

## 遗传算法和神经网络的重叠光谱解析

都月, 孟晓辰\*, 祝连庆

北京信息科技大学光电测试技术北京市重点实验室, 生物医学检测技术及仪器北京实验室, 北京 100101

**摘要** 随着光谱分析及荧光检测技术的快速发展, 单色荧光标记已无法对细胞样本进行精准判断, 必须采用双染色或多色荧光标记来分析细胞内部结构。然而, 使用光谱测量方法进行多色荧光分析时, 由于通常使用多种标记物同时对待测细胞进行标记, 发射光谱会产生部分光谱重叠, 为了准确对其进行分析, 需将重叠峰分解为独立谱峰。针对光谱重叠现象, 提出了遗传算法优化 BP 神经网络(GA\_BP)的重叠峰解析算法。首先确定了 BP 神经网络具体结构, 并对重叠峰信号进行二次微分预处理, 确定重叠峰中单峰个数及单峰位置, 将其作为重叠峰信号的特征值送入 BP 神经网络的输入层; 其次将 BP 神经网络权值及阈值初始化, 利用遗传算法全局搜索的优势, 进行算法初始种群及种群规模等最优参数的选取, 通过选择、交叉、变异等一系列遗传进化操作进行寻优计算, 得到包含 BP 神经网络最优权值和阈值的个体; 然后确定网络最优参数并进行相应网络训练, 使优化后的 BP 神经网络可从输出节点处获得独立单峰的峰宽及强度; 最后结合二次微分处理得到的重叠峰特征值, 即可分离出单个谱峰。以随机生成的多组高斯重叠峰数学模型作为实验数据进行仿真实验, 结果表明该方法具有较高的精确度。其中, 双峰重叠峰及三峰重叠峰分解后峰强度及峰宽的最大相对误差分别为 0.30%, 3.57% 和 0.64%, 3.83%; 同时也可对四峰重叠峰进行较为准确的分解。此外, 将 GA\_BP 网络模型与未经优化的 BP 神经网络模型作对比, 结果表明 GA\_BP 网络运行 5 步后即可达到预设的误差值, 而未经优化的网络模型则需 19 步方可达到, 进一步证明 GA\_BP 网络模型收敛更快且误差较低。由此可见, GA\_BP 算法在重叠光谱分析中有较好的效果, 并可应用于其他能谱重叠峰的分解, 与传统方法相比具有明显的优势, 具有一定的实用价值。

**关键词** 生物医学; 重叠峰分解; 神经网络; 遗传算法

**中图分类号:** TN911.74 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2020)07-2066-07

### 引言

随着生物医学的快速发展, 光谱分析已被广泛应用在细胞免疫学、微生物学、分子遗传学等方面。其中, 荧光检测法由于灵敏度高、信息量丰富等特点, 已发展成为一种科学且有效的研究手段。近年来, 荧光检测的仪器种类不断增加, 发展迅速, 已从最初的间接免疫荧光染色逐步发展到多色荧光分析, 通过测定细胞内的多种荧光, 在一次实验中同时测得多种分量。由于待测细胞通常携带两种或者两种以上的荧光素, 尽管它们的发射峰不相同, 但发射光谱范围仍会有一定的重叠。在此情况下, 进一步对光谱进行研究分析的实现难度非常大, 所以需要准确的分离光谱重叠峰。

重叠峰的分离一直是光谱研究、电化学分析等领域中重点研究的重点方向。目前常用的分峰方法主要分为数学法及化学计量学两类。数学法具有直观、计算速度快等特点, 适合进行在线实时分离<sup>[1]</sup>, 但有较大的局限性, 通常要求构成重叠峰的各单峰均为相似的, 故而已逐渐被代数法所取代。在代数法中, 常用的方法是主成分曲线拟合分析<sup>[2]</sup>、多元回归分析<sup>[3]</sup>、神经网络分析<sup>[4-5]</sup>、智能算法分析<sup>[6, 10]</sup>及小波变换<sup>[2, 4]</sup>等。其中, 神经网络具有非线性适应性信息处理的能力, 克服了传统人工智能方法的缺陷; 遗传算法(genetic algorithm, GA)由于其全局寻优能力、自适应调整搜索方向等优势, 已被广泛应用于机器学习、信号处理等各方面。针对重叠光谱峰分解这一难点问题, 本文将遗传算法与神经网络相结合, 避免了网络收敛于局部最优并陷入局部极小等情

