

引用格式: LIU Chi, LIU Jihong, BU Zichuan, et al. Rapid Demodulation of Ellipsometric Spectra of Non-Absorbing Thin Films: a Cauchy Model-Powered Neural Network Method[J]. Acta Photonica Sinica, 2026, 55(7):0730002

刘驰,刘继红,卜子川,等. 面向无吸收薄膜的椭偏光谱快速解调:柯西模型驱动神经网络方法[J]. 光子学报, 2026, 55(7): 0730002

面向无吸收薄膜的椭偏光谱快速解调:柯西模型驱动神经网络方法

刘驰,刘继红*,卜子川,李嘉琪,张家庚,景佳豪,任凯利,严学文,董军

(西安邮电大学 电子工程学院, 西安市 710121)

摘 要:光谱椭偏技术因其高精度、非破坏性等优势,已成为测量纳米薄膜参数的重要手段。传统的椭偏光谱数据解调方法依赖于复杂的物理模型进行非线性迭代拟合,存在计算效率低、过程烦琐的局限,难以满足现代工业的快速检测需求。基于神经网络的数据处理方法具有高效的数据处理能力,为解决此问题提供了新的可能,但用于神经网络训练的数据集获取困难,这制约了该方法在光谱椭偏数据解调中的应用。针对这一问题,本文面向单层无吸收介质薄膜的椭偏光谱数据解调,提出了一种柯西色散模型驱动的神经网络方法。该方法首先基于柯西模型生成折射率数据集,然后根据偏振光薄膜干涉理论生成椭偏光谱数据集,用于神经网络的训练。解调光谱数据时,神经网络从椭偏光谱得到薄膜厚度和柯西色散模型参数后,再根据柯西模型重构折射率。在多种厚度的 SiO_2 、 MgF_2 、 Si_3N_4 和 TiO_2 材料的实验结果表明:通过该方法训练出的神经网络对薄膜的平均厚度解调误差小于 1.5 nm,平均折射率解调误差小于 0.02,平均解调时间约为 6 ms,其误差与效率均优于 LM 算法。该方法为获取椭偏光谱数据集和使用神经网络快速解调椭偏光谱提供了一种可行的途径。

关键词:光谱椭偏法;神经网络;柯西色散模型;数据集;无吸收薄膜

中图分类号: TN247

文献标识码: A

doi: 10.3788/gzxb20265507.0730002

0 引言

由于无损、非接触以及高精度等优点,光谱椭偏法(Spectroscopic Ellipsometry, SE)已经成为纳米薄膜厚度和光学常数的标准表征方法^[1-3],被广泛应用于光伏、集成电路等行业的镀膜质量检测中^[4-5]。SE 测量基于偏振光在薄膜上反射前后偏振状态的变化,具体表现为多波长下两个正交偏振分量 p 和 s 的振幅变化比 ϕ 和相位差 Δ ,即椭偏光谱 $(\phi, \Delta) \sim \lambda$,通过算法解调该光谱可以得到薄膜的厚度、折射率等信息^[6-8]。

对椭偏光谱进行解调是一个复杂的反演问题,难点在于无法从 $(\phi, \Delta) \sim \lambda$ 直接获取薄膜参数,而是需要求解偏振光薄膜干涉的逆过程^[9]。最经典的解调方法是基于预设的薄膜光学结构模型以及介电函数模型,通过非线性回归分析迭代优化模型参数,使模拟椭偏光谱与实验椭偏光谱之间的误差达到最小,从而反演出薄膜信息^[10]。然而在该方法中,模型的选择高度依赖使用者的经验,需要从 Cauchy^[8]、Lorentz^[11]、Tauc-Lorentz^[12] 等模型中做出合理选择,而且优化迭代对初值敏感,易陷入局部最优,影响结果的稳定性,另外,复杂的迭代过程导致计算效率较低。这些限制导致经典解调方法操作复杂、耗时较长,难以满足工业生产中大批量薄膜样品的快速检测需求。

神经网络能够建立椭偏光谱到薄膜信息的直接映射关系,避免了经典方法中人工依赖性强、求解效率低等问题,表现出巨大的潜力^[13-16]。目前基于神经网络的椭偏光谱解调方法研究中,训练数据集通过理论模

基金项目:国家自然科学基金(12274341),陕西省重点研发计划项目(2024QY2-GJHX-39),陕西省科技厅重点研发计划(2025CY-YBXM-114),陕西省科技厅重点研发计划(2025CY-YBXM-077)

第一作者:刘驰(2002—),男,学生,硕士,主要研究方向为光电传感与测量。Email: liuchi0257@163.com

通讯作者:刘继红(1978—),男,教授,博士,主要研究方向为光电传感与测量。Email: ong1705@xupt.edu.cn

收稿日期: 2026-01-18; **录用日期:** 2026-05-07

<http://www.photon.ac.cn>

拟的方式生成,即从数据库获取介电函数数据,使用平移和缩放的方法对其扩展,再结合偏振光薄膜干涉理论生成椭圆偏光谱数据集。由于数据库包含的材料分布比较离散,覆盖精细度不足,基于此类数据源生成的椭圆偏光谱在特征空间中分布不均,存在明显盲区,从而削弱了神经网络的鲁棒性。此外,数据库中的折射率数据相对固定,难以针对特定任务进行优化与增强,灵活性较差。

针对神经网络方法中存在的问题,本文面向无吸收薄膜的椭圆偏光谱数据解调,提出了一种基于柯西色散模型的神经网络解调方法。在该方法中,首先根据薄膜材料色散程度确定柯西模型参数范围,在参数范围中等间隔采样生成折射率数据集,然后根据偏振光薄膜干涉理论生成椭圆偏光谱数据集。基于此数据集,训练出以薄膜厚度和柯西模型参数为解调目标的神经网络(Cauchy-model-Powered Spectroscopic Ellipsometry Neural Network, CPSENET)。这一结合数据集生成与网络输出参数设计的整体框架,构成了柯西模型驱动的椭圆偏光谱神经网络解调方法(Cauchy-model-Powered Spectroscopic Ellipsometry Neural Network Method, CPSENNM)。

使用CPSENET对多种厚度的 SiO_2 、 MgF_2 、 Si_3N_4 和 TiO_2 薄膜椭圆偏光谱测试集进行解调,其结果验证了CPSENNM的有效性。并且CPSENET的解调精度及耗时优于列文伯格-马夸尔特算法(Levenberg - Marquardt, LM)。在讨论中,对CPSENET的损失函数权重参数进行优化,并分析了CPSENNM的定向数据增强优势,最后分析了材料吸收对该方法性能的影响并给出了适用材料的消光系数最大阈值。

1 理论基础

1.1 偏振光薄膜干涉理论

偏振光发生薄膜干涉效应后,其两个正交分量 p 和 s 的振幅变化量比值 ψ 、相位变化量差值 Δ 可表示为公式(1):

$$\tan \Psi e^{i\Delta} = \frac{r_p}{r_s} = \left(\frac{r_{01,p} + r_{12,p} e^{-i2\beta}}{1 + r_{01,p} r_{12,p} e^{-i2\beta}} \right) / \left(\frac{r_{01,s} + r_{12,s} e^{-i2\beta}}{1 + r_{01,s} r_{12,s} e^{-i2\beta}} \right) \quad (1)$$

