折射率传感

Research on Surface Plasmon Polariton Interferometer Sensing Method Based on BP Neural Network

Qiuyue Hao*, Runling Peng, Haifeng Hu#

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Apr. 14th, 2025; accepted: May 7th, 2025; published: May 14th, 2025

*第一作者。

#通讯作者。

文章引用: 郝秋月, 彭润玲, 胡海峰. 基于 BP 神经网络的面等离激元干涉仪传感方法研究[J]. 建模与仿真, 2025, 14(5): 304-314. DOI: 10.12677/mos.2025.145395

Abstract

This study proposes a refractive index sensing prediction method integrating micro-nano optical simulations with deep learning algorithms. For traditional slit-groove interferometric structures, interference fringe images under varying refractive index environments were generated through theoretical modeling and numerical simulations. Principal component analysis was employed for feature extraction and dimensionality reduction of the image data, constructing a low-dimensional feature space to simplify input complexity. On this basis, a BP neural network model based on Bayesian regularized backpropagation was designed, where regularization constraints were introduced to optimize network parameters, effectively mitigating overfitting and achieving high-precision mapping from interference fringes to refractive indices. Simulation results demonstrated that the model exhibits a prediction error below the order of 10^{-4} , showcasing exceptional stability and generalization potential. However, limited by the idealized simulation data with finite sample size, future work should enhance the model's adaptability to practical complex scenarios through experimental data augmentation and noise embedding. This research provides a novel methodological reference for intelligent analysis and design of optical sensing systems.

Keywords

Surface Plasmon Polariton, Micro-Nano Optics, Optical Interferometer, BP Neural Network, Refractive Index Sensing

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

表面等离激元(Surface Plasmon Polaritons, SPP)作为一种在金属-介质界面传播的电磁波与自由电子集体振荡耦合形成的特殊模式,因其对周围介质折射率的极端敏感性和亚波长尺度下的强局域场增强特性,已成为高灵敏度光学传感与纳米光子学器件的核心研究对象[1]-[3]。近年来,基于 SPP 的干涉仪(如马赫-曾德尔干涉仪[4]、法布里-珀罗腔[5]等)通过将 SPP 的相位调控与干涉效应结合,在生物分子检测、环境监测及微纳尺度光场操控等领域展现出独特优势[6]-[8]。然而,传统 SPP 干涉仪普遍面临结构复杂、模式损耗高、灵敏度与集成化难以兼顾等问题,限制了其实际应用潜力。针对这一挑战,研究者提出通过微纳结构的几何优化(如狭缝、凹槽等)调控 SPP 的激发、传播与干涉过程,其中倾斜狭缝与凹槽的耦合设计因其可增强局域场、降低辐射损耗并实现多维度相位调控而备受关注[9] [10]。本文提出一种基于倾斜狭缝-凹槽结构的表面等离激元干涉仪,通过倾斜狭缝的非对称激发增强 SPP 耦合效率,结合凹槽结构的相位调制功能,构建高效、紧凑的 SPP 干涉传感结构。因为该结构的凹槽结构参数设置决定着能否产生干涉条纹,有必要先将理论计算的参数进行建模与仿真,现有研究多通过电磁仿真软件(如 COMSOL [11]、FDTD Solutions [12])获取特定结构的光场分布,得到应有的透射光干涉条纹。传统光学传感方法需依赖人工提取干涉条纹特征(如间距、相位)并基于简化物理模型计算灵敏度,其精度受限于经验公式的线性近似假设与复杂场分布的实际偏差。本研究提出的神经网络框架,通过端到端学习干涉条纹图像与目标折射率的全局非线性映射关系,彻底规避了繁琐的中间处理计算步骤。近年来,深度学习技术为光学结构的数据处理及分析提供了新思路,例如通过神经网络建立折射率与光学响应的映射模型,

可实现较高精度的折射率预测。其中反向传播(Backpropagation, BP)神经网络作为一种典型的多层前馈型人工神经网络, 凭借其强大的非线性映射能力和自学习特性, 已经成为解决复杂模式识别与回归预测问题的核心工具之一。自从 BP 神经网络通过不断迭代调整网络权值与阈值, 实现了从输入到输出的高精度非线性拟合后, 在图像处理、信号预测、工业控制等领域展现出显著优势[13]。本研究针对不同折射率下仿真得到的干涉强度值以及生成的干涉条纹图, 构建基于 BP 神经网络的折射率图像特征预测模型, 通过引入贝叶斯正则化算法(trainbr)作为网络训练函数防止过拟合, 结合主成分分析(PCA)降低输入维度等策略, 为干涉条纹的处理及传感研究提供新的方法路径。