收稿日期: 2019-06-23, 修订日期: 2019-10-10

基金项目: 国家自然科学基金项目(61605010), 北京信息科技大学促进高校内涵发展科研水平提高项目(521191092), 北京信息科技大学“勤信人才”培育计划项目(QXTCP C201903)资助

作者简介: 都月, 1996年生, 北京信息科技大学仪器科学与光电工程学院硕士研究生 e-mail: duyue2013@163.com

\* 通讯联系人 e-mail: mengxc@bistu.edu.cn

况,提高了全局的搜索能力<sup>[8]</sup>,对构成重叠峰的几个单峰参数进行准确预测,以实现光谱重叠峰的准确分离。

### 1 算法实现

遗传算法优化神经网络的重叠峰解析算法整体流程如图 1 所示。

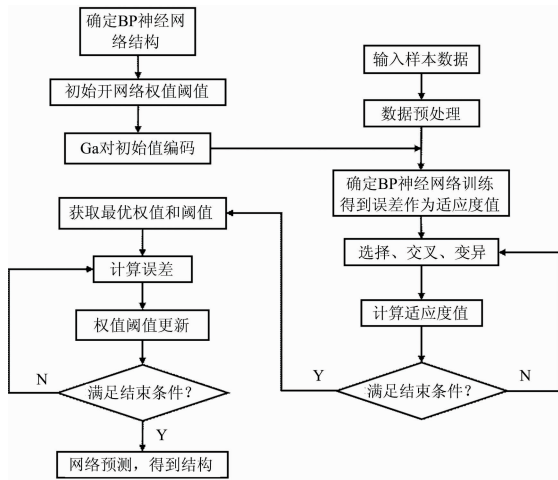


图 1 GA\_BP 算法解析重叠峰流程图  
Fig. 1 Flow chart of overlapping peaks parsed by GA\_BP algorithm

考虑到部分重叠峰信号分离度较小,为初步判断单峰个数及其所对应位置,进行重叠峰解析之前,需对重叠峰进行预处理来增大其分离度。随后根据神经网络自适应调整参数这一特性,结合遗传算法搜索网络的最优参数并作出相应改进,对神经网络进行学习和训练,从而使预测输出不断逼近期望输出,以实现重叠光谱信号的准确识别及分离。

#### 1.1 重叠峰信号预处理

微分法因其独特的性质,能准确的判别出函数中曲率发生变化的点,同时能够较为准确的解析信号内部结构,运算方便,操作简单<sup>[7]</sup>。故本文考虑对重叠光谱信号进行微分预处理,得出曲线的相关参数(如极值点等),进一步分析重叠峰的相关特性。

一般情况下,通常采用二次微分来对原始信号进行预处理。如图所示,图 2(a)是原始重叠峰信号,横坐标为波长范围,纵坐标为峰强度,该重叠峰由五个单峰叠加而成。图 2(b)是其二次微分后得到的曲线,该重叠峰经过二次微分预处理之后,构成重叠峰的各单峰可被初步分辨,与原信号相比峰形变窄,半峰宽变小,峰位置不变。每相邻两单峰的峰谷处,分别对应二次微分曲线的极大值点,由此可确定重叠峰中单峰的个数和其对应的位置。

#### 1.2 GA\_BP 算法原理

重叠峰信号经二次微分预处理后,可准确得到各个单峰的位置、强度及峰宽等信息,这是分离重叠峰的关键。但是各特征值之间的联系很难用数学模型计算,因此考虑神经网络来学习其内部复杂的关系,来逼近未知的输入-输出映射。

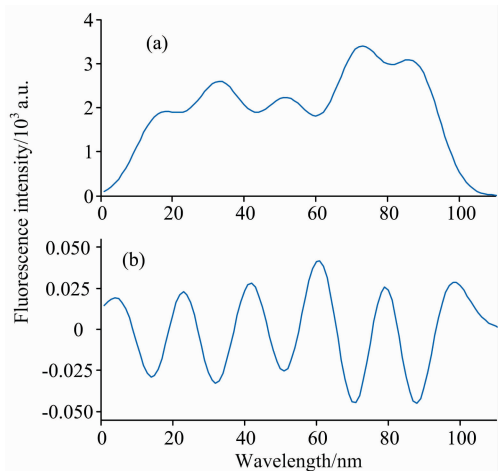


图 2 重叠峰信号及其二次微分曲线  
Fig. 2 Overlapping peaks and its quadratic differential curve