其中 $r_{xy,z}$ 为 $z(z \in \{p, s\})$ 分量在 $x, y(x, y \in \{0, 1, 2\})$,其中0、1和2分别表示空气、薄膜和基底)界面处的反射系数, β 为干涉光程差,计算方式为公式(2):

$$\beta = \frac{2\pi N d \cos \theta}{\lambda} \quad (2)$$

其中,薄膜的复折射率 $N = n - ik$, n 和 k 分别为薄膜的折射率和消光系数, d 为薄膜的厚度, θ 为偏振光的入射角度, λ 为波长。多波长处的 ψ 和 Δ 数据构成椭圆偏光谱 $(\psi, \Delta) \sim \lambda$ 。

根据式(1)、(2)中薄膜厚度 d 、复折射率 N 和 $(\psi, \Delta) \sim \lambda$ 之间的关系,通过数据解调算法即可解调出薄膜的厚度和光学常数等信息。本文的应用对象为无吸收材料,即 $k = 0$, N 可以简化为 n 。

1.2 Cauchy 介电函数模型

Cauchy模型是描述无吸收材料色散特性常用的经验模型,具有表达形式简单,参数数目少的特点,具体表达式为公式(3):

$$n = A + \frac{B}{\lambda^2} + \frac{C}{\lambda^4} + \dots \quad (3)$$

在Cauchy模型中,各阶系数对折射率的影响随阶数升高而降低,一般只考虑前三阶 A 、 B 和 C 。 A 决定折射率的基线值, B 控制折射率随波长变化的幅度, C 用来修正由 A 和 B 描述的理论折射率与实际折射率之间的偏差。由于大多材料在紫外波段具有明显的吸收特性,因此光谱范围选择为380 nm-1100 nm的可见光到近红外波段。

2 数据生成与神经网络设计

2.1 椭圆偏光谱数据生成

本文所设计的基于柯西色散模型的椭圆偏光谱数据生成方法包含以下四个步骤:

步骤1:判断材料的吸收特性。如果材料的消光系数大于阈值,则此方法不适用。

步骤2:确定柯西模型参数范围。当材料吸收特性小于阈值时,则根据折射率范围 $[n_{\min}, n_{\max}]$ 确定柯西参数 A 的范围 $[A_{\min}, A_{\max}]$, B 的范围 $[B_{\min}, B_{\max}]$ 通过折射率变化范围 $[\Delta n_{\min}, \Delta n_{\max}]$ 和柯西模型按公式(4)计算得到,其中 λ_1 和 λ_2 分别为光谱范围波长上下界。 C 的参数空间 $[C_{\min}, C_{\max}]$ 设定为经验取值范围 $[1 \times 10^8, 5 \times 10^9] \text{ nm}^4$,该范围覆盖常见无吸收光学材料在可见至近红外波段的高阶色散特性。若待测材料的折射率范围及其变化范围未知,则将参数 A 与 B 的采样范围分别设定为常见光学材料的典型取值区间: $[1.3, 2.5]$ 和 $[1 \times 10^3, 3 \times 10^4] \text{ nm}^2$,以确保能够覆盖足够的折射率变化范围,增强模型的解调能力。

$$[B_{\min}, B_{\max}] = \left[\frac{\Delta n_{\min}}{\frac{1}{\lambda_2^2} - \frac{1}{\lambda_1^2}}, \frac{\Delta n_{\max}}{\frac{1}{\lambda_2^2} - \frac{1}{\lambda_1^2}} \right] \quad (4)$$

步骤3:生成折射率数据集。确定参数 A 、 B 和 C 的取值范围后,采用等间隔采样与随机采样相结合的方式生成数据集。首先在 A 和 B 的参数空间内进行等间隔采样,得到 A 和 B 的组合,然后对每一组 (A, B) ,在 C 的参数空间内随机取值并进行组合,最后使用柯西公式生成折射率数据集。

步骤4:生成椭偏光谱数据集。根据已知信息确定薄膜厚度 d 的参数范围 $[d_{\min}, d_{\max}]$,在此范围内进行等间隔离散采样,得到厚度取值集合。将厚度集合中的每个值与折射率数据集中的每一组折射率组合,使用公式(1)生成模拟椭偏光谱。为增强数据的真实性,对模拟光谱添加随机噪声,最终获得可用于网络训练与验证的椭偏光谱数据集。

需要说明的是,虽然通过此方法生成的部分折射率数据没有实际的材料与之对应,但其所反映的色散特性仍遵循柯西色散模型,并且处于合理的范围之内,因此通过这些折射率数据生成的椭偏光谱仍能有效地训练神经网络。与折射率数据库中无吸收材料的折射率数据相比,此方法生成的数据在折射率空间中分布更均匀、更密集,能有效填补真实数据集的稀疏区域,减少模型盲区。

本文按照基于柯西色散模型的椭偏光谱数据生成方法得到训练神经网络的数据集,对 MgF_2 、 SiO_2 、 Si_3N_4 和 TiO_2 这四种可见光到红外波段无吸收薄膜材料的椭偏光谱进行解调。根据它们的折射率范围以及折射率变化幅度计算出 A 、 B 的参数范围分别为 $[1.35, 2.0]$ 和 $[2.5 \times 10^3, 2 \times 10^4] \text{ nm}^2$,并在此范围内分别等间隔取50、20个数据点。膜系结构设置为空气/光滑薄膜/硅基结构,薄膜的厚度 d 在 $[100, 200] \text{ nm}$ 范围等间隔取50个数据点,式(2)中偏振光入射角 θ 设置为 70° 。根据公式(1)生成椭偏光谱数据,并对每个数据添加大小为 $[-3, 3]$ 的随机噪声。通过此方式共生成了50000条椭偏光谱数据,每条椭偏光谱数据包含721个离散波长点的 (ψ, Δ) 响应及其对应的薄膜参数标签 (d, A, B, C) 。训练神经网络时随机抽取其中的80%作为训练集,另20%作为验证集。

2.2 神经网络模型选取与训练

ResNet作为一种性能优异的经典神经网络架构,被广泛应用于分类、回归及检测等任务中。CPSENET选用的网络结构具体为ResNet-18^[17],将其各模块设置为一维,以适应椭偏光谱的数据特性。CPSENET以两条椭偏光谱作为网络输入,输出结果是薄膜厚度 d 及柯西模型参数 A 、 B 和 C 。相比于网络直接输出多波长点的材料折射率,该输出参数设计有效降低了模型复杂度。

$$\begin{aligned} \text{loss} = & \left(W_d \cdot (d_{\text{de mod}} - d_{\text{theo}}) \right)^2 + \left(W_A \cdot (A_{\text{de mod}} - A_{\text{theo}}) \right)^2 \\ & + \left(W_B \cdot (B_{\text{de mod}} - B_{\text{theo}}) \right)^2 + \left(W_C \cdot (C_{\text{de mod}} - C_{\text{theo}}) \right)^2 \end{aligned} \quad (5)$$

在CPSENET的训练参数设置中,优化器选用Adam,batch size设置为256,初始学习率设置为0.01,每过50轮学习率下降至之前的十分之一,一共训练200轮。实验中采用如公式(5)所示的加权均方误差(Weighted Mean Square Error, WMSE)来平衡各参数数量级的差异,其中,厚度权重 $W_d = 1$,柯西模型参数 A 、 B 和 C 的权重分别为 $W_A = 1$, $W_B = 1 \times 10^{-3}$ 和 $W_C = 1 \times 10^{-9}$ 。下标de mod和theo表示网络对光谱的解调结果及理论数据。

2.3 网络评估方法

本文采用直接参数误差-重构光谱误差联合评估方法,来反映神经网络解调结果的准确性及可靠性。在直接参数误差评估中,根据神经网络解调出的薄膜厚度 d 计算厚度误差 d_{error} ,将解调出的柯西模型参数 A 、 B 和 C 转换为折射率 n ,并计算折射率误差 n_{error} ,以此避免网络输出的柯西模型参数与折射率之间的割裂