2. 结构原理及模型

2.1. SPP 结构原理

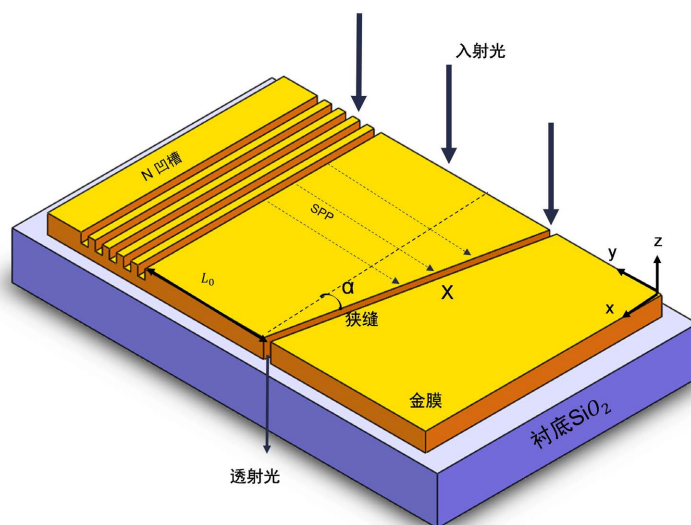


Figure 1. Schematic diagram of the slit-groove structure
图 1. 狭缝 - 凹槽结构原理图

在本文中, 我们提出了一种基于等离子体干涉仪的非光谱折射率传感方案, 如图 1 狭缝 - 凹槽结构原理图所示, 干涉仪的核心由两个功能单元构成: 1) 周期排列的金属凹槽阵列, 其周期性与入射光波长匹配, 用于高效激发表面等离子体激元(Surface Plasmon Polaritons, SPP); 2) 倾斜角为 α 的纳米狭缝阵列, 其空间位置经过特定角度设计, 用于调控 SPP 的相位积累并产生干涉效应。在垂直入射条件下, 入射光波与周期性的凹槽阵列结构发生耦合, 激发的 SPP 沿金属 - 介质界面传播至倾斜纳米狭缝处。由于纳米狭缝的倾斜排布, 不同空间位置(X 轴方向)的凹槽与狭缝之间的传播距离存在差异($\Delta L = X \cdot \tan \alpha$), 导致 SPP 在到达狭缝时累积与位置相关的相位延迟。

$$\Delta \phi = \frac{2\pi}{\lambda_{\text{spp}}} \cdot \Delta L \quad (1)$$

与此同时, 直接透过金属薄膜的传播光(未经 SPP 调制的背景光)与 SPP 介导的透射光在狭缝出口处发生相干叠加, 形成强度周期性调制的干涉条纹。当覆盖介质(如待测液体或气体)的折射率 n 发生变化时, SPP 的传播常数:

$$\beta = \frac{2\pi}{\lambda} \sqrt{\frac{\epsilon_m \cdot n^2}{\epsilon_m + n^2}} \quad (2)$$

(其中 ε_m 为金属介电常数)随之改变,进而影响相位积累量 $\Delta\phi$, 其中 λ 为入射光波长, λ_{spp} 为 SPP 的波长, 满足 $\lambda_{spp} = 2\pi/\beta$ 。这种相位变化将导致干涉条纹整体沿 X 轴方向发生平移, 其偏移量与折射率变化量 Δn 呈线性关系。通过数字图像相关算法精确测量条纹位移量, 即可实现无需光谱分辨的折射率定量检测。

2.2. 模型构建

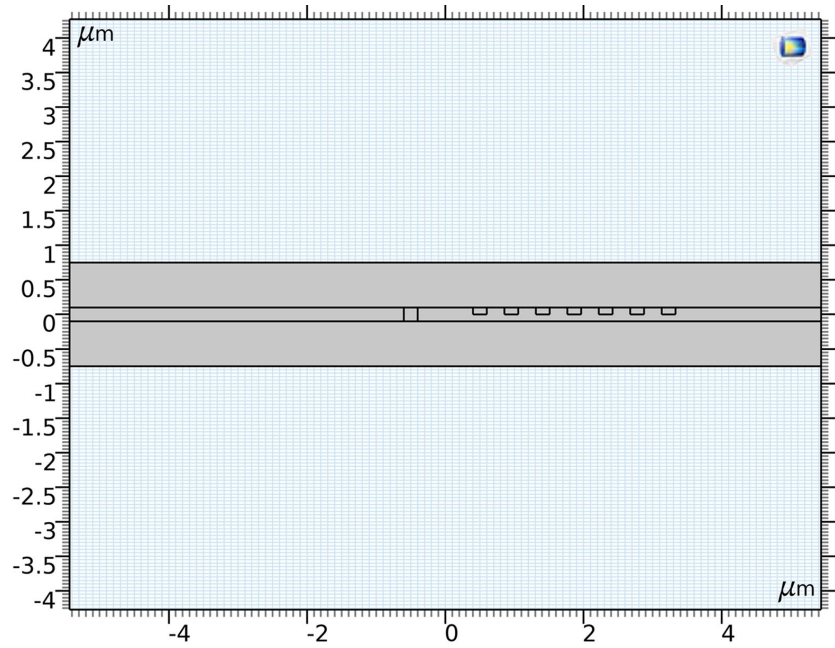


Figure 2. The 2D model of interferometer structure in COMSOL
图 2. 干涉仪结构在 COMSOL 中的 2D 模型

首先, 我们利用 COMSOL Multiphysics 软件进行数值仿真, 因凹槽的周期性结构建立侧面结构二维模型(图 2)即可得到相关的干涉信息, 分为三层, 底部为玻璃衬底, 中间是金属膜, 金属膜上面为介质层。为了更清楚地了解每个参数的含义, 建立干涉仪结构的侧面图, 如图 3 所示, 蓝色部分为折射率为 1.45 的玻璃衬底, 黄色层是金属膜, 金属膜上面为介质层, 周期性阵列结构为凹槽, 凹槽和狭缝的宽度 W 均为 200 nm, 凹槽的深度为 70 nm。通过满足波矢匹配公式(3)计算得出, 在正入射的条件下, 入射角 $\theta = 0$, 在介质折射率为水溶液($n = 1.33$)时, 凹槽的周期为 0.7λ 可以激发 SPP。介质层的金属膜选择光学性能稳定且抗腐蚀性强的金膜, 金膜厚度为 200 nm, 金膜的介电常数为 $\varepsilon = -13.692 + 1.0613 i$, 狭缝与凹槽的初始距离 L_0 为 3 μm 。