BP 神经网络主要特点是信号前向传递,误差反向传播,根据神经网络预测误差不断调整网络的权值和阈值来使网络输出的均方误差最小,达到预测输出不断逼近期望输出的目的<sup>[8-9]</sup>。故采用 BP 神经网络进行单峰提取。

研究表明,神经网络的初始权值和阈值是影响网络的学习和训练的重要因素<sup>[8]</sup>。但由于其较强的敏感性和随机性,一定程度上会导致网络输出不稳定。针对此问题,考虑使用遗传算法对其进行优化。遗传算法是一类模拟达尔文生物进化论,借鉴生物界自然选择和遗传学的随机优化算法,在其搜索过程中,通过模拟一系列进化操作,如选择、交叉、变异等,搜索最优个体并得到问题的最优解<sup>[7-8]</sup>。因此,利用遗传算法自适应性、自学习性、全局搜索等优势,对 BP 神经网络的初始权值和节点阈值进行寻优计算,确定其最优值后进行网络训练并准确预测输出。

结合重叠光谱峰分解,GA\_BP 算法具体由以下步骤(1)~(3)实现:

(1)构建 BP 神经网络:建立由输入层、输出层和隐含层构成的多层感知器的 BP 神经网络,确定网络的输入层节点数  $n$ 、隐含层节点数  $l$ ,隐含层激励函数  $f$ ,输出层节点数  $m$ ,将输入层与隐含层之间的连接权值  $\omega_{ij}$  及隐含层阈值  $a$  初始化,确定学习速率和激活函数等参数。将样本数据送入输入节点,经隐含层逐层处理后,由式(1)计算隐含层输出  $H_j$ 。

$$H_j = f\left(\sum_{i=1}^n \omega_{ij} x_i - a_j\right) \quad j = 1, 2, \dots, l \quad (1)$$

将隐含层与输出层之间的权值  $\omega_{jk}$  及输出层阈值  $b$  初始化,计算 BP 神经网络预测输出  $O_k$

$$O_k = \sum_{j=1}^l H_j \omega_{jk} - b_k \quad k = 1, 2, \dots, m \quad (2)$$

根据网络预测输出  $O_k$  和期望输出  $Y$ ,按式(3)计算网络预测误差。计算每次输入样本数据的误差值,若两者误差相对较大,则误差开始反向传输<sup>[9]</sup>,并调整神经网络参数,直到系统收敛。

$$e_k = Y_k - O_k \quad k = 1, 2, \dots, m \quad (3)$$

构造的多层 BP 神经网络系统中,输入层  $X$  为重叠峰信

号经二次微分后所得到的特征参数如峰高、峰位及标准差(以下称为峰宽),神经网络经训练及优化后,即可从输出节点  $Y$  提取出重叠峰信号中需要分离的单峰的实际参数(如单峰宽度和强度)。

(2)遗传算法优化:遗传算法优化过程中,种群内包含了待解决问题所有潜在的解集,每个种群都由经编码后的一定数量的个体组成,且每个个体均含有 BP 神经网络所有初始权值和阈值。采用实数编码,将每个个体看作一个实数串,将问题的解映射为生物进化中的染色体形式,进而生成初始群体。使用适应度值用来评价染色体的优劣,定义个体适应度值  $F$  为 BP 神经网络经训练后系统的预测输出和期望输出之间的误差绝对值和,

$$F = k \left( \sum_{i=1}^m \text{abs}(y_i - o_k) \right) \quad (4)$$

式(4)中,  $y_i$  为网络第  $i$  个节点的期望输出;  $o_k$  为其预测输出;  $k$  为系数。进化过程中,采用轮盘赌选择方法,选出适应度高的个体。由于采用实数编码方式,故交叉操作中采用实数交叉法,循环多次利用交叉算子,设置合适的交叉及变异概率,最终得到包含网络初始权值和阈值的最优解。