感,将 d_{error} 和 n_{error} 作为衡量标准来直接检验网络解调结果的准确性。在重构光谱评估中,将 d 和 n 代入公式(1)重构椭圆偏光谱,并计算重构光谱误差 ψ_{error} 和 Δ_{error} ,通过重构光谱误差来验证网络的输出结果是否满足偏振光薄膜干涉理论。

3 实验结果

本实验使用英伟达 NVIDIA GeForce RTX 4090D GPU 耗时 8min 完成对 CPSENET 的训练,每一轮训练中,网络在训练集和验证集的 WMSE 如图 1 所示。训练过程呈现为两个阶段:(1)快速收敛阶段(0-50 轮),WMSE 在训练初期快速下降,网络建立椭圆偏光谱与薄膜参数的初步映射关系;(2)精细优化阶段(50-200 轮),WMSE 进入平稳下降期,伴随小幅度波动,网络的参数被精细调整,逐步优化映射关系,直至 WMSE 收敛且稳定。

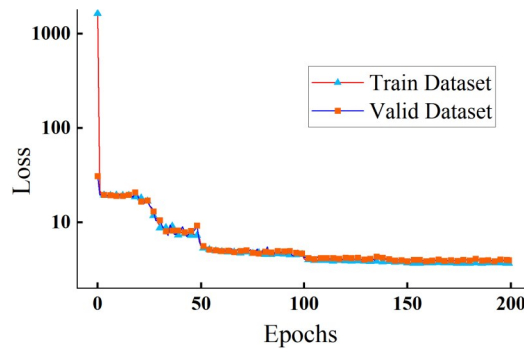


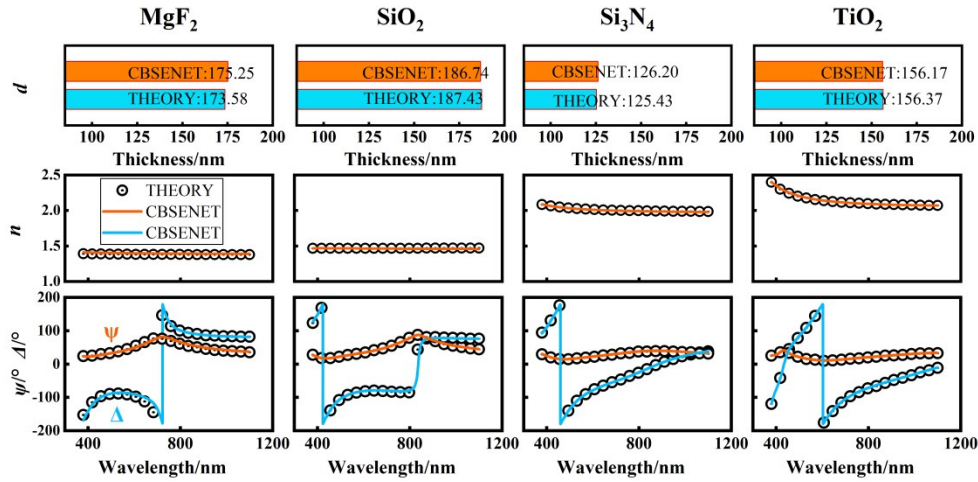
图 1 模型训练过程中的 WMSE
Fig. 1 WMSE Evolution in Model Training

使用 CPSENET 对椭圆偏光谱个例进行解调以直观呈现神经网络输出结果。根据公式(1)生成了[100, 200] nm 范围内随机厚度的 MgF_2 、 SiO_2 、 Si_3N_4 和 TiO_2 薄膜材料的椭圆偏光谱,并添加大小在[-3, 3]之间的随机噪声。CPSENET 对其解调出的 d 、 A 、 B 和 C 参数如表 1 所示,解调结果与理论值之间高度吻合,且误差整体相对较小,表明该网络解调结果的合理性。虽然 CPSENET 对 MgF_2 薄膜椭圆偏光谱解调出的 B 和 C 与理论值之间的误差相对较大,但是由于 B 和 C 本身的变化范围比较大,且属于柯西模型的高阶项,对折射率的影响较小,因此不会造成较大的折射率误差。基于表 1 中的参数重构的折射率与椭圆偏光谱如图 2 所示,对比重构光谱与理论光谱可见,二者在整个光谱范围内均呈现高度一致性,进一步验证了该网络输出结果的物理可靠性。

表 1 CPSENET 对随机厚度 MgF_2 、 SiO_2 、 Si_3N_4 及 TiO_2 薄膜的椭圆偏光谱的直接解调结果

Table 1 Direct Demodulation Results of CPSENET for Random-Thickness MgF_2 , SiO_2 , Si_3N_4 , and TiO_2 Thin Films					
	method	d/nm	A	$B \times 10^{-3}/\text{nm}^2$	$C \times 10^{-9}/\text{nm}^4$
MgF_2	THEORY	173.88	1.3756	2.71	0.071
	CPSENET	174.25	1.3879	4.30	0.004
SiO_2	THEORY	187.43	1.4460	4.77	0.151
	CPSENET	186.74	1.4557	4.40	0.371
Si_3N_4	THEORY	125.43	1.9734	14.16	0.273
	CPSENET	124.97	1.9667	14.28	0.483
TiO_2	THEORY	156.37	2.0521	20.67	3.988
	CPSENET	155.90	2.0479	22.65	4.141

在训练神经网络时,将生成的 50000 条椭圆偏光谱数据按照 8:2 的比例随机划分为训练集和验证集,并对 CPSENET 进行训练,对数据集进行随机划分的方式可能会影响神经网络的性能稳定性。为评估数据集划分对模型稳定性的影响,本文进一步采用 5 折交叉验证对其进行分析。具体地,将数据集等分为 5 折,每次将其中一折作为验证集、其余四折作为训练集,按照 2.2 节的参数设置对神经网络进行训练。在每折的神经

图2 CPSENET对随机厚度MgF₂、SiO₂、Si₃N₄和TiO₂薄膜的椭偏光谱的解调误差Fig. 2 Demodulation Error by CPSENET for MgF₂, SiO₂, Si₃N₄, and TiO₂ Films with Random Thicknesses

网络训练过程中,取神经网络在验证集上达到最小损失时的模型参数作为该折的训练结果,并计算该模型在当前验证集上的误差 d_{error} 、 n_{error} 、 ψ_{error} 和 Δ_{error} 。重复上述过程5次后,计算5次误差的均值 μ 、标准差 σ 以及变异系数(Coefficient of Variation, CV),以此量化数据集划分对CPSENET性能稳定性的影响(CV的计算方式如公式(6)所示)。

$$\text{CV} = \frac{\sigma}{\mu} \times 100\% \quad (6)$$

CPSENET的5折交叉验证结果如表2所示。其中, d_{error} 的均值为0.49 nm,标准差为0.0073 nm,CV为1.47%; n_{error} 的均值为0.0072,标准差为0.0002,CV为3.34%; ψ_{error} 的均值为3.27°,标准差为0.0573°,CV为1.73%; Δ_{error} 的均值为4.48°,标准差为0.0924°,CV为2.05%。四个评价指标的变异系数均小于5%,其中 d_{error} 的CV低至1.47%,显示出良好的稳定性。该结果表明,即使在数据集中选取不同的数据作为训练集和验证集,CPSENET的解调性能始终保持稳定,随机划分数据集的方式对模型性能的产生不会显著影响。