$$\beta = k_0 \sin \theta + \frac{2\pi}{p} \tag{3}$$

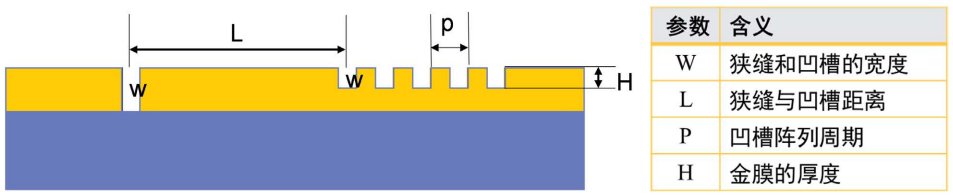


Figure 3. Parameters and their meanings corresponding to the interferometer structure
图 3. 干涉仪结构对应参数与含义

2.3. 数据集

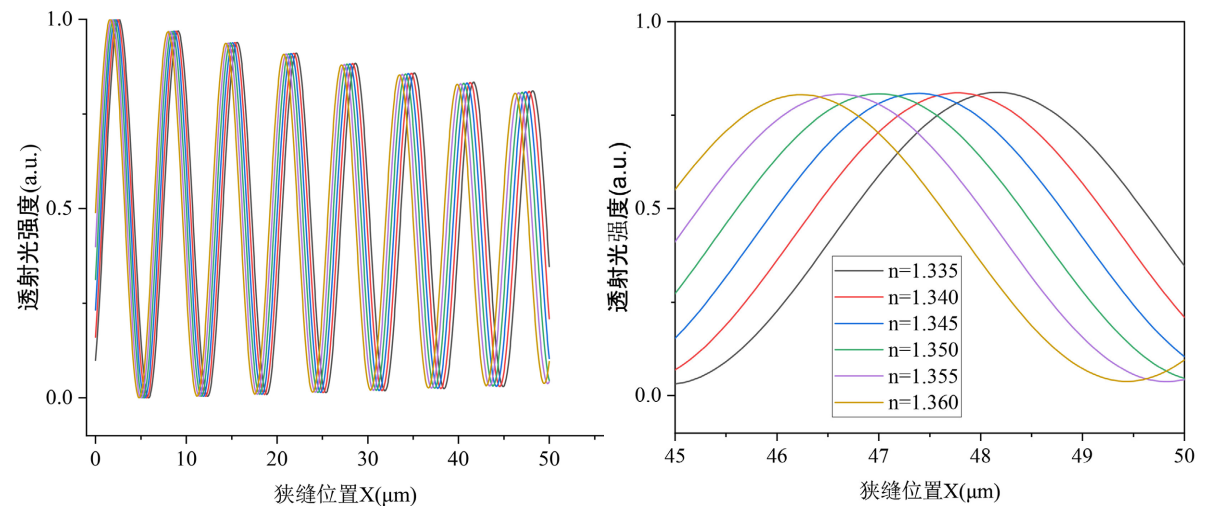


Figure 4. Variation of transmitted light intensity with different refractive indices
图 4. 不同折射率的透射光强度变化

入射波长为 650 nm 的光垂直入射到狭缝凹槽界面上，狭缝与凹槽之间的距离为 L ，参数化扫描范围 L 为 $3\ \mu\text{m}$ ~ $6.5\ \mu\text{m}$ ，对狭缝处的透射光进行采集。模拟不同折射率条件下的狭缝透射光强度变化(将介质层的折射率变化为 1.333~1.366)，我们得到了一组对应不同折射率的透射光强度值。如图 4 所示，这些强度值反映了光波在通过不同介质时的干涉效应，是后续分析和机器学习模型训练的基础数据。随后，我们采用 MATLAB 软件对仿真得到的透射光强度值进行后处理，将一维干涉强度分布利用伪彩色映射扩展为二维图像，因为狭缝与凹槽阵列之间的夹角为 α ，狭缝 - 凹槽的距离 L 与狭缝的位置存在关系： $L = L_0 + X \cdot \sin(\alpha)$ ，当狭缝 - 凹槽角度 $\alpha = 4^\circ$ ，初始狭缝 - 凹槽距 $L_0 = 3\ \mu\text{m}$ ，狭缝长度为 $50\ \mu\text{m}$ 时，根据仿真可得到狭缝位置(0~50 μm)随着折射变化的光强度变化图。每个折射率对应一个干涉条纹图，为构建机器学习模型提供了直观且信息丰富的输入数据。最后，我们将这些干涉条纹图作为神经网络模型的输入，进行训练和验证。通过学习不同折射率下的干涉图样特征，模型能够对未知折射率的样品进行准确预测，从而实现折射率的快速测定。