(3)BP 神经网络预测:通过网络训练,根据网络预测误差  $e$  更新权值  $\omega_{ij}$ ,  $\omega_{jk}$

$$\omega_{ij} = \omega_{ij} + \eta H_j (1 - H_j) x(i) \sum_{k=1}^m \omega_{jk} e_k$$

$$i = 1, 2, \dots, n; j = 1, 2, \dots, l \quad (5)$$

$$\omega_{jk} = \omega_{jk} + \eta H_j e_k$$

$$j = 1, 2, \dots, l; k = 1, 2, \dots, m \quad (6)$$

更新网络节点阈值  $a$  和  $b$

$$a_j = a_j + \eta H_j (1 - H_j) x(x) \sum_{k=1}^m \omega_{jk} e_k$$

$$j = 1, 2, \dots, l \quad (7)$$

$$b_k = b_k + e_k \quad k = 1, 2, \dots, m \quad (8)$$

神经网络进行训练后,将训练样本模式分布记忆在权值和阈值中,使预测输出值与期望值误差最小,从而达到准确预测的目的。对待测重叠峰信号,网络将从输入节点预测输出重叠峰的几个单峰峰宽和强度。结合各单峰位置,即可恢复独立单峰。

## 2 算法最优参数的选取

重叠峰解析过程中,遗传算法对 BP 神经网络的权值和阈值的寻优过程是得到准确分析结果的关键。个体适应度值作为影响初始种群质量的一个重要因素,决定了种群进化的方向。为提高初始种群质量,本文以初始种群的适应度值作为评价指标,设置适应度阈值为  $t$ ,分别计算随机生成的种群中每个个体的适应度值。若其适应度值大于  $t$ ,则将其作为初始种群的一部分进行下一步进化操作,否则被淘汰。所有个体计算完毕之后,选择适应度值相符的个体作为本次算法运行的初始种群。通过这种方式生成的初始种群质量较高,便于收敛到全局最优解。此外,可将算法前一次运行得到的最后种群作为下一次运行的初始种群,如此会得到更准

确的结果。

种群规模的大小同样影响遗传算法准确度。种群规模太小时,交叉运算作用大幅降低,导致难以产生新的个体,变异算子执行能力降低,并且对已生成的模型具有极大的破坏作用,阻碍种群中有效模式的正确传播,使得种群进化过程中不能产生所期望的预测数量;种群规模太大时,算法难以收敛,繁琐性增大且稳定性下降。为确保算法精度,在其他参数相同的情况下,依次选取种群规模为 20, 30, 40 及 50 的初始种群,分别计算遗传算法优化后的神经网络输出误差。如图 3 可知,在特定应用环境下,种群规模为 40 的时候,网络预测误差最小。鉴于此,后续的讨论均基于种群规模 40 之上。

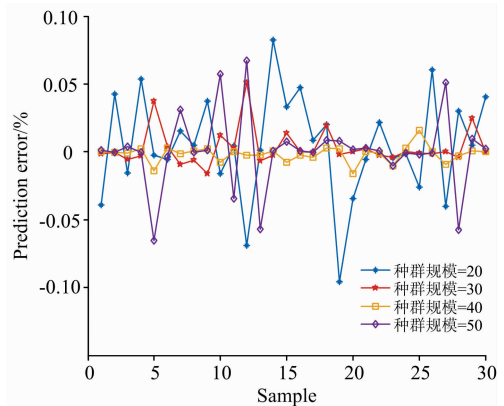


图 3 种群规模对神经网络预测误差的影响

Fig. 3 Effect of population size on prediction error of neural network

## 3 重叠峰分解

为检验本算法对重叠峰分析的有效性,使用 Matlab 软件,随机生成 200 组不含噪声且具有不同特征值的仿真重叠光谱作为样本数据。通过二次微分对重叠峰进行预处理,提取各个重叠峰曲线的特征参数,并结合单峰的峰宽及强度作为训练数据,并从中随机选取 50 组样本数据用来测试网络的预测性能。使用遗传算法优化后,用训练好的网络进行各单峰预测,并对预测结果加以分析。

### 3.1 双峰分解

以高斯模型  $ae^{-(x-b)^2/2c^2}$  作为单峰数学模型,其中  $a$  为峰强度,  $b$  为峰位,  $c$  为标准差(峰宽)。如图 4(a)所示,本节需要分解的重叠峰由两个高斯峰 A 和 B 叠加而成,两峰的峰强度、峰位及峰宽分别为  $a_1 = 1.7$ ,  $b_1 = 40$ ,  $c_1 = 80.6$ ;  $a_2 = 2.5$ ,  $b_2 = 60$ ,  $c_2 = 106.1$ 。对该信号进行二次微分预处理,得到的曲线如图 4(b)所示,横坐标为波长范围,纵坐标为峰强度。

构建神经网络,设置隐含层节点数为 7,学习速率为 0.1。对其进行遗传算法优化,设置交叉概率和变异概率为  $[0, 1]$  之间,最大迭代次数 50 次。随着迭代次数的增加,个体适应度值的变化情况如图 5 所示,经过 50 次迭代运算后算法结束,此时最优个体的适应度值为 0.176 4。