表2 CPSENET的5折交叉验证结果

Table 2 5-fold cross-validation results of CPSENET

	$d_{\text{error}}/\text{nm}$	n_{error}	$\psi_{\text{error}}/^{\circ}$	$\Delta_{\text{error}}/^{\circ}$
折1	0.52	0.0069	3.34	4.65
折2	0.49	0.0073	3.31	4.42
折3	0.50	0.0072	3.29	4.49
折4	0.49	0.0075	3.21	4.46
折5	0.47	0.0070	3.22	4.39
μ	0.49	0.0072	3.27	4.48
σ	0.0073	0.0002	0.0573	0.0924
CV	1.47%	3.34%	1.73%	2.05%

本文对MgF₂、SiO₂、Si₃N₄和TiO₂四种材料的薄膜,生成了厚度在[100, 200] nm(间隔1 nm)的椭偏光谱测试集,并添加大小在[-3, 3]之间的随机噪声,通过CPSENET对该测试集的解调结果反映其综合性能。以LM优化算法作为传统优化算法的代表作为对比实验,LM算法通过Python SciPy科学计算库中的标准优化求解器实现。表3展示了两种方法在测试集的直接参数误差-重构光谱误差评估统计平均结果。

在直接参数误差中,CPSENET对四种材料的平均厚度解调误差小于1.5 nm,平均折射率解调误差小于0.02,显著优于LM算法,对Si₃N₄和TiO₂的解调效果尤为突出。而LM算法的重构光谱误差略优于CPSENET,这是因为LM算法以最小化光谱拟合残差为优化目标,在重构光谱误差评估中自然具备优势。尽管如此,CPSENET的重构光谱误差仍与LM算法处于相近水平,表明CPSENET的解调结果与椭偏光谱

之间满足偏振光薄膜干涉理论,其参数反演的物理可靠性得到了验证。

在计算效率方面,LM算法对四种材料的求解总耗时分别为43.63 s、24.08 s、34.55 s、26.56 s,且求解过程严重依赖于初始值设置,稳定性较差。相比之下,CPSENET的单次解调时间稳定在6 ms左右,解调100条光谱总耗时仅为0.6 s,解调速度比LM算法提高了数十倍,展现出显著的高效性与稳定性优势。

表3 CPSENET及LM算法对多种厚度MgF₂、SiO₂、Si₃N₄和TiO₂薄膜椭圆偏光谱的平均解调误差统计数据及耗时对比
Table 3 Comparison of Average Demodulation Errors and Computational Time between CPSENET and LM Algorithms for MgF₂, SiO₂, Si₃N₄, and TiO₂ Thin Films with Various Thicknesses

	method	$d_{\text{error}}/\text{nm}$	n_{error}	$\psi_{\text{error}}/^{\circ}$	$\Delta_{\text{error}}/^{\circ}$	time/s
MgF ₂	CPSENET	0.78	0.0084	3.63	8.80	0.60
	LM	0.84	0.0175	2.71	6.55	43.63
SiO ₂	CPSENET	0.78	0.0053	3.61	6.05	0.60
	LM	1.65	0.0166	2.75	6.76	24.08
Si ₃ N ₄	CPSENET	0.49	0.0047	3.56	4.54	0.60
	LM	4.89	0.0719	2.70	5.08	34.55
TiO ₂	CPSENET	1.47	0.0195	3.89	7.39	0.60
	LM	5.89	0.0885	2.73	5.23	26.56

CPSENET和LM算法对测试集的解调结果详细误差如图3所示。在Si₃N₄和TiO₂两种材料的厚度解调误差中,当厚度大于150 nm时,LM算法的误差随厚度增加呈现显著上升趋势,CPSENET的厚度解调误差虽然也有所增大,但是整体表现出更优的稳定性。在折射率解调结果中,CPSENET对不同材料的解调误差表现均处于较低水平,而LM算法则对高折射率表现出明显的敏感性,在Si₃N₄和TiO₂材料上的解调误差明显增大。整个波段的解调结果中,CPSENET在380 nm - 450 nm的短波范围的折射率解调误差明显上升,对应的重构光谱误差 Δ_{error} 也显著增大。该现象主要源于柯西色散模型在紫外波段的固有局限性:随着波长减小,高阶参数B和C对折射率的影响急剧增强,导致微小的高阶项偏差也会对短波区域及重构光谱误差造成较大影响。

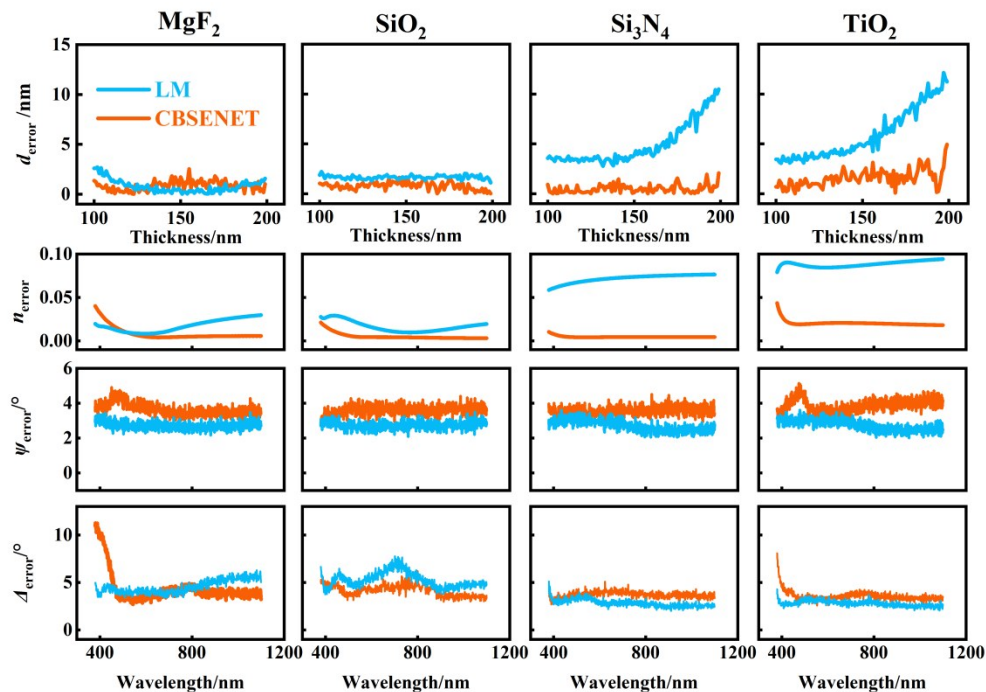


图3 CPSENET对多种厚度MgF₂、SiO₂、Si₃N₄和TiO₂薄膜的椭圆偏光谱的解调误差统计结果
Fig. 3 Demodulation Error Statistics of CPSENET for Spectroscopic Ellipsometry Data of MgF₂, SiO₂, Si₃N₄, and TiO₂ Thin Films with Various Thicknesses

4 讨论

4.1 加权损失函数权重优化

本文在网络训练过程中,使用WMSE以平衡柯西模型中各阶参数数量级之间的差异,在此采用逐参数寻优的方法进一步确定权重的最优取值。以 W_d 为基准,固定其为1,设置柯西参数的初始权重为 $W_A = 1$, $W_B = 1 \times 10^3$, $W_C = 1 \times 10^9$,对柯西参数权重进行依次优化并将其更新,再对下一个参数继续寻优,直到所有参数完成寻优。为评估所得权重配置的优劣,按照训练集的生成方式构建了椭偏光谱测试集,其中 d 、 A 和 B 的采样数量分别设置为10、10和5,并以CPSENET在此测试集的解调误差作为评判指标。