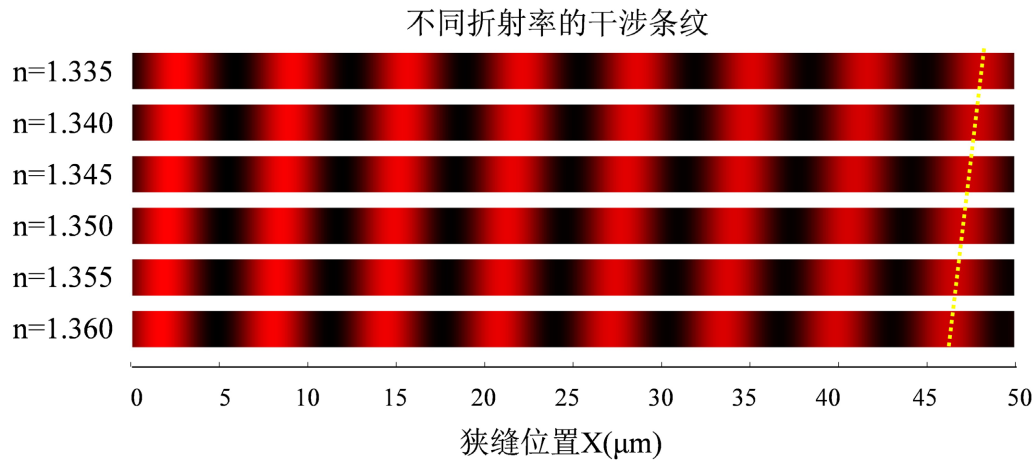


Figure 5. Changes in interference fringes with different refractive indices
图 5. 不同折射率的干涉条纹变化

图5 为介质层的折射率为 $n = 1.335$ 、 $n = 1.340$ 、 $n = 1.345$ 、 $n = 1.350$ 、 $n = 1.355$ 和 $n = 1.360$ 时的狭缝不同位置的透射光强度变化，可以看到狭缝处的透射光强度会随着介质层的折射率变化而变化。因此收集每一组折射率下光强度值对应的干涉条纹图，共收集折射率 $1.333 \sim 1.366$ ，步进为 0.001 的 34 幅干涉条纹图作为神经网络的训练输入数据集。

2.4. 网络架构

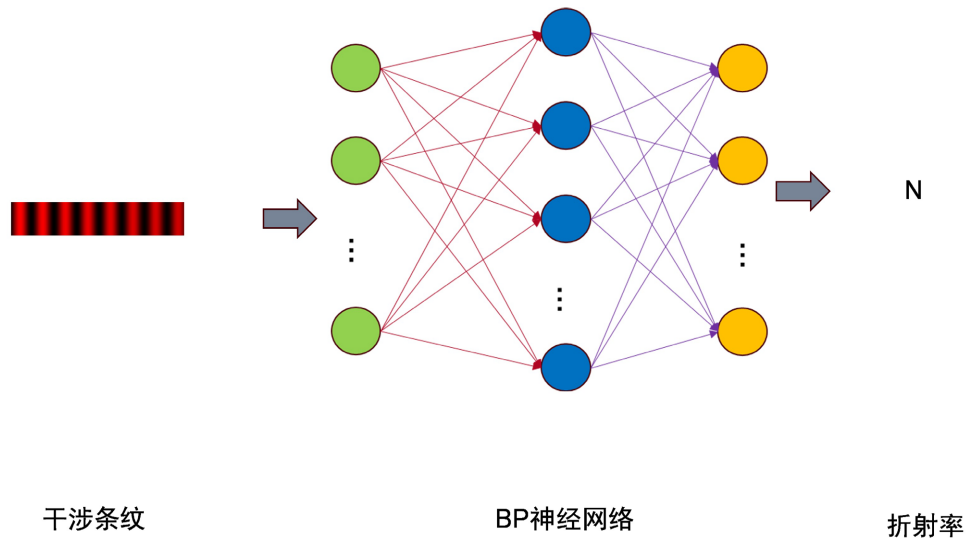


Figure 6. Flowchart of refractive index identification method based on BP model
图6. 基于BP模型的折射率识别方法流程

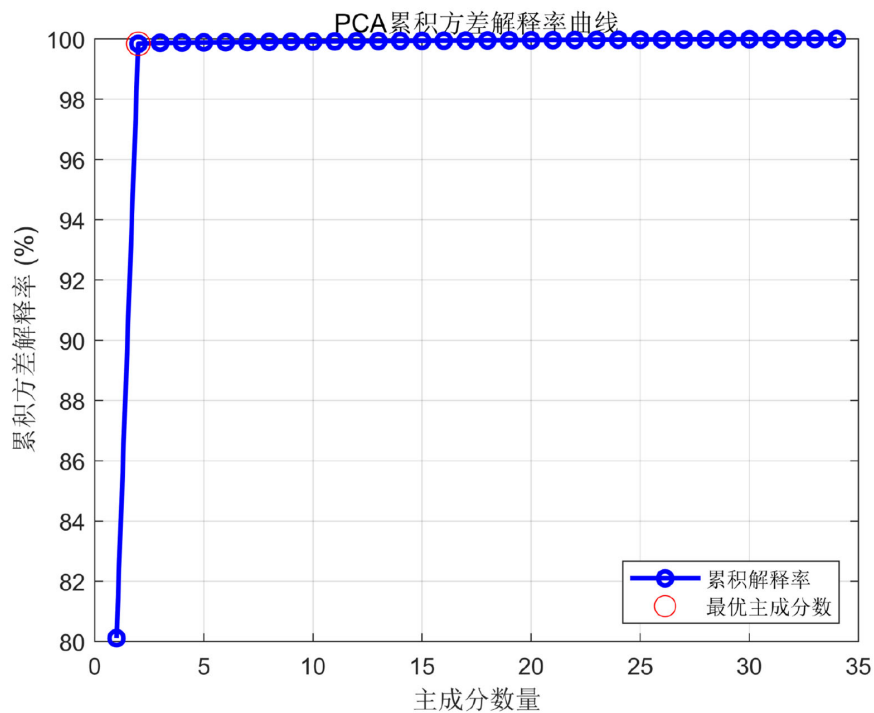


Figure 7. PCA cumulative explained variance curve
图7. PCA 累积方差解释率曲线图

BP 神经网络(Backpropagation Neural Network)是一种基于误差反向传播算法的多层前馈网络,其通过非线性映射与梯度优化实现对复杂数据关系的建模。该网络由输入层、隐藏层及输出层构成,各层神经元通过权重矩阵与偏置项全连接,形成从输入到输出的逐层信息传递机制。在前向传播阶段,输入数据通过加权求和与激活函数的非线性变换逐层传递,最终生成预测输出。