定义神经网络预测误差为

$$W = \frac{Y_{\text{预测值}} - Y_{\text{真实值}}}{Y_{\text{真实值}}} \times 100\% \quad (9)$$

把最优权值和阈值赋给神经网络,经训练后计算神经网络预测误差。其中图 6(a)表示 A 峰和 B 峰强度的预测误差,图 6(b)表示其峰宽的预测误差。由图可知,峰强度预测误差稳定在 0.01% 以内,标准偏差(峰宽)预测误差稳定在 0.20% 以内。

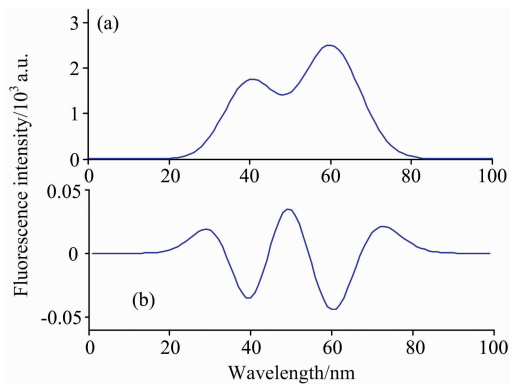


图 4 双重叠峰及其二次微分曲线  
Fig. 4 Two overlapping peaks and its quadratic differential curves

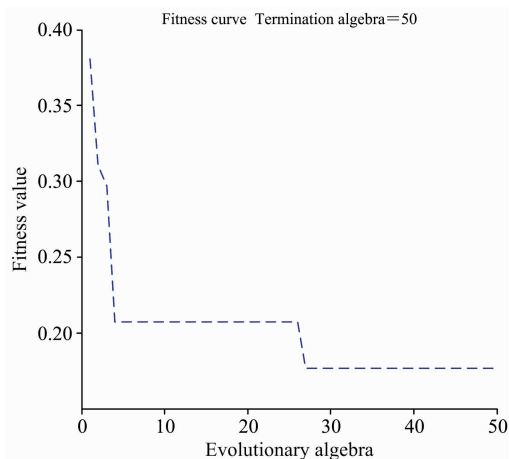


图 5 适应度值变化过程  
Fig. 5 Fitness changes

通过描述的神经网络反向传播算法,结合二次微分法确定的单峰峰位,可对已提取特征值的重叠峰进行分离操作。表 1 为双峰重叠解析后预测值与真实值对比,图 7 为原始重叠峰与分峰后的各单峰曲线。由此可见,使用该方法分解后单峰可以清晰地显示出来,且叠加后的曲线与原始重叠峰相差甚小,具有很高的精度。

### 3.2 三峰分解

需要分解的重叠峰如图 8(a)所示,其由三个峰强度、峰值及峰宽分别为  $a_1 = 2.0, b_1 = 35, c_1 = 100; a_2 = 1.7, b_2 = 52, c_2 = 84.5; a_3 = 2.9, b_3 = 70, c_3 = 93.8$  的高斯峰 A, B 和 C 叠加而成,可见重叠严重。对该信号进行二次微分后,得

到的曲线如图 8(b)所示。

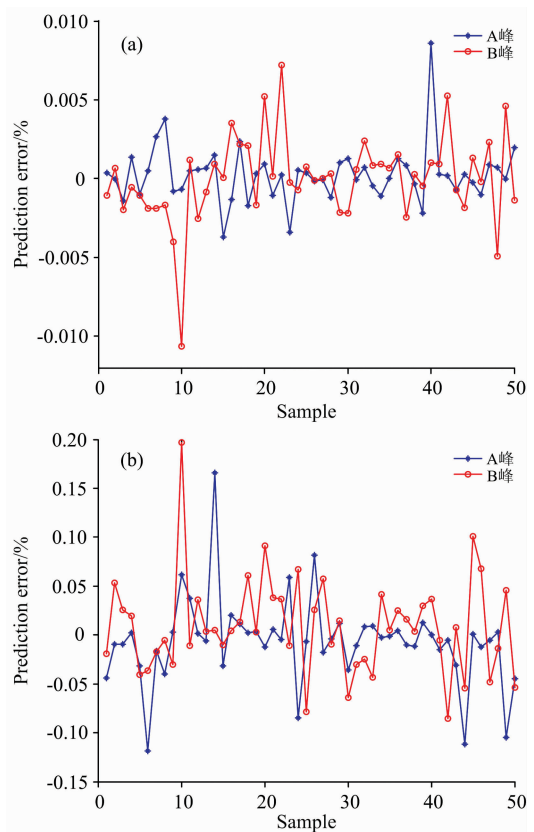


图 6 (a) 峰强度预测误差; (b) 峰宽预测误差  
Fig. 6 (a) Prediction error of peak strength; (b) Prediction error of peak width

