

〈综述与评论〉

基于卷积神经网络的红外光谱建模分析综述

王 堃, 史 勇, 刘池池, 谢 义, 蔡 萍, 孔松涛
(重庆科技学院 机械与动力工程学院, 重庆 401331)

摘要: 红外光谱技术存在着数据预处理复杂、预测精度不高, 且难以处理大量非线性数据的问题, 适用于卷积神经网络进行处理。本文首先分析了卷积神经网络应用在红外光谱上的优点, 并对卷积神经网络结构组成进行简单的概述。然后针对卷积神经网络在光谱分析建模中的输入数据维度问题进行详细阐述; 针对模型设计中卷积核参数的影响、多任务处理模型以及训练过程中的优化方法进行综述。最后分析了该研究的优点与不足, 并展望了未来的发展趋势。

关键词: 红外光谱; 卷积神经网络; 维度; 建模

中图分类号: O657.33

文献标识码: A

文章编号: 1001-8891(2021)08-0757-09

A Review of Infrared Spectrum Modeling Based on Convolutional Neural Networks

WANG Kun, SHI Yong, LIU Chichi, XIE Yi, CAI Ping, KONG Songtao

(Chongqing University of Science & Technology, College of Mechanical and Power Engineering, Chongqing 401331, China)

Abstract: Convolutional neural networks are used to solve problems such as complex data preprocessing, low prediction accuracy, and difficulty in dealing with a large amount of nonlinear data in infrared spectroscopy. Moreover, owing to their strong feature extraction ability and good nonlinear expression ability, the application of convolutional neural networks in the modeling of infrared spectrum analysis has attracted attention. In this study, the advantages of the application of a convolutional neural network for the infrared spectrum are analyzed, and the structure and composition of the convolutional neural network are briefly summarized. Then, the dimension problem of the input data in the spectral analysis modeling of the convolutional neural network is described in detail. This paper reviews the influence of convolution kernel parameters in the model design, multi-task processing model, and optimization methods in the training process. Finally, the advantages and disadvantages of this research are analyzed, and future development trends are discussed.

Key words: infrared spectroscopy, convolutional neural network, dimension, modeling

0 引言

随着化学计量学和仪器测量技术的发展, 红外光谱学已经可以同时分析多个参数, 仪器设备也在向小型化升级, 从而实现合并到单一平台分析数据^[1-5]。光谱学如今对化学领域中的质量控制与产品分析具有重要作用。在石油化工、农业、医药、烟草、食品、环境监测、航空航天等领域得到了广泛的应用^[6]。例如天津港东科技公司利用红外光谱法检测新冠肺炎防护口罩的质量, 对口罩产品进行辅助评价^[7]。然而, 由于光谱信息变量与要处理的各种变量之间存在共线性

关系, 因此, 测定产品中特定的化学成分的浓度是很困难的^[8]。

红外光谱学在数据分析过程中主要包括数据预处理与建立模型两个步骤。光谱预处理主要有基线校正、散射校正、平滑和归一化等 4 个步骤, 每个步骤又有若干种方法^[9]。一方面通过试错法选择预处理方法会增加复杂度。另一方面, 采集光谱数据不稳定, 影响因素较多, 环境、产源、设备等因素发生变化时, 光谱信号的噪声也会发生改变, 以至于预处理方法产生不适应性, 导致模型预测能力的下降。在对光谱数据建立模型时所用的各种方法, 如支持向量机^[10], 人工

收稿日期: 2020-08-16; 修订日期: 2020-10-26.

作者简介: 王堃 (1980-), 男, 博士研究生, 主要研究方向为传热反问题, E-mail: 3938630@qq.com.

通信作者: 孔松涛 (1969-), 男, 四川人, 教授, 研究生导师, 博士。研究方向: 流体流动与传热、工业大数据分析及钻井与石油装备, E-mail: kst@tom.com.

神经网络^[11], 随机森林^[12], 立体派回归树, 偏最小二乘回归^[13]等, 都是在探索处理大量输入变量的能力。

以上模型的性能取决于光谱预处理技术的使用。如何选择一种良好的数据预处理方法仍是主要关注的问题^[14]。而机器学习学界则提出一个更为直接的解决方案, 即对数据进行学习, 从而识别并挖掘数据中隐藏的潜在信息^[15]。随着对机器学习不断地研究与发展, 它相比传统的方法具有以下优势: ①具有非线性建模能力; ②具有良好的泛化能力; ③对预处理的要求较少。

卷积神经网络 (convolutional neural networks, CNN) 是近年来备受关注的深度学习的一个子集, 在计算机视觉领域有着重要的作用^[16]。在光谱学的应用中, 卷积神经网络作为一种端到端的模型, 可以采取较少的数据预处理甚至不进行数据预处理, 直接对光谱信号进行采集、训练、验证、测试, 可避免需要掌握过多的专业知识, 降低了操作难度。例如卷积核函数不仅可以提取近红外光谱的抽象特征, 还可以选择特征最明显的波长, 并消除没有信息的波长。Chen Yuanyuan^[17]等人研究发现直接将采集到的全部原始光谱信息作为输入, 不对红外光谱进行波长选择时, 建立一个基于卷积神经网络的近红外光谱端到端的定量分析模型, 并以玉米近红外数据集为实验, 验证了该方法的有效性。且广义性能更优于基于波长选择的传统方法。Ni C.^[18]等人也建立了一个一维变权卷积神经网络, 可以自动将网络注意力集中到重要的光谱波长上, 对马尾松幼苗叶片氮含量进行了预测, 结果证明模型的泛化能力得到了提高。

卷积神经网络的应用使红外光谱技术得到了提升, 本文针对卷积神经网络在红外光谱技术上的应用方法以及影响因素进行了综述, 并详细分析了该方法的优点和局限性, 以及未来的发展趋势。