本文中使用的是 BP 神经网络模型, BP 神经网络包括的输入层、隐藏层和输出层如图 6,其中绿色圆表示输入层,蓝色圆表示隐藏层,橙色圆表示输出层。因为生成的干涉条纹图有(1200×120 像素)大小,在图片进行网络输入时,会将图片展平成一维向量,34 张图片会有(34×144 000)向量的输入,因为过大的向量输入会导致计算资源的浪费和模型训练时间的延长,接下来,需要使用 PCA 算法构建低维空间的维数。PCA 算法的作用是降低样本空间的维数,同时尽可能地保留干涉条纹图的主要特征,这有助于去除冗余信息,提高模型的学习效率和收敛速度。经过 PCA 处理后的图片保留了(34×33)向量的主成分,通常选择 95%累计贡献率的主成分。如图 7 所示,在主成分为 2 时可达到,最终经过处理后的输入数据为(34×2)向量。

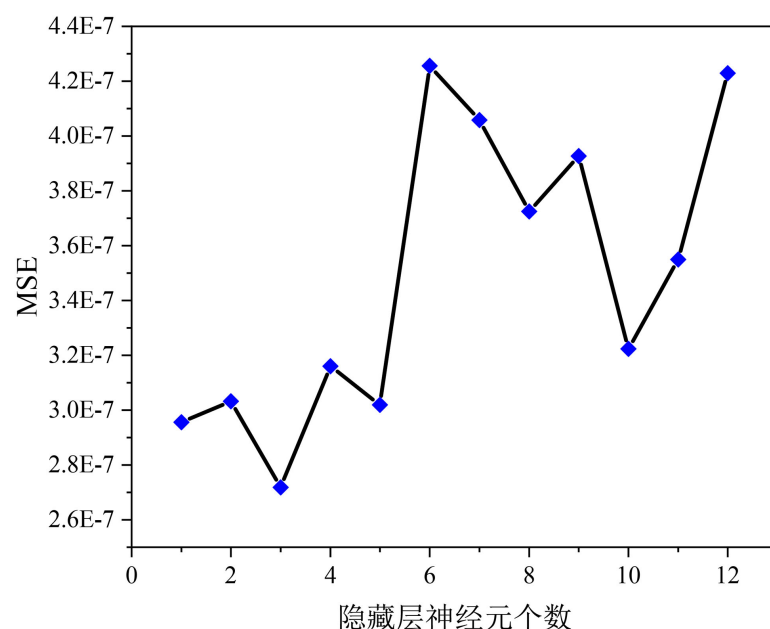


Figure 8. The impact of hidden layer neuron counts on MSE

图 8. 隐藏层神经元数量对 MSE 的影响

因为输入和输出确定以后,有必要确定隐藏层的神经元个数,隐藏层神经元数量过多可能会导致过拟合,如果神经元数量太少,则可能导致模型欠拟合,即模型无法充分学习数据中的模式。根据一般确定隐藏层的计算式子 $h = \sqrt{n + m} + c, c \in [1, 10]$, n 为输入层神经元个数, m 为输出层神经元的个数, h 为隐藏层的神经元个数,使用网格搜索法将隐藏层神经元的个数设置为(1~12),用 MSE 作为评价指标,图 8 为不同个数下的损失情况,在数量为 3 时损失最小,故隐藏层神经元个数设定为 3。BP 神经网络的结构设置为 2-3-1,其中输入层有 2 个神经元,隐藏层有 3 个神经元,输出层有 1 个神经元。采用贝叶斯正则化算法(trainbr)作为网络训练函数,该算法通过引入正则化因子动态平衡模型复杂度与拟合精度,固定学习率为 0.01,以精细调节权重更新步长,维持训练过程的数值稳定性。设定收敛阈值为目标均方误差 $MSE \leq 10^{-6}$,即期望将误差降至极小值,训练将继续,直到误差小于该设定值时停止。

神经网络的数据集划分包括训练集、验证集和测试集。训练集用于训练模型,调整权重。验证集用

于在训练过程中评估模型，调整超参数，比如学习率、层数等，防止过拟合。测试集则是最终评估模型性能的数据，确保模型泛化能力。本文将 70%的数据集进行一般神经网络的自划分(0.7:0.15:0.15)，将数据集的 30%进行单独测试。

3. 结果分析

图 9 为训练集、验证集和测试集在训练过程中的损失情况，训练曲线显示，模型在训练过程中，验证集性能表现出色。验证集的最佳 MSE 在第 18 个 epoch 时达到 9.0464×10^{-8} ，显著低于设定的目标值 10^{-6} ，表明模型在验证集上的预测误差极小，性能远超预期。同时，模型仅需 18 个 epoch 即可达到最优值，说明训练效率高，参数优化过程迅速且稳定。在评估过拟合与泛化能力方面，假设训练集与验证集的 MSE 曲线均持续下降并在最佳点后趋于平稳，且验证集未出现明显回升，这表明模型未出现过拟合现象。若测试集的 MSE 接近验证集，则进一步证明模型具有优秀的泛化能力。在训练过程的稳定性方面，总 epoch 数为 24，模型在第 18 个 epoch 后性能不再显著提升，这反映了模型在较短时间内收敛，避免了不必要的计算资源消耗。结合贝叶斯正则化(trainbr)和固定学习率(0.01)，这些设置有效控制了权重更新幅度，保障了训练的稳定性。