表 1 双峰重叠解析结果

Table 1 Analysis results at two overlapping peaks

	真实值	预测值	相对误差/%
$a_1$	1.7	1.701 2	0.07
$a_2$	2.5	2.492 6	-0.30
$c_1$	80.6	77.719 7	-3.57
$c_2$	106.1	105.655 6	-0.42

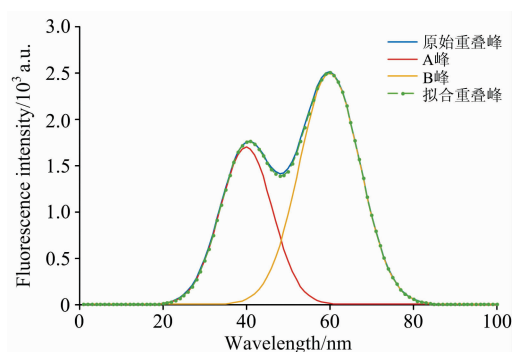


图 7 原始重叠峰、分离单峰及拟合重叠峰  
Fig. 7 Original spectrum, decomposed peaks and fitted overlapping peaks

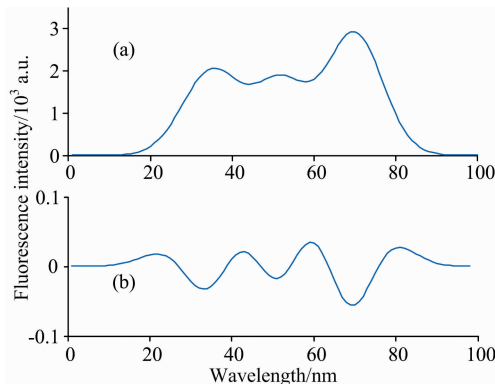


图 8 三峰重叠峰及其二次微分曲线

Fig. 8 Three overlapping peaks and its quadratic differential curves

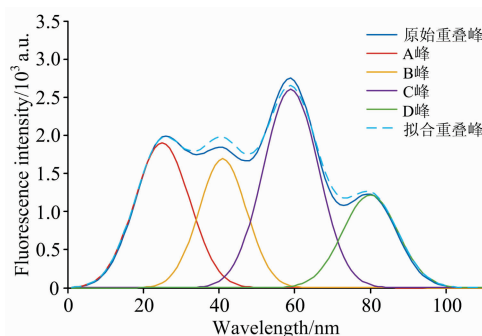


图 10 原始重叠峰、分离单峰及拟合重叠峰  
Fig. 10 Original spectrum, decomposed peaks and fitted overlapping peaks

利用 GA\_BP 算法对其进行分解, 此时调整交叉概率大小为  $[0, 3]$  之间, 其余参数保持不变。随着迭代次数的增加, 适应度值变化逐渐趋于平稳, 迭代 50 次算法结束后最优适应度值为 0.272 9。表 2 为原始重叠峰与分解后各峰的特征值对比, 图 9 为原始重叠峰、分解后各单峰及单峰叠加后的曲线。可见, 使用 GA\_BP 算法对三峰叠加而成的重叠峰进行分解时, 仍具有较高精度。

表 2 三峰重叠解析结果

Table 2 Analysis results at three overlapping peaks

	真实值	预测值	相对误差/%
a1	2.0	2.004 9	0.25
a2	1.7	1.710 9	0.64
a3	2.9	2.897 4	-0.09
c1	100.0	102.395 2	2.40
c2	84.5	87.734 3	3.83
c3	93.8	97.093 4	3.51

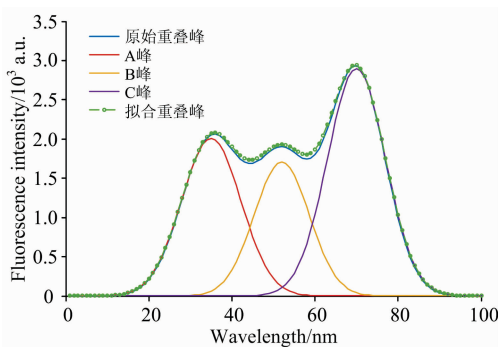


图 9 原始重叠峰、分离单峰及拟合重叠峰  
Fig. 9 Original spectrum, decomposed peaks and fitted overlapping peaks