1 卷积神经网络结构

Le Cun^[19]等人在 1989 年首先提出了卷积神经网络的概念, 并在 1998 年对其实现了改进, 该架构被命名为 LeNet-5, 其网络模型如图 1 所示。一个典型的 CNN 模型有一个输入层、几个隐藏层 (卷积层、池化和全连接) 和一个输出层。随后 Hinton 和 Alex Krizhevsky 在 2012 年设计出了 AlexNet^[20], 首次在 CNN 中成功应用了 ReLU、Dropout 和 LRN 等 Trick, 大大提高了 CNN 的性能。卷积神经网络是一个多层非全连接的神经网络, 在正向传播过程中利用卷积层和池化层相互交替学习, 提取原始光谱数据的特征; 反向传播过程中利用梯度下降算法最小化误差函数来

调整参数, 完成权值更新^[21]。

1.1 输入层

输入层用于接收待处理的多维原始光谱数据。针对红外光谱, 输入层通常需要指定宽度、高度通道。如果输入的数据是光谱图, 那么通道的数量通常设置为 3 个, 以说明通道的颜色 (红色、绿色和蓝色)。如果输入数据是一维光谱信号, 则通道数设置为 1。

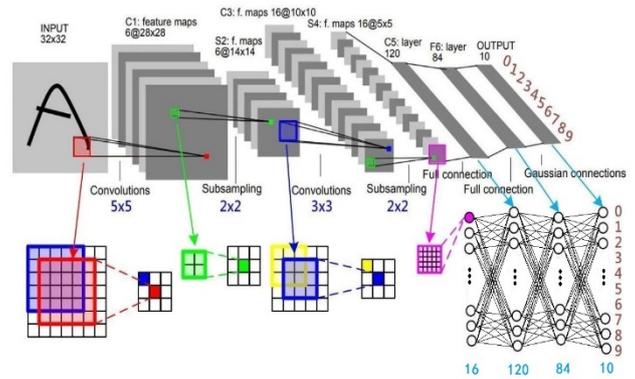


图 1 LeNet-5 网络模型示意图^[19]

Fig.1 Schematic diagram of LeNet-5 network model

1.2 卷积层

在卷积层中, 对输入的光谱信息进行特征提取, 由一个在输入端滑动的卷积核组成。在公式上, 卷积层可以表示为:

$$y^j = f(b^j + \sum_i k^{ij} * x^i) \quad (1)$$

式中: x^i 和 y^j 分别是第 i 个输入映射和第 j 个输出映射; k^{ij} 是特征图 i 和 j 之间的卷积核; $*$ 表示卷积; b^j 是第 j 个特征图的偏置参数。一个单独的卷积层能够识别简单的特征, 当增加更多的层时, 网络能够提取更加复杂和抽象的特征^[22]。对光谱信号做卷积运算时, 需考虑光谱信号的维度与卷积核的维度的适应性, 具体方法将在第 3 章中介绍。

1.3 激活函数

激活函数可以在复杂的学习模式中做出决策, 使学习过程更快。卷积特征图的激活函数定义式:

$$T_i^k = f_A(F_i^k) \quad (2)$$

式中: F_i^k 是卷积运算的输出, 传递给激活层; $f_A()$ 会通过非线性函数运算返回第 k 层输出 T_i^k 。通常使用的几个非线性组合激活函数, 如 sigmoid, tanh, maxout, ReLU 和 ReLU 的变体如 leaky ReLU、ELU 和 PReLU^[23-27]。光谱信号包含非常多的非线性数据, 因此激活函数的存在可以使光谱数据有更好的非线性表达。

1.4 池化层

池化层用于降低特征图的维数, 同时保留输入的

特征信息。它通过减少参数数量和防止过拟合来降低计算成本,从而提高了网络的通用性和准确性。池化层的计算过程可表示为:

$$x_{\beta}^{\gamma} = \text{pool}(x_{\beta}^{\gamma-1}) \quad (3)$$

式中: $x_{\beta}^{\gamma-1}$ 和 x_{β}^{γ} 分别是第 $\gamma-1$ 层、第 γ 层输出特征上相应的值; pool 为所选择的池化方式,包括最大值池化、平均池化、L2、重叠和空间金字塔合并等^[28-29]。

1.5 批次归一化

批次归一化用于解决与特征图中内部协方差平移有关的问题。公式(4)表示对变换后的特征图 T_l^k 的批次归一化。

$$N_l^k = \frac{F_l^k - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}} \quad (4)$$

式中: N_l^k 表示归一化特征图; F_l^k 是输入特征图; μ_B 和 σ_B^2 分别表示小批次特征图的均值和方差。批次归一化通过将特征图值设为零均值和单位方差来统一其分布^[30]。

1.6 Dropout

Dropout 在网络中引入了正则化,通过随机跳过某些单元或以一定概率跳过某些连接,从而提高网络的泛化性^[31]。某些单元或连接的随机丢弃会产生若干稀疏的网络结构,最后以较小的权值选出一个具有代表性的网络,视其为所有拟议网络的近似值^[32]。

1.7 全连接层

CNN 在经过卷积与池化过后,与全连接层相连接。全连接层通常用于网络末端的分类任务^[33]。与卷积与池化不同,它从前一层获取输入后,从全局分析特征^[34]。然后将特征进行非线性组合,进行数据分类^[35]。全连接层如图 2 所示。