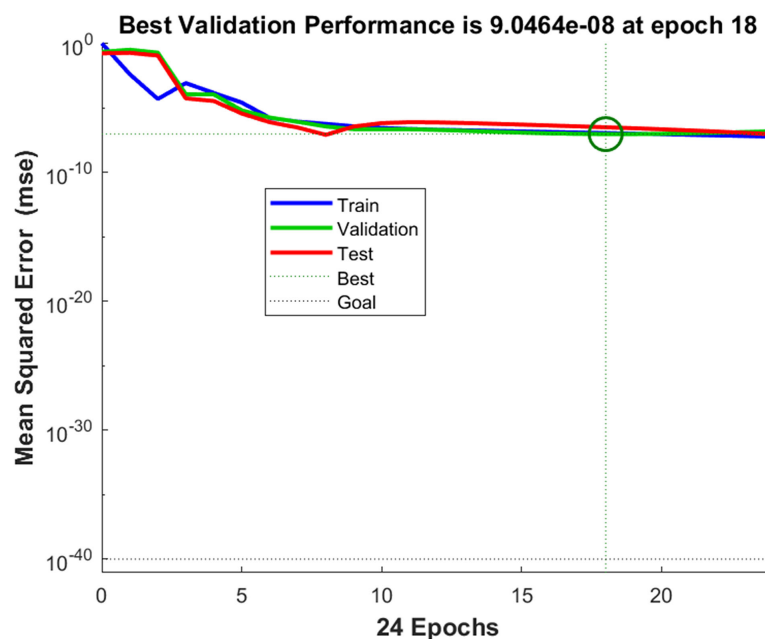


Figure 9. Variation of loss during the training process

图 9. 训练过程的损失变化

图 10 中展示了模型在训练、验证、测试及整体数据集上的拟合效果，揭示了模型在不同数据子集上的性能表现。训练集的相关系数 $R = 1$ ，表明模型在训练数据上实现了完美的拟合，没有残差，即模型能够精确地从输入特征中预测目标变量。验证集的相关系数 $R = 0.99982$ ，测试集的相关系数 $R = 0.99998$ ，以及整体数据集的相关系数 $R = 0.99997$ ，均非常接近于 1，这表明模型在未见过的数据上也具有极高的预测准确性，显示出模型具有出色的泛化能力。图中数据点与拟合直线的紧密重合进一步证实了模型预测值与实际值之间具有极高的一致性，几乎无偏差。这种高度一致性表明模型不仅在训练集上表现良好，而且在新数据上的预测同样精确，这在实际应用中是非常理想的特性。此外，模型在各个数据集上的高 R 值表明模型没有出现过拟合或欠拟合的现象，这通常是模型选择和调优过程中需要避免的问题。过拟

合会导致模型在训练数据上表现良好，但在新数据上表现差，而欠拟合则意味着模型过于简单，无法捕捉数据中的复杂关系。综上所述，该模型在训练、验证、测试及整体数据集上的表现均较好，显示出模型的稳健性，这为模型在实际应用中的有效性和可靠性提供了坚实的证据。这些结果表明，所选模型不仅能够准确地从输入特征中预测目标变量，而且能够在新数据上保持这种预测能力，这对于任何依赖于模型预测的应用场景都是至关重要的。