不同参数权重训练的网络在测试集的解调误差如图4所示。在 W_A 从1增加至200的过程中, n_{error} 、 ψ_{error} 和 Δ_{error} 持续下降,当 $W_A > 100$ 后趋于稳定, d_{error} 呈现先下降后上升的趋势,并在 $W_A = 100$ 时取得最小值。综合各误差的变化趋势,确定 $W_A = 100$ 为参数 A 的最优权重。

在 W_B 从 0.1×10^{-3} 增加至 10×10^{-3} 的过程中,各项误差均呈现出先下降后上升的趋势,并在 $W_B = 1 \times 10^{-3}$ 时同时达到最小值。因此,确定 $W_B = 1 \times 10^{-3}$ 为参数 B 的最优权重。

在 W_C 从 0.01×10^{-9} 增加至 100×10^{-9} 的过程中,当 W_C 小于 10×10^{-9} 时, d_{error} 的变化较小, n_{error} 、 ψ_{error} 和 Δ_{error} 呈现先下降后上升的趋势,且在 W_C 为 5×10^{-9} 和 10×10^{-9} 时处于最低水平。当 W_C 大于 10×10^{-9} 时,各项误差均显著增大。由此可见, W_C 的最优取值在 $[5 \times 10^{-9}, 10 \times 10^{-9}]$ 之间。在柯西模型中参数 C 主要用于折射率微调,且对最终解调结果影响有限,因此可以认为 W_C 在 $[5 \times 10^{-9}, 10 \times 10^{-9}]$ 范围内取值均比较合理。

综上,通过逐参数权重优化,确定损失函数的最优权重配置为: $W_d = 1$, $W_A = 100$, $W_B = 1 \times 10^{-3}$, W_C 的取值范围为 $[5 \times 10^{-9}, 10 \times 10^{-9}]$ 。经过权重优化后,损失函数中各项参数的贡献度与其重要性实现了良好匹配。

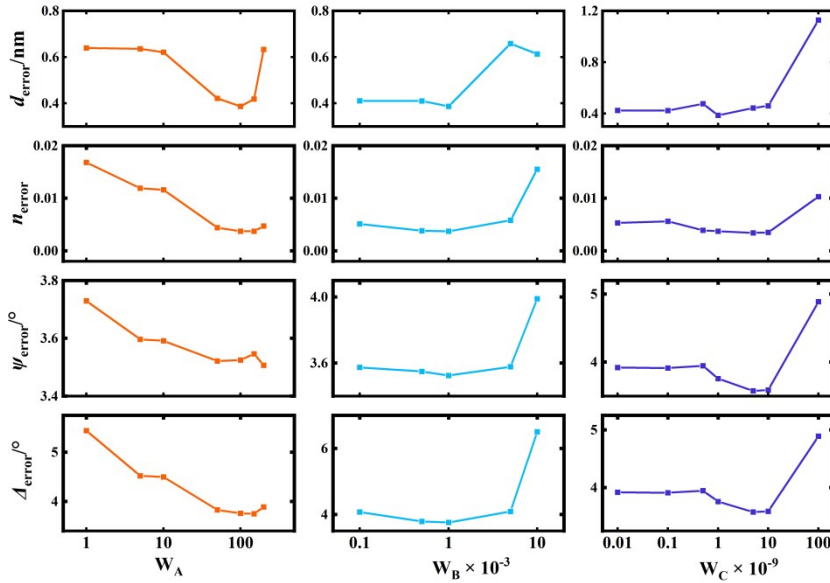


图4 不同损失函数权重配置下CPSENET的解调误差

Fig. 4 Demodulation Error of CPSENET under Different Loss Function Weight Configurations

4.2 定向数据增强

本文提出的基于柯西模型的参数化数据生成方法具备显著的定向数据增强优势:针对不同材料的光学特性及具体任务需求,通过调整柯西模型的参数范围,有针对性地生成训练数据,从而实现网络性能的定向提升与优化。

本文分别对 SiO_2 和 MgF_2 这两种低色散、低折射率材料,以及 Si_3N_4 和 TiO_2 这两种高色散、高折射率材料,实施了定向数据增强,并在此基础上重新训练CPSENET,得到两个专用网络模型:CPSENET-A(针对低色散材料)与CPSENET-B(针对高色散材料)。在构建CPSENET-A的训练集时,根据 SiO_2 和 MgF_2 的折

射率特性,设置 A 的参数范围为 $[1.3, 1.5]$, B 的参数范围为 $[1 \times 10^3, 6 \times 10^3] \text{ nm}^2$; 在构建 CPSENET-B 的训练集时,根据 Si_3N_4 和 TiO_2 的折射率特性,设置 A 的参数范围为 $[1.8, 2.2]$, B 的参数范围为 $[10 \times 10^3, 25 \times 10^3] \text{ nm}^2$ 。四种材料的厚度范围设置为 $[90, 210] \text{ nm}$,以增强网络对厚度边界区域的解调稳定性。在网络训练过程中,WMSE 的权重采用优化后的参数,其余数据集生成参数、网络结构及网络训练参数均与前文保持一致。

CPSENET-A 和 CPSENET-B 在测试集上的直接参数误差和重构光谱误差如图 5 所示。相比于 CPSENET,两专用模型的直接参数误差与重构光谱误差均显著降低。具体而言:CPSENET-A 对 SiO_2 和 MgF_2 的 d_{error} 大幅度降低,表现出更好的厚度解调稳定性,并且短波段处的 n_{error} 和 Δ_{error} 被有效改善;CPSENET-B 主要提升了对 TiO_2 薄膜的解调性能, d_{error} 和全波段的 n_{error} 明显降低。通过数据增强后的数据集训练出的专用网络,在对应的材料上表现出更好的解调性能。

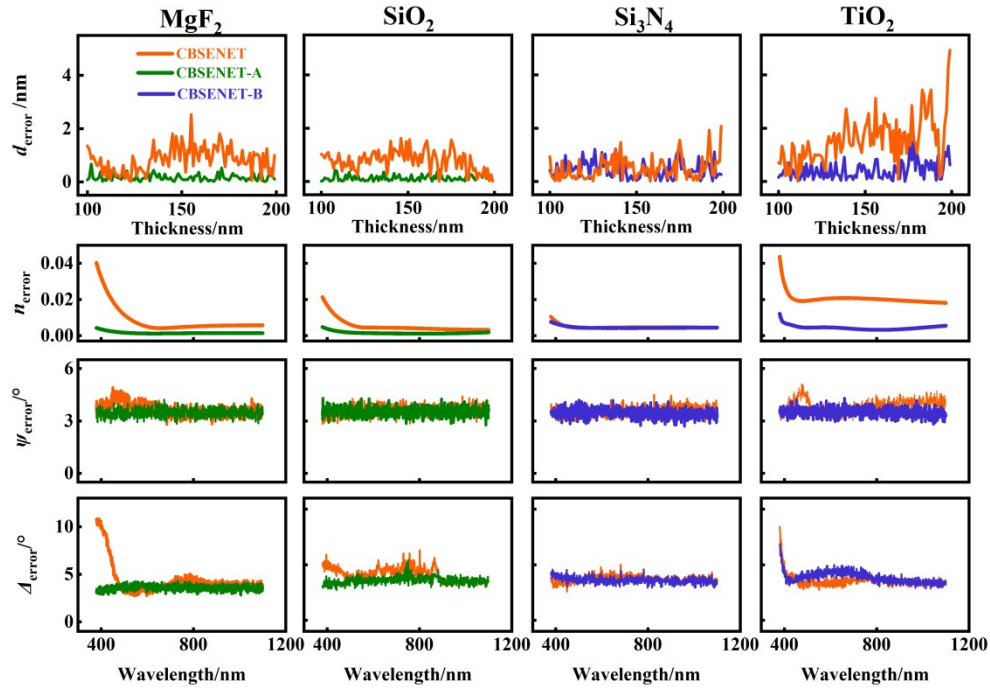


图 5 CPSENET-A 和 CPSENET-B 分别在低折射率及低色散材料和高折射率及高色散材料上的解调误差
Fig. 5 Demodulation Errors for CPSENET-A (on Low Refractive Index and Dispersion) and CPSENET-B (on High Refractive Index and Dispersion) Materials