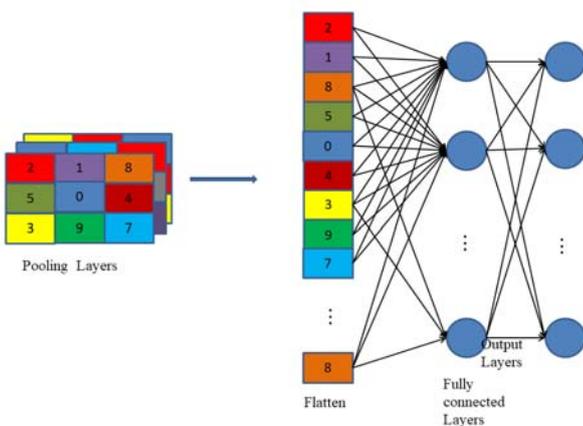


图 2 全连接层示意图

Fig.2 Diagram of full connection layer

2 维度转换

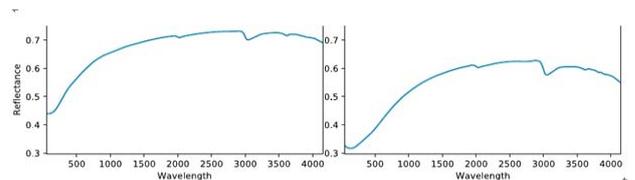
现在,世界上有很多流行的免费开源深度学习框架,例如 Caffe、Tensorflow、Torch 和 MatConv。默认情况下,这些框架的输入信号是二维(2D)或三维(3D)图像,在对图像的处理上有很强的能力。光谱信号是一维信号图,不能直接使用上述模型。所以为了更好地利用 CNN 模型,需要对数据或模型进行维度转换。

2.1 光谱信号转变为光谱图

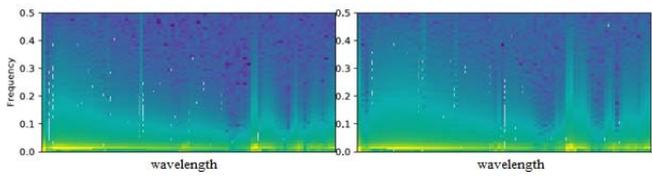
由于 CNN 模型适用于二维信号输入,将光谱转变为光谱图的二维表示,从而实现 CNN 的适用。早在 1974 年, Potter 等人^[36]就得到了声谱图,使人们可以看到声音。声谱图是二维空间中信号的一种表示,其中信号的大小由像素的值表示。而光谱图通常是通过将信号分解成重叠的片段,并对其进行短时快速傅里叶变换来生成^[37]。将生成的光谱图应用在二维卷积神经网络模型中,进而对光谱图上的信息进行预测。

J. Padarian 等人^[38]在研究使用深度学习预测土壤属性中,使用了 Hann 窗口^[39],片段长度为 100,有 50 个重叠的观测值,采样频率为 1。生成光谱图后,将其转换为对数刻度,从而生成了一个 2D 频谱,成功地从长度为 4200 的向量转换为 51×83 的矩阵(频率 \times 波长),如图 3 所示,从而利用 CNN 对一维光谱信号进行建模。Ng Wartini^[40]采用 J. Padarian 提出的方法,转换成光谱图,同样实现了 CNN 的应用,且 CNN 模型表现优于 Cubist 和 PLSR 模型,这与 Padarian 等人发现的结果一致。此外,在 Padarian 等人的研究基础上又测试了一维光谱作为输入的性能,并对二者进行比较,结果发现将一维光谱输入到 CNN 模型比将光谱转换为二维光谱图作为输入性能更优秀。

Qibin Wang 等人^[41]利用 L. Wen^[42]等人提出的一种将一维信号转换为二维图像以满足二维 CNN 输入要求的信号转换方法,从而提出了一种二维深度卷积神经网络,用于快速评估可靠性和预测轴承使用寿命。结果显示在轴承可靠性评估和可靠性预测方面具有良好的准确性和快速计算能力,特别是在时间消耗方面比其他深度学习网络要短。



(a) 矿质土壤(有机碳含量 0.5%)的原始光谱数据 (b) 有机土壤(20%有机碳含量)的原始光谱数据
(a) Raw spectral data of mineral soils (organic carbon content 0.5%) (b) Raw spectral data for organic soils (20% organic carbon content)



(c) 矿质土壤(有机碳含量 0.5%) 的对数刻度的光谱图
(d) 有机土壤(20%有机碳含量) 的对数刻度的光谱图
(c)A log scale spectrogram of a mineral soil (organic carbon content 0.5%)
(d)A spectrogram of an organic soil (20% organic carbon content) on a log scale

图3 光谱数据编码^[38]

Fig.3 Spectral data coding

2.2 光谱信号转变为光谱信息矩阵

一维的光谱信息之间的相关性很难被 CNN 模型发现，而二维光谱信息矩阵可以求取数据之间的自相关，对光谱数据之间的空间信息与原始光谱特征都可以实现特征提取。二维光谱信息矩阵的输入也可以适应卷积层的操作要求，增强了模型的表达能力^[43]。

设 X 为某一样本的光谱数据向量，且为列向量形式，则该样本的二维光谱信息矩阵为：

$$S = XX^T \quad (5)$$

假设 X 是三维光谱数据列向量，则一个典型的二维光谱信息矩阵为：

$$S = \begin{bmatrix} a_1^2 a_1 a_2 a_1 a_3 \\ a_2 a_1 a_2^2 a_2 a_3 \\ a_3 a_1 a_3 a_2 a_3^2 \end{bmatrix} \quad (6)$$

在统计学中，该方阵即为原光谱数据的信息矩阵。信息矩阵的特性就是能够囊括原始数据中的所有信息，并且能够充分体现数据方差和协方差的相对大小。该方法较一维光谱信息输入一维卷积模型而言，对原始光谱信息有着更全面的表达，且符合 CNN 模型的数据处理要求。在光谱分析研究方面，二维光谱信息矩阵与二维相关光谱等高图相似，其特性是能够从一定程度反映光谱数据的波动情况。更有利于 CNN 学习光谱数据的内部结构，取得更好的模型表达效果。王璨^[44]等人在预测土壤含水率时、温馨^[45]在检测水果糖度时，都使用该方法使一维光谱信号成功适用于二维 CNN 模型上，通过训练模型取得了较好的预测结果。

2.3 构建一维卷积网络

二维卷积神经网络在图像上的特征提取有着独到的优势，1D-CNN 同 2D-CNN 不同，输入数据不需要对光谱信号进行处理，在维度上更具适用性；此外，1D-CNN 模型结构更加紧凑，在硬件上的要求更低，使成套配置实时高效、低成本成为可能；值得注意的