当重叠峰由四个及以上高斯峰叠加而成时, 由于样本数据增多, 相应增大了种群规模数量为 50。设置最大迭代数为 100, 分峰结果如图 10 所示, 分解精度较前稍差。由于遗传算法具有随机性, 算法每次运行的结果不全相同。

### 3.3 对比分析

为了验证 GA\_BP 算法的准确性, 使用同样的样本数据, 与未经优化的 BP 神经网络作对比, 分析了两种网络模型在训练过程中的收敛度和误差大小。两种神经网络结构参数均相同。

计算网络均方差值 (mean square error, MSE), 将其作为网络训练性能指标, 来对比训练过程中两种模型的误差情

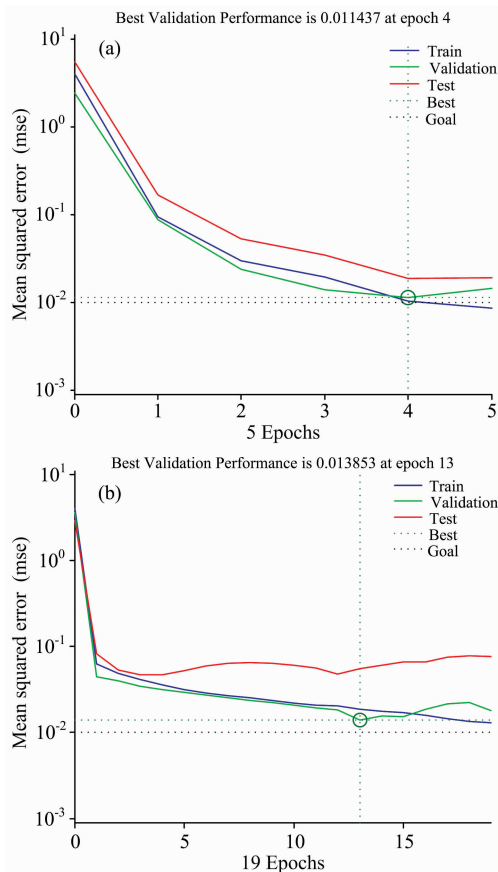


图 11 神经网络收敛度对比

(a): GA\_BP 神经网络; (b): 未经优化的 BP 神经网络

Fig. 11 Contrast of convergence of neural networks

(a): GA\_BP neural network;

(b): Unoptimized BP neural network

况。将两种模型的误差目标均设为 0.01, 下图中分别表示了两种网络模型的网络训练过程中的误差变化。其中图 11(a) 表示 GA\_BP 神经网络, 图 11(b) 表示未经优化的 BP 神经网络。分析可知, GA\_BP 神经网络运行 5 步后, 达到了预设的误差值目标; 而未优化的 BP 神经网络需经过 19 步后方可达到。可以看出 GA\_BP 神经网络模型收敛更快, 误差较小。

## 4 结 论

为了分析多色荧光发射光谱产生的重叠峰, 采用遗传算法优化神经网络(GA\_BP)算法, 利用遗传算法全局搜索寻优的特点及神经网络良好的拟合和预测能力, 对 BP 神经网络

初始权值和阈值进行寻优计算, 确定网络最优参数并进行相应网络训练, 最终分解出独立谱峰。使用此方法进行分析时充分考虑到遗传算法的随机性, 对初始种群及种群规模进行了相应的选取及调整, 确保算法搜索到最大概率意义下的最优网络参数, 提高了算法收敛能力。结果表明, GA\_BP 网络模型预测得到的单峰参数与真实值较为接近, 双峰重叠情况下最大相对误差为 3.57%, 三峰重叠情况下最大相对误差为 3.83%, 具有较高的准确度。同时将 GA\_BP 神经网络模型与未经优化的 BP 神经网络模型作对比, 其中 GA\_BP 网络模型运行 5 步后即可达到预设误差目标, 进一步验证了该算法的收敛度及准确性, 故可使用该算法对重叠光谱进行准确的定量分析。