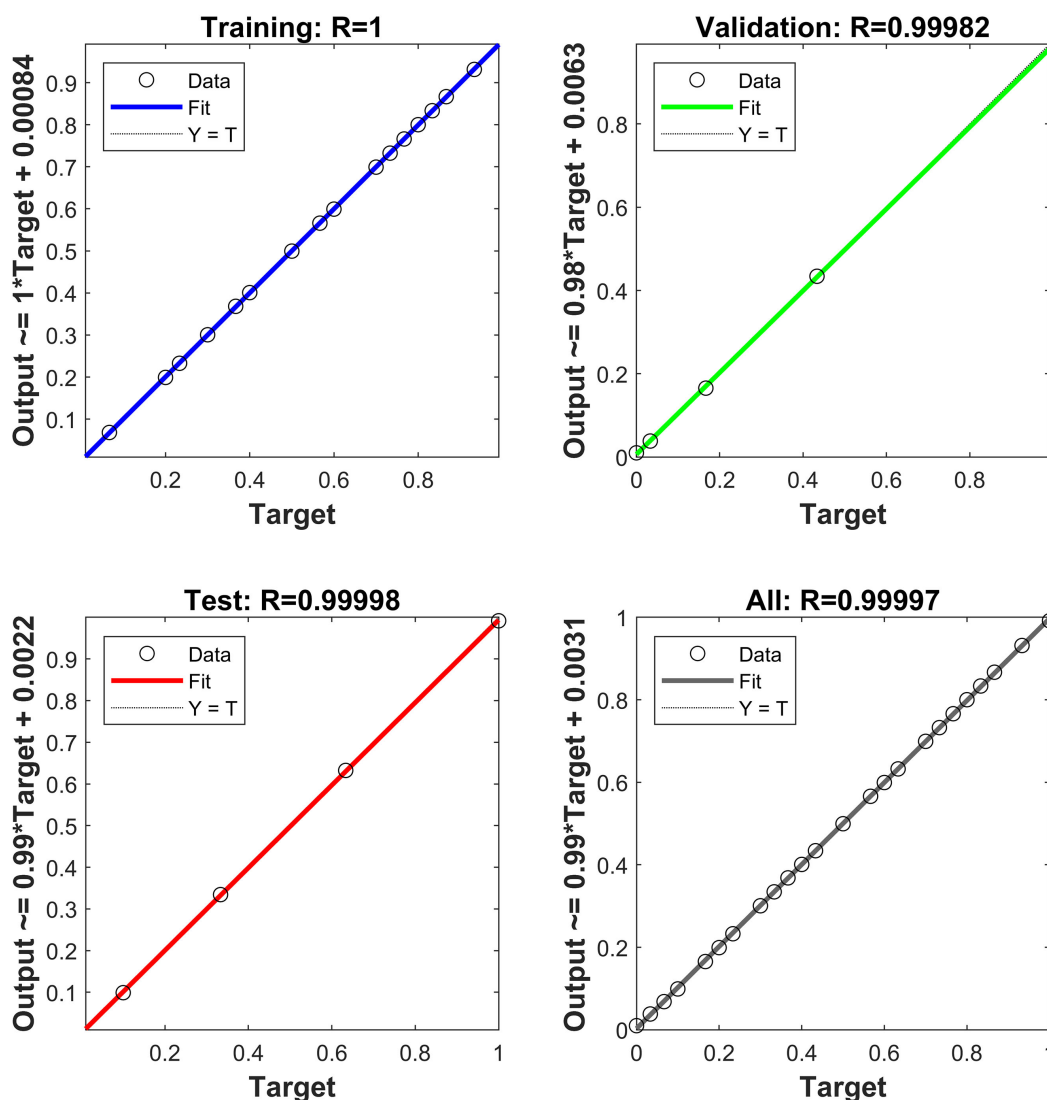


Figure 10. Correlation during the training process

图 10. 训练过程的相关性

图 11 中展示的是未进行训练的 30%测试部分，显示了折射率值预测结果与真实值的对比。横轴表示预测样本的编号，纵轴表示折射率值。图中蓝色菱形标记表示模型预测值，红色圆圈标记表示真实值。两条折线几乎重合，显示出模型预测的高精度。图 11 显示预测结果的决定系数 $R^2 = 0.99997$ ，这一数值非常接近 1，表明预测值与真实值之间具有极高的线性相关性，几乎完美地拟合了真实数据。这进一步证实了模型在测试集上具有优异的预测性能。综合来看，该模型在测试集上的预测表现较好，能够准确地预测折射率值，显示出模型具有很高的泛化能力和预测准确性。这种高精度的预测对于实际应用中的性

能评估和决策制定具有重要意义。图 12 为预测值与真实值之间的误差，该柱状图展示了模型在不同样本索引下的预测误差分布，其中误差值主要集中在 0 附近，表明模型的预测值与真实值之间的差异较小，整体预测性能较好。尽管样本索引为 3 和 7 的误差为负值，显示模型在这些样本上的预测值略低于真实值，而样本索引为 4 和 8 的误差为正值且相对较大，表明在这些样本上的预测值略高于真实值，但整体误差较小，可以看到预测的误差在 10^{-4} 量级以内，已经能够满足识别精度。