是，卷积神经网络对数据集的大小有着较高的要求，而一维数据由于维度上的特点，在训练中可以降低对样本数量的要求也可实现较好的拟合。1D-CNN 的关键就是构建一维卷积核，体现在卷积核大小参数上的设置为 $n \times 1$ 。离散信号 $f(n)$ ， $g(n)$ 的定义如下：

$$S(n) = (f * g)[n] = \sum_{m=0}^{N-1} f(m)g(n-m) \quad (7)$$

式中： N 为信号 $f(n)$ 的长度； $S(n)$ 为卷积结果序列，长度为 $\text{len}(f(n)) + \text{len}(g(n)) - 1$ 。

Serkan Kiranyaz^[46]综述了紧凑一维 CNN 的主要信号处理应用，包括心电图监测的应用、振动信号的应用、电力机械和电路的应用、复杂度分析的计算。并表明采用适当的系统方法，压缩一维 CNN 可以在计算复杂度最小的情况下获得最优的性能。Salim Malek^[7]等人在研究用于光谱信号回归的一维卷积神经网络中，提出了一种利用一维 CNN 进行化学计量数据分析的新方法。考虑输入数据为一维数据，在卷积层上应用了一维卷积核，在训练阶段修改正向传播和反向传播的方程，从而使 2D-CNN 适应了光谱数据的单维性质。Chen Yuanyuan^[17]等人在研究基于集成卷积神经网络的红外光谱定量分析建模中，为了解决一维数据在 CNN 上的不适应性，做了如下转换：假设 1D 的近红外信号是一个特殊的 2D 图像，即图像只包含一行（或一列）。相应地，设计了一维卷积核函数来匹配输入的近红外光谱信号。Chen Yuanyuan^[47]等人又使用同样的方法使数据适用，提出了一种新的基于卷积神经网络的近红外光谱端到端定量分析建模方法。Liu jin chao^[48]等人在利用卷积神经网络识别拉曼光谱的研究中，同样在 CNN 中训练了一维卷积核，也成功实现了拉曼光谱一维数据应用在二维 CNN 上。鲁梦瑶^[49]在使用 CNN 对近红外光谱分类中，将网络中方形矩阵卷积核改为一维向量卷积核，使之适用于近红外光谱数据。

综上，尽管近红外光谱信号作为一维信号与二维 CNN 具有不适应性，但是可以将光谱信号转换为光谱图、构建光谱信息矩阵或构建一维卷积核，训练一维 CNN，使 CNN 应用在光谱信号的分析上成为可能。在研究 GoogLe Net 和 VGG 神经网络结构的时候，他们在某些层中都设置了大小为 1×1 的卷积核，这可以增加网络的深度，增加非线性。可以发现，一维 CNN 模型在光谱信号上的预测效果要优于二维 CNN 模型，但他们的性能都优于传统的方法。值得注意的是一维卷积是针对单方向信息的特征提取，而信息矩阵的构建可以在一维卷积核的基础上，包含更多的空间信息，可以反应信号的波动情况。

3 模型设计

设计 CNN 是一个需要频繁重复的过程,它包括对各个参数的选定,如使用层的数量和类型,以及学习率。在建立模型时,各类参数对模型有着较为重要的影响,需要对其进行分析。模型的训练与优化关系着模型的预测性能,针对不同的目的需要对方法进行综合考虑。本章对模型中结构参数的影响进行了讨论,分析了多任务网络模型的优点,讨论了训练过程中存在的一些问题以及解决办法。

3.1 卷积核参数的影响

原始近红外光谱与卷积核函数卷积运算的模式如图 5 所示可以明显发现,当卷积核函数在整个红外光谱范围内移动时,卷积核函数会自动提取每个红外光谱窗口的局部特征。有两个问题需要关注,首先,卷积核函数的实际形式可能不是高斯分布,它对红外原始数据有自己独特的识别方式,并能提取出抽象的特征;其次,如图 4 所示,移动步长与核大小的关系会影响卷积运算的结果。如果移动步长小于核尺寸(图 4(a)),说明在整个近红外光谱范围内存在一定的重叠区域;如果移动步长等于核尺寸(图 4(b)),则意味着整个近红外光谱范围被平均划分为若干个与间隔 PLS (iPLS) 方法相似的区间;如果移动步长大于内核大小(图 4(c)),则会丢失一些有用的信息;因此,参数设定过程中应尽量避免这种情况。

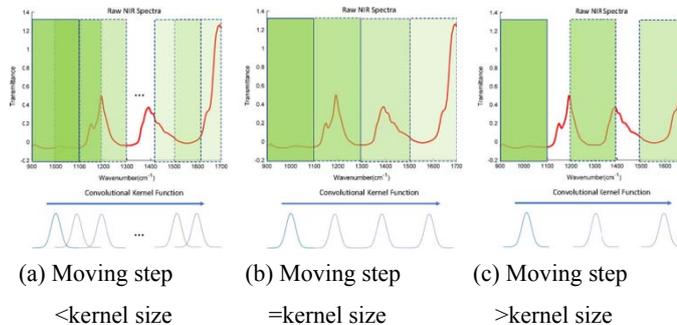


图 4 一维卷积核提取原始红外光谱局部特征模式图^[46]
Fig.4 One dimensional convolution kernel extraction of original IR local feature pattern

鲁梦瑶^[49]在利用一维近红外光谱 CNN 模型对烟叶进行产区分类时,检验了卷积核尺寸对模型性能的影响,得到模型分类正确率随卷积核尺寸变化的曲线如图 5 所示,从而得出卷积核尺寸对模型的判别结果影响不大的结论。Wartini Ng^[40]指出卷积核是用于特征检测的权重矩阵,对于输出特征图的大小有着决定作用。Chen Yuanyuan^[47]等人则认为随着卷积核大小的增加,定量模型的预测性能会提高。这是因为,如果卷积核太小,就很难在每个内核窗口中提取有效的信