针对不同厚度范围的薄膜,调整数据集生成参数中 d 的取值范围,同样可以实现定向数据增强。将 d 范围扩展至 $[50, 300] \text{ nm}$,并将其划分为三个子区间,分别为 $[50, 100] \text{ nm}$, $[100, 200] \text{ nm}$ 和 $[200, 300] \text{ nm}$,对扩展及划分的厚度范围重新生成了椭圆偏振光谱数据,并分别对 CPSENET 进行训练。在数据生成和网络训练过程中,WMSE 的权重采用优化后的参数,其余设置均与前文保持一致。为验证对薄膜厚度进行定向数据增强的有效性,对 MgF_2 、 SiO_2 、 Si_3N_4 和 TiO_2 四种材料的薄膜,生成了厚度在 $[50, 300] \text{ nm}$ (间隔 1 nm) 的椭圆偏振光谱测试集,并添加大小在 $[-3, 3]$ 之间的随机噪声。随后,分别利用不同厚度区间对应的数据集训练得到 CPSENET-1 ($[50, 300] \text{ nm}$)、CPSENET-2 ($[50, 100] \text{ nm}$)、CPSENET-3 ($[100, 200] \text{ nm}$) 和 CPSENET-4 ($[200, 300] \text{ nm}$),使用训练出的网络对测试集中对应厚度范围的椭圆偏振光谱进行解调,通过解调误差来评估模型的解调性能。

CPSENET-1、CPSENET-2、CPSENET-3 和 CPSENET-4 的解调误差如图 6 所示。在全厚度范围 $[50, 300] \text{ nm}$ 的解调任务中,CPSENET-1 对 MgF_2 和 SiO_2 在 $[50, 100] \text{ nm}$ 范围内的 d_{error} 较大,当 d 较小时尤为明显。其原因在于,这两种材料的 n 较小,当 d 也较小时, d 变化所引起的椭圆偏振光谱变化非常小,模型难以分辨不同厚度薄膜之间的特征差别,从而导致 d_{error} 增大。此外,CPSENET-1 对四种材料在 $[200, 300] \text{ nm}$ 范围内的 d_{error} 均较大。这是因为当 d 较大时,光谱细节起伏更为剧烈,特征更加精细,光谱复杂度增加,对网

络的特征提取能力提出了更高要求,进而导致 d_{error} 上升。相比之下,经过厚度数据增强之后的 CPSENET-2、CPSENET-3 和 CPSENET-4 分别对 $[50, 100]$ nm、 $[100, 200]$ nm 和 $[200, 300]$ nm 三个子区间的椭偏光谱进行解调,其 d_{error} 均显著降低,并展现出更高的稳定性。随着厚度增大,三个模型的 d_{error} 整体呈现上升趋势,但其与理论厚度的相对平均误差分别为 0.41%、0.26% 和 0.36%,仍保持稳定且处于较低水平。与 CPSENET-1 相比,CPSENET-2、CPSENET-3 和 CPSENET-4 不仅在 d_{error} 上有所下降,其 n_{error} 也同步降低。这表明,通过对薄膜厚度进行定向数据增强,网络多方面的性能得到了有效优化。

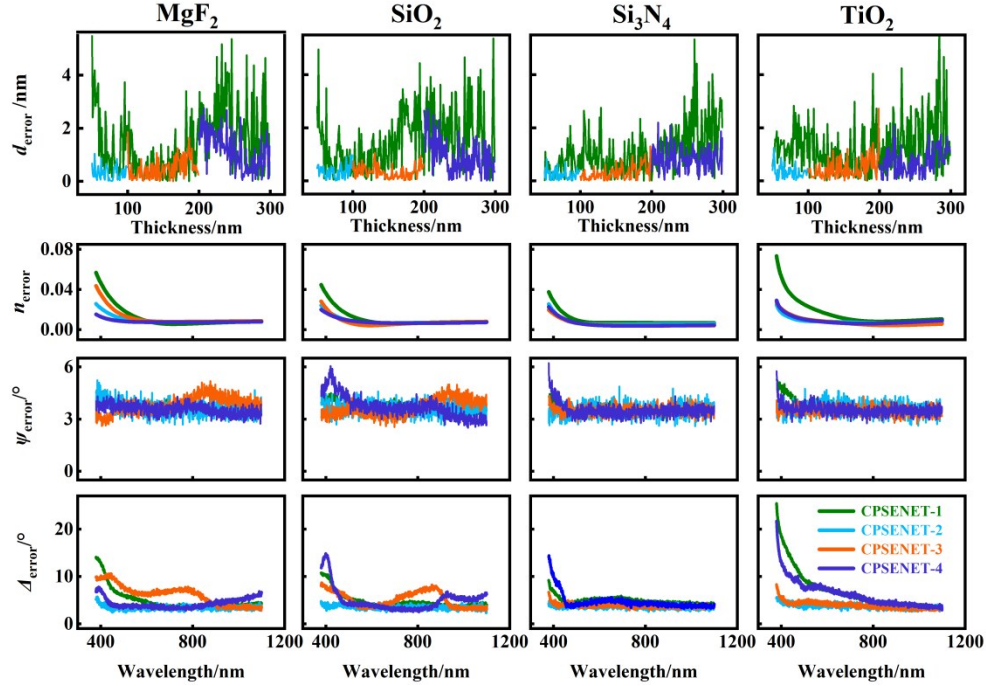


图6 定向厚度增强后的 CPSENET 分别对不同厚度四种材料薄膜的椭偏光谱的解调结果

Fig. 6 Demodulation results of ellipsometric spectra for four material films with different thicknesses using thickness-directed enhanced CPSENET

4.3 材料吸收特性对网络性能的影响

本文的方法主要面向无吸收薄膜的光谱椭偏数据解调所设计,然而在实际应用中,材料会存在不同程度的吸收特性,即消光系数 k 不等于 0。因此,本节将讨论材料吸收特性对 CPSENET 解调性能的影响。不同吸收程度的材料选择的是电子跃迁能量分别为 2.74 eV、2.60 eV、2.48 eV、2.33 eV 的 $\text{Si}_{1-x}\text{N}_x$ ^[18],它们在 380 nm 波长处的吸收系数分别为 0.019、0.052、0.087、0.163,并分别在 440 nm、460 nm、480 nm、500 nm 处下降至 0。对这四种材料根据公式(1)生成多种厚度的椭偏光谱各 100 条,并添加大小在 $[-3, 3]$ 之间的随机噪声,通过 CPSENET 对这些椭偏光谱上的解调误差反映材料吸收特性对网络性能的影响。