## References

- [ 1 ] YE Guo-yang, XU Ke-jun(叶国阳, 徐科军). Chinese Journal of Scientific Instrument(仪器仪表学报), 2015, 36(2): 439.
- [ 2 ] FENG Fei, WANG Fu-bei, XIE Fei, et al(冯 飞, 王府北, 谢 非, 等). Acta Photonica Sinica(光子学报), 2015, 44(6): 113.
- [ 3 ] HU Ruo-mu, WANG Zhen-zhen, LIU Ren-wei, et al(胡若木, 王珍珍, 刘人玮, 等). Acta Photonica Sinica(光子学报), 2018, 47(8): 39.
- [ 4 ] XU Xi-rong, PAN Zi-qi, LI Xing-hua, et al(徐喜荣, 潘子琦, 李兴华, 等). Journal of Dalian University of Technology(大连理工大学学报), 2019, 59(1): 97.
- [ 5 ] TAO Wei-liang, LIU Yan, WANG Xian-pei, et al(陶维亮, 刘 艳, 王先培, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2017, 37(12): 3664.
- [ 6 ] Du Hua, Chen Wuhui, Zhu Qingjun, et al. Applied Radiation and Isotopes, 2018, 141: 149.
- [ 7 ] Li Lili, Wu Yanling, Dong Xuehu, et al. Analysis of Overlapping Peaks of GA Algorithms with Early Experience in Integration. ICIEA, 2018.
- [ 8 ] LIU Guang-da, WEI Xing, ZHANG Shang, et al(刘光达, 魏 星, 张 尚, 等). Journal of Biomedical Engineering(生物医学工程杂志), 2019, 36(1): 24.
- [ 9 ] Gao Yuyang, Qu Chao, Zhang Kequan. Energies, 2016, 9(10): 1.
- [ 10 ] HU Yao-gai, ZHANG Xiao-xing, ZHAO Zheng-yu, et al(胡耀垓, 张晓星, 赵正予, 等). Journal of Chongqing University(重庆大学学报), 2012, 35(5): 76.

# Overlapping Spectral Analysis Based on Genetic Algorithms and BP Neural Networks

DU Yue, MENG Xiao-chen\*, ZHU Lian-qing

Beijing Laboratory for Optoelectronic Measurement Technology, Beijing Laboratory for Biomedical Testing Technology and Instruments, Beijing Information Science and Technology University, Beijing 100101, China

**Abstract** With the rapid development of spectroscopy and fluorescence detection technology, monochrome fluorescence labeling is unable to analyze cell samples accurately and has been gradually replaced by two-color or multi-color fluorescence labeling. In the multicolor fluorescence analysis, since the cells were labeled with a variety of fluorescein usually, partial spectral overlap will occur in the emission spectrum, which need to be decomposed into an independent spectral peak to analyze accurately. Aiming at it, optimized BP neural network based on genetic algorithm (GA\_BP) were used for overlapping spectral peak analysis. Firstly, the concrete structure of BP neural network was determined, and the overlapping peak was pre-processed by quadratic differential to find out the number and positions of single peaks as the characteristic value of overlapping peaks to be the input layer of BP neural network; in addition, weights and thresholds of BP neural network were Initialized, and optimal parameters like initial population and population size of the genetic algorithm were selected by using the advantage of global search; after a series of genetic evolution operations like selecting, crossing and mutating, the individuals containing the optimal weights and thresholds of BP neural network were obtained; and then the optimal parameters of the network were selected to carry out network training,

which the width and intensity of the independent peak can be calculated from the output node of the optimized BP neural network; finally, combined with the eigenvalues of overlapping peak identified by quadratic differential, independent spectral peak can be separated. The randomly generated Gaussian overlapping peaks model was used as experimental simulation data, and the decomposition experiments showed high precision of the peak intensity and peak width. Wherein, the maximum relative error of decomposition of two overlapping peaks was 0.30% and 3.57%, and which of the three overlapping peaks was 0.64% and 3.83%. It can also be decomposed when the four overlapping peaks. Moreover, compared the GA\_BP network model with the unoptimized BP neural network model, the results showed that the GA\_BP network could reach the preset error value after five steps, while the unoptimized network model takes 19 steps. This further proves that the GA\_BP network model converged faster with a fairly high precision that can be widely used for the decomposition of spectral and other overlapping peaks, which has a certain practical value compared with traditional methods.

**Keywords** Biomedicine; Overlapping peaks decomposition; Neural networks; Genetic algorithm

(Received Jun. 23, 2019; accepted Oct. 10, 2019)

\* Corresponding author