息。同时指出卷积核数对模型性能的影响很小,因此我们可以将其设置为一个较小的值,以减少计算量,并验证了步长大小的变换对预测能力的影响很小。

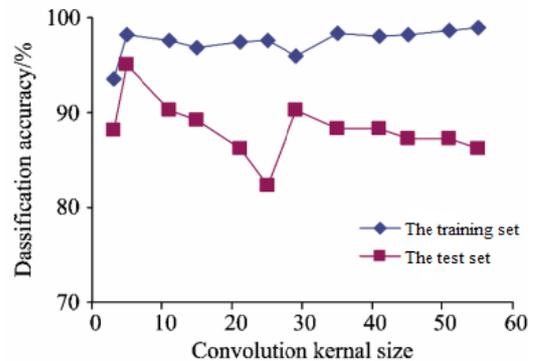


图 5 不同卷积核尺寸的 NIR-CNN 模型判别结果^[49]
Fig.5 The discrimination results of NIR-CNN model with different convolution kernel sizes

综上所述,模型中的参数直接影响了模型的最终性能。在进行模型参数的设置时,要注意避免卷积核移动步长大于卷积核尺寸,从而导致部分信息的缺失。此外,在卷积核的尺寸的设置既不能太小,因其无法获取数据间隐藏的有效信息特征;也不能太大,因其光谱信息上的局部特征可能会丢失。卷积核的最佳尺寸需要在测试中不断验证,找到最合适的范围值。

3.2 多任务网络

CNN 具有在单一网络和训练过程中预测多种特性的能力,并且这种能力十分切合光谱学在分析物质中的应用。并且 Ruder^[50]指出,虽然当一个模型专注于一个单一任务时可以获得普遍可接受的性能,但是共享相关任务的信息可能会改进模型,使用多任务模型可以降低过度拟合的风险。更值得注意的是,准确性随着任务数量的增加而不断增加^[51]。多任务模型的使用已经得到广泛的研究,并得到了良好的效果。

J. Padarian^[38]等人提出了一种多任务网络来预测不同的土壤属性。架构有一系列的 4 个共享卷积层,每个属性随后有一系列的一个卷积核和一个全连接层,如图 6。网络的头部(“公共层”)是一系列卷积层和最大池化层。网络的这一部分由所有目标土壤属性共享,能够了解光谱图的结构。在“共层”提取光谱图数据的一般表示之后,信息被定向到 6 个不同的分支,每个分支对应一个目标土壤属性。发现结果同 Ramsundar 等人^[52]在他们的药物研究中所描述的相同,即准确性随着任务数量的不断增加而提高,这使得使用同一个模型同时从单一光谱中预测多种特性成为可能。

其中,“公共层”表示由所有预测属性共享的层。每个分支对应一个预测土壤性质,对应一个卷积层

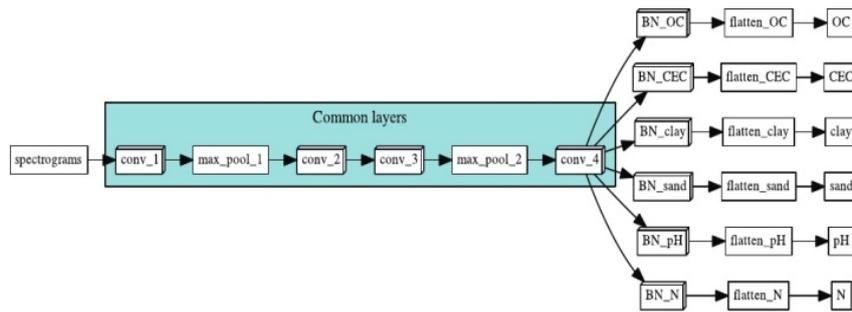


图6 多任务网络的体系结构^[41]

Fig.6 The architecture of a multitasking network

(BN: 瓶颈层, 降低数据的维度) 和一个大小为 1 的全连通层, 对应最终预测。

为多个输出预测创建多个模型不是最优的, 因为从输入中提取的数据可能是冗余的, 并且输入之间的相关性可能不会被保留。多任务学习可以通过在神经网络中拥有一些共享层和保持几个非共享的特定于任务的输出层来实现。这在简单性和计算时间方面有重要的意义, 而且还具有实现协同的能力。Ramsundar 等人^[51]和 Padarian 等人^[38]都发现, 与单任务设置相比, 多任务分配时模型性能显著提高。但是 Padarian 等人发现多任务 CNN 在较小的数据集上并不有效, 甚至表现出比传统的模型更差的性能。所以数据集的大小对于深度学习多任务网络有着很重要的影响。

Wartini Ng^[40]在通过多任务模型研究土壤性质相关性中, 发现无论输入光谱数据的类型如何, 多任务 CNN 被发现在保持数据集的相关系数值方面效果最好, 这些值由最低的绝对差异显示。而传统的 PLSR 和 Cubist 模型分别对每种土壤特性进行了预测, 结果显示土壤特性之间没有保持相关性。虽然多任务 CNN 在学习过程中并没有明确的使用输出的相关结构, 但是多任务优化了所有的输出变量, 从而间接地保持了相关性。保持输出相关性是多任务建模的目标之一^[51]。

3.3 训练

在训练过程中, 为了找到网络的最优权值, 需要使用数据集对网络进行多次训练。通常, 在使用 CNN 时, 数据是成批处理的, 这就使得整个大型数据集可以在同一批次内完成训练, 而完成整个数据集的一次训练称为一次迭代。鲁梦瑶^[49]与 DU Jian^[53]都验证了随着迭代次数的增加, 模型预测性能逐渐提高, 迭代次数到达一定次数时, 性能基本不发生改变, 这表示模型性能已达到最佳。

机器学习, 尤其是深度学习, 是一种受数据集规模大小影响较大的方法。模型的训练依赖于数据集。在训练过程中, 数据集分为训练集、验证集和测试集,

训练集用于拟合或训练模型; 验证集用于参数选择的预测误差; 测试集用于对模型的误差进行评估。

在光谱数据的获取上, 目前仍是复杂且不稳定的, 训练样本较少会使模型的训练更加容易过拟合, 使得预测能力较差, 所以对数据集进行数据增强是有必要的。

Jinchao Liu^[48]为了减少数据量的需求, 使用了扩增, 这是增加 CNN 训练集大小的一种非常常见的方法。提出以下的数据扩充程序: ①将每个频谱向左或向右随机移动几个波数; ②加入随机噪声, 与每个波数的大小成比例; ③对于具有多个光谱的物质, 将属于同一物质的所有光谱的线性组合作为扩增数据。