CPSENET 对这些椭偏光谱数据的解调结果如图 7 所示。 d_{error} 和 n_{error} 随着吸收系数的增大而增大,原因在于:训练集中不包含有吸收材料的椭偏光谱,CPSENET 没有学习到处理材料吸收对解调结果影响的能力,材料吸收特性的增大等同于引入了噪声,因此,CPSENET 对有吸收材料的解调性能会下降。此外, n_{error} 在短波段的误差增加非常剧烈,其原因除了在第 4 节中说明的柯西模型在短波段的固有局限性以外,还由于 K-K 关系的约束^[19],柯西模型不能准确描述材料在有吸收波段的折射率。随着材料吸收特性的增加, ψ_{error} 和 Δ_{error} 也明显增加,CPSENET 的解调结果与椭偏光谱之间的偏振光薄膜干涉理论自洽性恶化,从侧面反映了解调结果可信度的下降。

随着 $\text{Si}_{1-x}\text{N}_x$ 材料消光系数的增大,CPSENET 的解调误差变化趋势如图 8 所示。解调误差随材料吸收特性的增强呈现非线性变化,在吸收超过阈值后急剧增加:当材料最大消光系数小于 0.05 时,CPSENET 的直接参数误差与重构光谱误差均未出现明显上升趋势,且误差水平与无吸收材料相当;而当最大消光系数超过 0.05 时,各项误差迅速增大。此结果表明:当材料的最大消光系数小于 0.05 时,CPSENET 仍具有较好

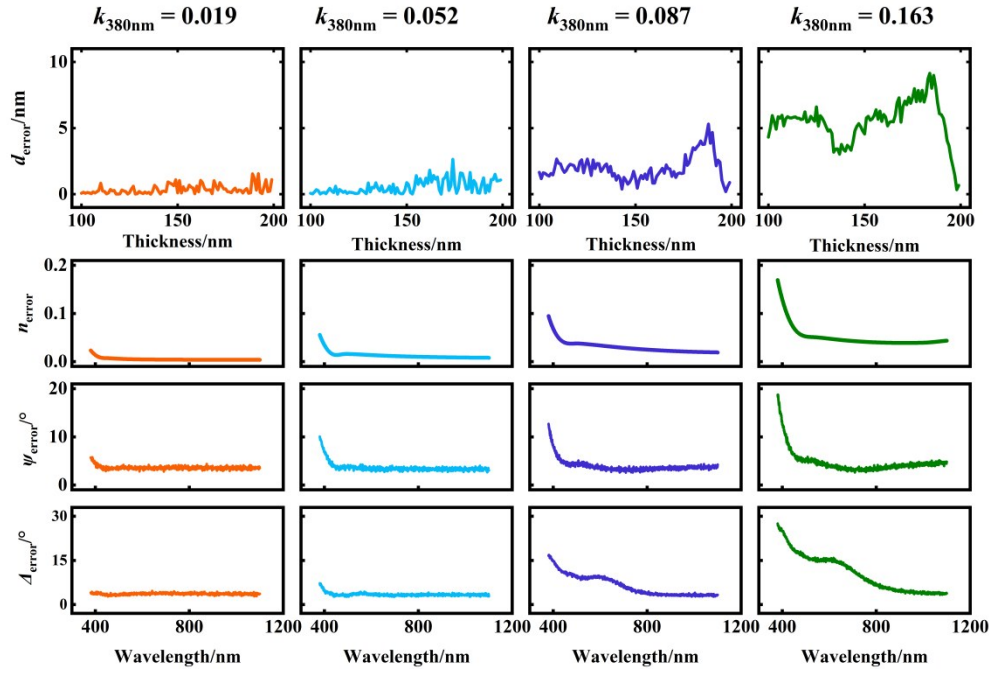


图7 CPSENET对不同消光系数的 $\text{Si}_{1-x}\text{N}_x$ 薄膜椭偏光谱的解调误差

Fig. 7 Demodulation errors of CPSENET for ellipsometric spectra of $\text{Si}_{1-x}\text{N}_x$ thin films with different extinction coefficients

的解调性能;当消光系数超过该阈值,CPSENET的解调性能将随着吸收特性增加而急剧退化,解调结果可信度降低。

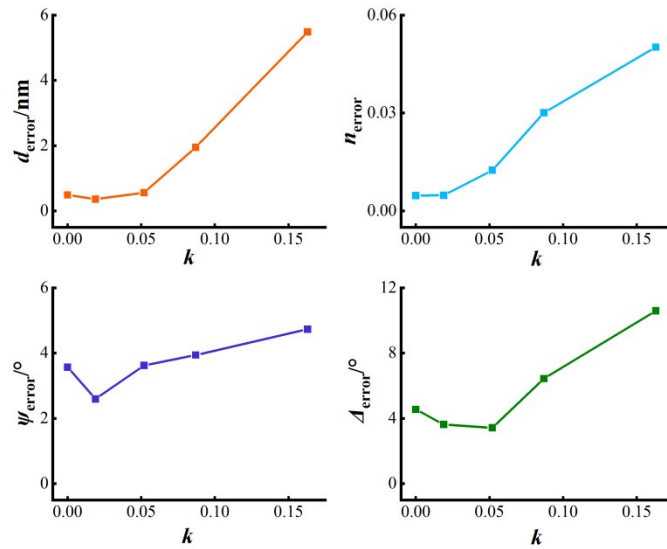


图8 消光系数对CPSENET解调 $\text{Si}_{1-x}\text{N}_x$ 薄膜椭偏光谱结果的影响

Fig. 8 Impact of Extinction Coefficient Variation on CPSENET Performance for Ellipsometric Spectra of $\text{Si}_{1-x}\text{N}_x$ Thin Films

5 总结

本文提出了一种基于柯西色散模型的光谱椭偏数据解调神经网络方法CPSENNM。该方法首先利用柯西色散模型生成折射率数据,无需依赖折射率数据库,再根据偏振薄膜干涉公式生成椭偏光谱数据集,据此训练出神经网络CPSENET。CPSENET建立椭偏光谱到薄膜厚度 d 和柯西模型参数 A 、 B 和 C 的映射,再根据柯西模型得到折射率信息。对四种典型材料 MgF_2 、 SiO_2 、 Si_3N_4 和 TiO_2 的椭偏光谱解调结果表明:CPSENET具有优于LM算法的解调误差,并且解调耗时远小于LM算法。同时,CPSENNM能根据材料的折射率特性以及薄膜厚度定向增强训练数据,从而提升网络对特定折射率和厚度范围薄膜材料椭偏光谱的