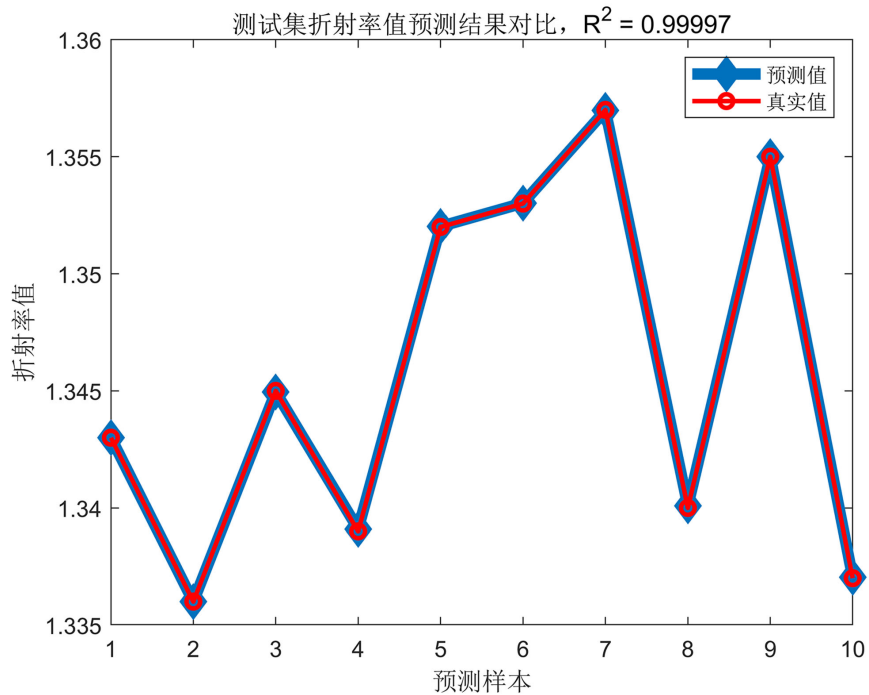


Figure 11. Comparison of refractive index prediction results on the test set

图 11. 测试集折射率预测结果对比

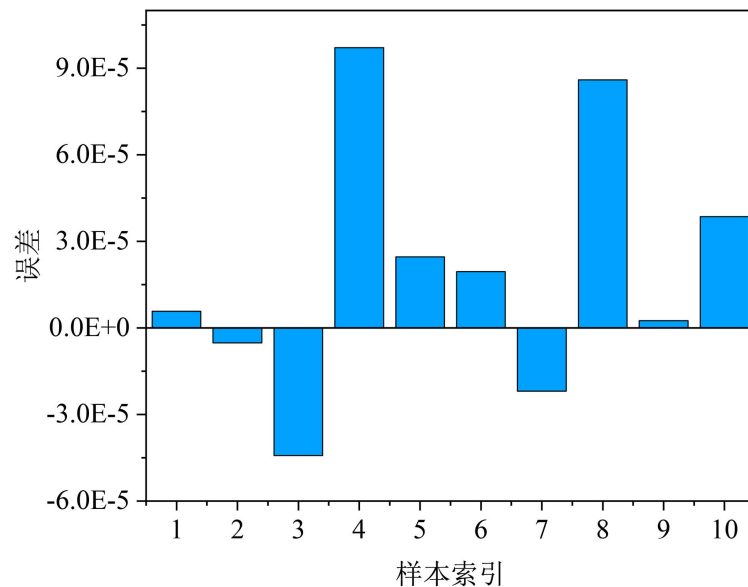


Figure 12. Error between predicted values and true values on the test set

图 12. 测试集预测值与真实值的误差

4. 结论

本研究将传统的狭缝-凹槽传感干涉结构进行了模拟仿真, 根据理论计算的凹槽周期进行建模和仿真, 并成功得到了干涉条纹, 将仿真得到的数据生成对应的干涉条纹图, 并结合 BP 神经网络作为处理工具进行图像处理和预测, 利用 PCA 算法构建低维空间的维数, 简化计算量, 并采用贝叶斯正则化反向传播的训练函数 `trainbr`, 通过引入正则化项, 避免过拟合, 建立的网络模型稳定且对折射率预测结果准确, 对于以后实验得到的干涉条纹图分析提供了一种新的研究方法。不足的点在于由于仿真得到的数据为理想值, 不同折射率下的干涉条纹变化规律仅需要每种环境的干涉数据值即可进行训练预测, 这使得样本量过少, 过拟合的风险增加, 因此在实际实验过程中应增加样本量, 使网络模型充分学习每种折射率的特征, 增加模型的泛化能力。总之, 本研究为折射率传感预测提供了一种有效方法。

参考文献

- [1] Zhang, J., Zhang, L. and Xu, W. (2012) Surface Plasmon Polaritons: Physics and Applications. *Journal of Physics D: Applied Physics*, **45**, Article 13001. <https://doi.org/10.1088/0022-3727/45/11/113001>
- [2] Maier, S.A. and Atwater, H.A. (2005) Plasmonics: Localization and Guiding of Electromagnetic Energy in Metal/Dielectric Structures. *Journal of Applied Physics*, **98**, Article 011101. <https://doi.org/10.1063/1.1951057>
- [3] Ritchie, R.H. (1957) Plasma Losses by Fast Electrons in Thin Films. *Physical Review*, **106**, 874-881. <https://doi.org/10.1103/physrev.106.874>
- [4] Liu, S., Wang, Y., Xie, T., *et al.* (2025) Highly Sensitive Optical Fiber Temperature Sensor Based on Mach-Zehnder Modal Interferometer with Taper Structure. *Optics Communications*, **582**, Article 131640. <https://doi.org/10.1103/PhysRev.106.874>
- [5] 张仁龙, 孔德旭, 张嘉伟, 等. 基于光子晶体光纤和空芯光纤的法布里-珀罗干涉型气压传感器[J]. 光学学报, 2024, 44(21): 54-63.
- [6] Homola, J. (2008) Surface Plasmon Resonance Sensors for Detection of Chemical and Biological Species. *Chemical Reviews*, **108**, 462-493. <https://doi.org/10.1021/cr068107d>
- [7] Yesilkoy, F., Terborg, R.A., Pello, J., Belushkin, A.A., Jahani, Y., Pruneri, V., *et al.* (2017) Phase-Sensitive Plasmonic Biosensor Using a Portable and Large Field-of-View Interferometric Microarray Imager. *Light: Science & Applications*, **7**, Article 7152. <https://doi.org/10.1038/lsa.2017.152>
- [8] 施伟华, 王华晋, 上官铭宇. 金纳米粒子增强表面等离子体共振传感[J]. 激光与光电子学进展, 2024, 61(13): 199-205.
- [9] Eftekharinia, B., Moshaii, A. and Dabirian, A. (2017) Design of a Slit-Groove Coupler for Unidirectional Excitation of the Guided Surface Plasmon Polaritons through a Plasmonic Slot Waveguide. *Plasmonics*, **12**, 131-138. <https://doi.org/10.1007/s11468-016-0238-x>
- [10] 孙书超, 孟阔. 凹槽修饰的太赫兹亚波长狭缝透射特性[J]. 北京信息科技大学学报(自然科学版), 2023, 38(5): 67-73, 87.
- [11] 李丙宏, 祁云平. 一种基于 COMSOL 的 SPP 频域建模研究[J]. 电子世界, 2018(20): 78, 80.
- [12] 陈淑婷, 张乾坤, 杨县超, 等. 高灵敏度低折射率的双开环通道 PCF-SPR 传感器设计[J]. 仪表技术与传感器, 2023(11): 7-11.
- [13] LeCun, Y., Bengio, Y. and Hinton, G. (2015) Deep learning. *Nature*, **521**, 436-444. <https://doi.org/10.1038/nature14539>