鲁梦瑶^[49]则讨论了采样间隔对模型训练的影响, 这是针对存在较多光谱数据点时, 要考虑分析选择合适的采样间隔, 采样间隔过大, 准确率会下降; 采样间隔过小, 则会增加网络参数, 增加每次迭代的时间。Salim Malek^[8]等人为克服训练样本数量有限而导致的过拟合问题, 引入了粒子群优化算法来估计一维 CNN 的不同参数 (权值), 以代替标准的反向传播算法。

Du Jian^[53]等人为了改善光谱数据维度较高与训练样本较少之间的不平衡, 在模型中采用 ReLU (Rectified Linear Unit) 和 Dropout 方法。ReLU 用来增强模型表达能力并加速收敛。Dropout 适用于光谱数据高维输入的特征, 可以提高整个模型的鲁棒性。并在训练结果中发现神经元较少时, 模型收敛速度较慢, 随着神经元个数的增加, 收敛速度明显加快, 但同时训练时间也明显增加, 训练误差曲线波动较大。

在训练中, 根据基于梯度的优化方法调整权重, 即对参数相对于误差的偏导数进行评估并将参数调整到最小误差值。参数沿误差梯度的变化率由学习率控制。如果学习率太高, 权重会随着每次迭代变化过多, 这将使参数在最优解附近波动, 或者只是发散; 如果学习率太低, 参数可能永远不会收敛。Jinchao Liu^[48]等人在训练 CNN 模型时, 使用了 Adam 的训练算法^[54], 一种随机梯度算法的变种, 实现了较好的权重调整。

Chen Yuanyuan^[47]等人在建模过程中,提出了使用自举随机抽样机制,能够保证定量分析模型的稳定性(也称“鲁棒性”)优于传统方法。并且将 L1 范数正则化优化算法与 BP 神经网络结合,分别建立每个数据集的定量分析模型后,发现 CNN 模型的广义性能比基于波长选择的模型性能要好,这证明 CNN 模型经过较好的训练后,具有更好的波长选择能力。

4 总结

在光谱学分析领域,深度学习方法具有非常大的应用潜力^[55]。本文针对卷积神经网络在光谱分析建模中的应用进行综述。很多研究都已验证过 CNN 模型在光谱分析上的性能要优于传统的建模方法。此外,传统的建模方法需要对数据进行各种数据预处理,这不仅会影响到模型的性能还需要人们掌握更多的专业知识,增加分析难度。但目前的研究已证明不经过数据预处理的 CNN 模型也可以很好地实现预测能力甚至更优于传统方法。

CNN 模型应用到红外光谱中的好处包括:1)多任务学习能力,该模型可以同时预测多个输出,同时保持所有输出之间的相关性,并对组成特性做出准确的估计;2)多维输入利用率:该模型可以合并不同维度的多个输入通道;3)更高的性能:与其他回归模型如三次回归、偏最小二乘回归等相比,CNN 模型具有更高的精度性能。但是 CNN 模型也有其局限性:1)深度学习需要大量的数据集,否则很容易过拟合;2)CNN 模型所使用的参数数量较大,微调过程会很困难;3)由于需要训练许多参数,因此对硬件的要求较高。

综上所述,卷积神经网络在光谱分析上的应用是值得肯定的,并且已经取得较为理想的发展,但是仍然存在一些问题和困难,例如:光谱数据的最佳输入方式、模型的最佳优化算法、光谱数据集的限制,以及网络模型结构的功能性创新等都值得我们做更多的研究与探索。

参考文献:

[1] Gente R, Busch S F, Eva-Maria Stübling, et al. Quality control of sugar beet seeds with THz time-domain spectroscopy[J]. *IEEE Transactions on Terahertz ence & Technology*, 2016, **6**(5):754-756.

[2] Przybylek P. A new method for indirect measurement of water content in fibrous electro-insulating materials using near-infrared spectroscopy[J]. *IEEE Transactions on Dielectrics and Electrical Insulation*, 2016, **23**(3):1798-1804.

[3] Hiroaki I, Toyonori N, Eiji T. Measurement of pesticide residues in food based on diffuse reflectance IR spectroscopy[J]. *IEEE Transactions on Instrumentation and Measurement*, 2002, **51**(5): 886-890.

[4] Mignani A G, Ciaccheri L, Mencaglia A A, et al. Spectroscopy AS a

“green” technique for food quality and safety applications[C]//*Technical Digest of the Eighteenth Microoptics Conference of IEEE*, 2013: 1-2.

[5] Nishizawa S, Morita H, Iwamoto T, et al. Terahertz time-domain spectroscopy applied to nondestructive evaluation of pharmaceutical products[C]//2011 *International Conference on Infrared, Millimeter, and Terahertz Waves of IEEE*, 2011: 1-2.

[6] ZOU Xiaobo, ZHAO Jiewen, Povey M J W, et al. Variables selection methods in near-infrared spectroscopy[J]. *Analytica Chimica Acta*, 2010, **667**(1-2): 14-32.

[7] 周宣. 基于新型冠状病毒肺炎防护的医用口罩分类与使用[J]. *医疗装备*, 2020(15): 10-12.

ZHOU Xuan. Classification and use of medical masks based on new Coronavirus pneumonia protection[J]. *Medical Equipment*, 2020(15): 10-12.

[8] Malek S, Melgani F, Bazi Y. One-dimensional convolutional neural networks for spectroscopic signal regression[J]. *Journal of Chemometrics*, 2017: e2977.

[9] LIU Xuemei, ZHANG Hailiang, SUN Xudong, et al. NIR sensitive wavelength selection based on different methods[C]//2010 *International Conference on Mechanic Automation and Control Engineering*, 2010: 26-28.