解调性能。CPSENNM具有较好的解调性能以及灵活性,为无吸收薄膜椭偏光谱的快速解调提供了一种可行的方法。

参考文献

- [1] LI Guobin, HU Kun, ZHANG Taiwei, et al. Study on the preparation and spectroscopic ellipsometry of MoTe₂ quantum dot films[J]. Journal of Infrared and Millimeter waves, 2025, 44(02): 153-160.
李国彬, 胡坤, 张泰玮, 等. MoTe₂量子点薄膜的制备及椭圆偏振光谱研究[J]. 红外与毫米波学报, 2025, 44(02): 153-160.
- [2] HU Yi, LI Zijian, LIU Lei. Temperature-induced work function changes in NiMn₂O₄ thin films by spectroscopic ellipsometry[J]. Journal of Materials Science and Engineering, 2025, 49(01): 486-492.
胡轶, 李子建, 刘磊. 基于椭圆偏振光谱线法研究温度对NiMn₂O₄薄膜光电性能的影响[J]. 材料科学与工程学报, 2025, 43(03): 486-492.
- [3] SUN Wei, JIN Shangzhong, ZHANG Yin, et al. Design of film thickness and curvature measurement system based on ellipsometry spectrum and optical lever method[J]. Optical Communication Technology, 2025, 49(01): 94-100.
孙伟, 金尚忠, 张殷, 等. 基于椭偏光谱和光杠杆法的薄膜厚度及曲率测量系统设计[J]. 光通信技术, 2025, 49(01): 94-100.
- [4] SEBAK M A, GHALAB S, EL-TAHER A, et al. Studying the structural, optical spectroscopic ellipsometry and electrical properties of variable-CdS thickness/CdTe for solar cell applications[J]. Chalcogenide Letters, 2022, 19(6): 389-408.
- [5] ORJI N G, BADAROGLU M, BARNES B M, et al. Metrology for the next generation of semiconductor devices[J]. Nature Electronics, 2018, 1(10): 532-547.
- [6] AZZAM R M A, BASHARA N M, BALLARD S S. Ellipsometry and polarized light[M]. Amsterdam: North-Holland Publishing Company, 1978.
- [7] TOMPKINS H, IRENE E A. Handbook of ellipsometry[M]. Norwich: William Andrew, 2005.
- [8] FUJIWARA H. Spectroscopic ellipsometry: principles and applications[M]. Chichester: John Wiley & Sons, 2007.
- [9] ZAGHLOUL A R M, AZZAM R M A, BASHARA N M. Inversion of the nonlinear equations of reflection ellipsometry on film-substrate systems[J]. Surface Science, 1976, 56: 87-96.
- [10] JELLISON JR G E. Spectroscopic ellipsometry data analysis: measured versus calculated quantities[J]. Thin solid films, 1998, 313: 33-39.
- [11] MOORE E A, SMART L E. Optical properties of solids[M]. Oxford: Oxford university press, 2020.
- [12] JELLISON JR G E, MODINE F A. Parameterization of the optical functions of amorphous materials in the interband region[J]. Applied Physics Letters, 1996, 69(3): 371-373.
- [13] LIU J, ZHANG D, YU D, et al. Machine learning powered ellipsometry[J]. Light: Science & Applications, 2021, 10(1): 55.
- [14] ZHU P, ZHANG D, NIU X, et al. A lightweight neural network for spectroscopic ellipsometry analysis[J]. Advanced Optical Materials, 2024, 12(4): 2301381.
- [15] URBAN F K, BARTON D. Numerical ellipsometry: Artificial intelligence for rapid analysis of indium tin oxide films on silicon[J]. Journal of Vacuum Science & Technology A, 2024, 42(4): 043401.
- [16] URBAN F K, BARTON D. Numerical ellipsometry: Artificial intelligence for real-time, in situ absorbing film process control[J]. Journal of Vacuum Science & Technology A, 2024, 42(2): 023404.
- [17] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
- [18] FUJIWARA H, COLLINS R W. Spectroscopic Ellipsometry for Photovoltaics: Volume 2: Applications and Optical Data of Solar Cell Materials[M]. Springer, 2019.
- [19] ADACHI S. Optical properties of crystalline and amorphous semiconductors: Materials and fundamental principles[M]. Springer Science & Business Media, 2012.

Rapid Demodulation of Ellipsometric Spectra of Non-Absorbing Thin Films: a Cauchy Model-Powered Neural Network Method

LIU Chi, LIU Jihong*, BU Zichuan, LI Jiaqi, ZHANG Jiageng, JING Jiahao, REN Kaili,
YAN Xuewen, Dong Jun

(School of Electronic Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China)

Abstract: Valued for its nondestructive, non-contact operation and high precision, spectroscopic ellipsometry has established itself as the standard characterization method for measuring the thickness and optical constants of nanoscale thin films and is routinely employed in the quality inspection of coatings. The conventional spectroscopic ellipsometry demodulation approach is model-based. It involves constructing a pre-defined optical model for the thin film, complete with a dielectric function model, and then iteratively refining the model parameters via nonlinear regression analysis to extract the properties of the film. However, the complexity of this modeling and the iterative solving process result in low computational efficiency, making it difficult to meet the demands for rapid inspection of coated samples. Neural networks can establish a direct mapping from ellipsometric spectra to thin-film properties, thereby enabling rapid data demodulation. However, the application of neural networks in spectroscopic ellipsometry data demodulation is constrained by the challenges of obtaining sufficient training data and ensuring its high quality. To address those challenges, this paper proposes a neural network-based demodulation method for spectroscopic ellipsometry data of non-absorbing thin films, incorporating the Cauchy dispersion model as a physical constraint. In this method, a dataset of ellipsometric spectra was first generated based on the Cauchy dispersion model and the theory of polarized light interference in thin films. Subsequently, a neural network was trained using this dataset, with the film thickness and the parameters of the Cauchy model designated as its output targets. The dataset generation procedure was conducted as follows. First, the value ranges for the Cauchy model parameters (A , B , and C) were established based on the material's attainable refractive index range and its spectral variation. Within these ranges, parameters A and B were sampled at equal intervals, and C was randomly sampled. All possible combinations of (A , B and C) were then traversed to generate the corresponding refractive index data via the Cauchy dispersion model. Subsequently, the film thickness d was sampled at equal intervals within its predefined range. Each sampled thickness value was paired with every generated refractive index dataset. Following this, the ellipsometric spectra for each (d , n) pair were simulated according to the principles of polarized light interference in thin films. Finally, experimental noise was added to these ideal spectra to produce the final ellipsometric spectra dataset for network training. In this work, the ResNet-18 architecture was selected as the backbone network. It was specifically configured to be one-dimensional to accommodate the sequential nature of spectroscopic ellipsometry data. The two measured ellipsometric parameters are used as the input of network. The outputs of the network consist of the film thickness d and the three Cauchy model parameters (A , B and C). The wavelength-dependent refractive index is then reconstructed from these parameters using the Cauchy dispersion model. This design, as opposed to having the network directly output the refractive index at multiple discrete wavelengths, effectively reduces the complexity of model, and improves its learning efficiency. Experimental validation on four distinct materials (MgF_2 , SiO_2 , Si_3N_4 , and TiO_2) demonstrates that the proposed network achieves an average thickness demodulation error of less than 1.5 nm, an average refractive index error below 0.02, and an average computation time of approximately 6 ms. It significantly outperforms the conventional Levenberg-Marquardt (LM) algorithm in both accuracy and efficiency, demonstrating substantial advantages in both performance and stability for rapid characterization. The proposed demodulation framework also demonstrates the advantage of enabling targeted data augmentation. Specifically, tailored data augmentation strategies were implemented for two distinct material categories: low-dispersion, low-refractive-index materials and high-dispersion, high-refractive-index materials. This process yielded two specialized network models. Compared to the original general-purpose model, these specialized models demonstrate superior demodulation performance on their respective target material types. Finally, an analysis was performed to investigate the impact of material absorption on the demodulation accuracy and to determine the corresponding tolerance threshold. The results indicate that the absorption characteristic of material generally degrades the demodulation precision. Nevertheless, the proposed method maintains satisfactory accuracy when the extinction coefficient k is below 0.05, establishing a practical applicability boundary for absorption-free or weakly absorbing thin films. In conclusion, the proposed neural network framework addresses the two challenges of training data acquisition and fast demodulation in spectroscopic ellipsometry, offering a novel and efficient pathway for thin-film characterization.

Key words: Spectroscopic Ellipsometry; Neural Networks; Cauchy Dispersion Model; Dataset; Non-absorbing Thin Films

OCIS Codes: 300.6170; 130.2130; 240.0310; 240.2130

CSTR: 32255.14.gzxb20265507.0730002

Foundation item: National Natural Science Foundation of China (12274341), Key Research and Development Projects in Shaanxi Province (2024QY2-GJHX-39), Key Research and Development Program of Shaanxi Provincial Department of Science and Technology(2025CY-YBXM-114), Key Research and Development Program of Shaanxi Provincial Department of Science and Technology(2025CY-YBXM-077).