[10] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[C]//*Advances in Neural Information Processing Systems*, 2012: 1097-1105.

[11] Devos O, Ruckebusch C, Durand A, et al. Support vector machines (SVM) in near infrared (NIR) spectroscopy: focus on parameters optimization and model interpretation[J]. *Chemometrics and Intelligent Laboratory Systems*, 2009, **96**(1): 27-33.

[12] Demeulemeester J, Smeets D, Barradas N P, et al. Artificial neural networks for instantaneous analysis of real-time rutherford backscattering spectra[J]. *Nuclear Instruments and Methods in Physics Research*, 2010, **268**(10): 1676-1681.

[13] Lee S, Choi H, Cha K, et al. Random forest as a potential multivariate method for near-infrared (NIR) spectroscopic analysis of complex mixture samples: Gasoline and naphtha[J]. *Microchemical Journal*, 2013, **110**: 739-748.

[14] McCarty G W, Reeves J B, Reeves V B, et al. Mid-infrared and near - infrared diffuse reflectance spectroscopy for soil carbon measurement[J]. *Soil Science Society of America Journal*, 2002, **66**(2): 640-646.

[15] Gerretzen J, Szymańska E, Jansen J J, et al. Simple and effective way for data preprocessing selection based on design of experiments[J]. *Analytical Chemistry*, 2015, **87**(24): 12096-12103.

[16] Hubel D H, Wiesel T N. Receptive fields and functional architecture of monkey striate cortex[J]. *The Journal of Physiology*, 1968, **195**(1): 215-243.

[17] CHEN Yuanyuan, WANG Zhibin. Quantitative analysis modeling of infrared spectroscopy based on ensemble convolutional neural

- networks[J]. *Chemometrics and Intelligent Laboratory Systems*, 2018, **181**: 1-10.
- [18] NI C, WANG D, TAO Y. Variable weighted convolutional neural network for the nitrogen content quantization of Masson pine seedling leaves with near-infrared spectroscopy[J]. *Spectrochimica Acta Part A Molecular & Biomolecular Spectroscopy*, 2019, **209**: 32-39.
- [19] LeCun Y. The MNIST database of handwritten digits[EB/OL]. <http://yann.lecun.com/exdb/mnist/>, 1998.
- [20] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]//*International Conference on Neural Information Processing Systems*, 2012:1097-1105.
- [21] CHENG G, ZHOU P, HAN J. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, **54**(12): 7405-7415.
- [22] LeCun Y, Boser B E, Denker J S, et al. Handwritten digit recognition with a back-propagation network[C]//*Advances in Neural Information Processing Systems*, 1990: 396-404.
- [23] GU J, WANG Z, Kuen J, et al. Recent advances in convolutional neural networks[J]. *Pattern Recognition*, 2018, **77**: 354-377.
- [24] WANG T, WU D J, Coates A, et al. End-to-end text recognition with convolutional neural networks[C]//*Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012) of IEEE*, 2012: 3304-3308.
- [25] XU B, WANG N, CHEN T, et al. Empirical evaluation of rectified activations in convolutional network[J/OL]. *arXiv preprint arXiv:1505.00853*, 2015.
- [26] LeCun Y A, Bottou L, Orr G B, et al. *Efficient Backprop*[M]//*Neural Networks: Tricks of the Trade*, Springer, 2012: 9-48.
- [27] Nwankpa C, Ijomah W, Gachagan A, et al. Activation functions: Comparison of trends in practice and research for deep learning[J/OL]. *arXiv preprint arXiv:1811.03378*, 2018.
- [28] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, **37**(9): 1904-1916.
- [29] Boureau Y L, Ponce J, LeCun Y. A theoretical analysis of feature pooling in visual recognition[C]//*Proceedings of the 27th International Conference on Machine Learning (ICML-10)*. 2010: 111-118.
- [30] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift[J/OL]. *arXiv preprint arXiv:1502.03167*, 2015.
- [31] Hinton G E, Srivastava N, Krizhevsky A, et al. Improving neural networks by preventing co-adaptation of feature detectors[J/OL]. *arXiv preprint arXiv:1207.0580*, 2012.
- [32] Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. *The Journal of Machine Learning Research*, 2014, **15**(1): 1929-1958.
- [33] Khan A, Sohail A, Zahoor U, et al. A survey of the recent architectures of deep convolutional neural networks[J]. *Artificial Intelligence Review*, 2020, **53**(8): 5455-5516.
- [34] LIN M, CHEN Q, YAN S. Network in network[J/OL]. *arXiv preprint arXiv:1312.4400*, 2013.
- [35] Rawat W, WANG Z. Deep convolutional neural networks for image classification: a comprehensive review[J]. *Neural Computation*, 2017, **29**(1): 2352-2449.
- [36] Potter R K, Kopp G A, Green H C. *Visible Speech*, New York, 1947[J]. *D. Van Nostrand Co.*, 1962(8): 39.
- [37] Griffin D, Lim J. Signal estimation from modified short-time Fourier transform[J]. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1984, **32**(2): 236-243.
- [38] Padarian J, Minasny B, McBratney A B. Using deep learning to predict soil properties from regional spectral data[J]. *Geoderma Regional*, 2019, **16**: e00198.
- [39] Blackman R B, Tukey J W. The measurement of power spectra from the point of view of communications engineering[J]. *Bell System Technical Journal*, 1958, **37**(1): 185-282.
- [40] Ng W, Minasny B, Montazerolghaem M, et al. Convolutional neural network for simultaneous prediction of several soil properties using visible/near-infrared, mid-infrared, and their combined spectra[J]. *Geoderma*, 2019, **352**: 251-267.
- [41] WANG Q, BO Z, MA H, et al. A method for rapidly evaluating reliability and predicting remaining useful life using two-dimensional convolutional neural network with signal conversion[J]. *Journal of Mechanical Science and Technology*, 2019, **33**(6): 2561-2571.
- [42] WEN L, LI X, GAO L, et al. A new convolutional neural network-based data-driven fault diagnosis method[J]. *IEEE Transactions on Industrial Electronics*, 2017, **65**(7): 5990-5998.
- [43] 谢丽娟. 转基因番茄的可见/近红外光谱快速无损检测方法[D]. 杭州: 浙江大学, 2009.
- XIE Lijuan. Rapid non-destructive detection of Transgenic tomatoes by visible/near-infrared Spectroscopy[D]. Hangzhou: Zhejiang University, 2009.
- [44] 王璨, 武新慧, 李恋卿, 等. 卷积神经网络用于近红外光谱预测土壤含水率[J]. *光谱学与光谱分析*, 2018, **38**(1): 42-47.
- WANG Can, WU Xinhui, LI Xiangqing, et al. Application of convolutional neural network in near infrared spectroscopy to predict soil moisture content[J]. *Spectroscopy and Spectral Analysis*, 2018, **38**(1): 42-47.
- [45] 温馨. 基于深度学习的水果糖度可见/近红外光谱无损检测方法研究[D]. 北京: 北京交通大学, 2018.
- WEN Xin. A Nondestructive Testing Method for visible/near-infrared spectra of fruit Sugar Based on Deep learning [D]. Beijing: Beijing Jiaotong University, 2018.
- [46] Kiranyaz S, Ince T, Abdeljaber O, et al. 1-d convolutional neural networks for signal processing applications[C]//*2019 IEEE International*

- Conference on Acoustics, Speech and Signal Processing (ICASSP) of IEEE*, 2019: 8360-8364.
- [47] CHEN Y Y, WANG Z B. End-to-end quantitative analysis modeling of near - infrared spectroscopy based on convolutional neural network[J]. *Journal of Chemometrics*, 2019, **33**(5): e3122.
- [48] LIU J, Osadchy M, Ashton L, et al. Deep convolutional neural networks for Raman spectrum recognition: a unified solution[J]. *Analyst*, 2017, **142**(21): 4067-4074.
- [49] 鲁梦瑶, 杨凯, 宋鹏飞, 等. 基于卷积神经网络的烟叶近红外光谱分类建模方法研究[J]. *光谱学与光谱分析*, 2018, **38**(12):78-82.
- LU M Y, YANG K, SONG P F, et al. The study of classification modeling method for near infrared spectroscopy of tobacco leaves based on convolution neural network[J]. *Spectroscopy and Spectral Analysis*, 2018, **38**(12):78-82.
- [50] Ruder S. An overview of multi-task learning in deep neural networks [J/OL]. arXiv preprint arXiv:1706.05098, 2017.
- [51] ZHANG Y, YANG Q. A survey on multi-task learning[J/OL]. arXiv preprint arXiv:1707.08114, 2017.
- [52] Ramsundar B, Kearnes S, Riley P, et al. Massively multitask networks for drug discovery[J/OL]. arXiv preprint arXiv:1502.02072, 2015.
- [53] DU Jian, HU Bingliang, LIU Yongzheng, et al. Study on quality identification of macadamia nut based on convolutional neural networks and spectral features[J]. *Spectroscopy and Spectral Analysis*, 2018, **38**(5): 1514-1519.
- [54] Kingma D P, Ba J Adam: a method for stochastic optimization[J/OL]. arXiv preprint arXiv:1412.6980, 2014.
- [55] Acquarelli J, van Laarhoven T, Gerretzen J, et al. Convolutional neural networks for vibrational spectroscopic data analysis[J]. *Analytica Chimica Acta*, 2017, **954**: 22-31.

<书讯>

光电系统设计基础

本书于2021年5月由清华大学出版社出版,作者基于高校光电类专业培养高层次、综合性、创新性人才的需求,以及“新工科”课程建设的实际需要,从工程实际和教学出发,编著适合光学工程学科研究生和光电信息科学与工程专业高年级本科生的教材,同时也适合光电行业的技术、管理人员学习参考,提高光电系统总体技术设计能力。

本书是作者结合自己长期从事光电系统(装备)工作的经验,将基础理论与工程技术相结合,融合科研成果与心得体会的一部教材。

本书内容丰富,深入浅出;取材科研,源于工程;突出重点,应用性强;系统全面,概念清晰;通过示例,触类旁通;层次分明,易教易学。

《光电系统设计基础》与作者于2019年和2018年编著出版的《光电系统设计——方法、实用技术及应用》、《光电系统环境与可靠性工程技术》,一起构成一套有特色、有联系、有深度、有广度的专业用书,对从事光电系统总体设计、开发的行业技术人员,以及从事各种光电系统研究的科研人员,都具有很高的实用参考价值。

作者简介:

吴晗平,男,1964年10月生,工学博士,教授,博士生导师。任中国光学学会光学教育专业委员会常务委员,中国兵工学会光电子技术专业委员会委员,中国电子学会量子电子学与光电子学分会委员,教育部学位与研究生教育发展研究中心评议专家,湖北省仪器仪表学会副理事长,湖北省光学学会常务理事,以及湖北省重大科技成果转化与产业化项目、省自然科学基金项目、省创新创业战略团队项目评审专家等。同时担任《光学精密工程》《光子学报》《中国光学》《激光与红外》《应用光学》《红外技术》《现代电子技术》《光电技术应用》等专业学术期刊编委。曾从事光电系统(装备)的立项论证、技术研究、设计、研制、检验、试验等一线工作近二十年,主持、承担或参与完成了多个国防高新重大光电装备研制项目和军队科研课题。曾获省(部)级科学技术奖一等奖两次(其中国防科学技术奖一等奖一次),省(部)级科学技术进步奖三等奖四次。出版国家军用标准一项。独撰出版学术著作(专著)四部,独撰出版教育部高等学校电子信息类专业教学指导委员会规划教材两部。先后在国内外三十多种专业学术期刊上发表论文一百余篇。主要研究方向为光电系统总体技术与设计,环境与可靠性工程技术,红外与紫外技术等。

