

基于 OPO 脉冲激光激发光声光谱的真假血液分类鉴别

任重^{1, 2*}, 刘涛¹, 刘国栋^{1, 2}

1. 江西科技师范大学光电子与通信省级重点实验室, 江西 南昌 330038
2. 江西科技师范大学南昌市光电检测与信息处理重点实验室, 江西 南昌 330038

摘要 为了实现快速准确并可回收再利用地鉴别真血和假血, 采用光声光谱技术构建了一套血液光声检测系统并获取血样的光声信号。选取三种动物真血(马血、牛血和兔血)和两种假血(道具假血和红墨汁)共125组血样作为实验样本。获取了700~1064 nm波段内所有样本的光声信号和光声峰峰值谱。实验表明, 真血和假血的光声信号幅度、轮廓、峰值时间点和光声峰峰值均存在差异。为了实现高准确度的真假血液分类识别, 采用了遗传优化的小波神经网络(WNN-GA)算法, 对全波段100组样本进行训练, 并构建了类Morlet小波基函数, 然后对25组测试血样进行分类识别。利用遗传算法对WNN网络的权值、阈值和小波基函数平移、伸缩因子进行了优化, 同时通过调节两个学习率因子, 将真假血液的分类识别率提高了24%。采用主成分分析(PCA)对全波段血样光声峰峰值进行特征提取, 再利用WNN-GA算法进行训练和分类识别。结果表明, 在主成分个数为6时, PCA-WNN-GA融合算法可以使真假血液的分类识别率提高到100%。与另外6种分类识别算法相对比, 该融合算法的识别准确率明显占优。光声光谱技术联合PCA-WNN-GA算法, 可以准确地实现真假血液的分类鉴别。

关键词 光声光谱; 脉冲激光; 分类识别; 血液

中图分类号: O433.4 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2021)09-2734-08

引言

血液是生命系统最重要的组织之一, 应用十分广泛。在生物医学领域, 血液可用于缺血性输血、凝血酶和血液疾病等方面的治疗。在刑侦领域, 血液可以作为重要物证。同时, 血液在动物检验检疫和食品等领域也具有十分重要的应用价值。由于血液的重要性和高利润等因素, 目前出现了利用动物血充当人血制造和扰乱犯罪现场, 在可食用血液制品中也有不法分子利用假冒伪劣血来冒充的报道, 给社会治安、食品安全和医疗卫生等带来了不良影响, 对真血和假血的准确鉴别意义重大。传统鉴别真假血液的方法有感官法、液相/气相色谱法^[1]、生化分析法^[2]和电化学传感法等。由于假血颜色与真血类似, 从视觉、嗅觉等来进行判别会存在误差。液相/气相色谱法、生化分析法和电化学传感法均需要进行复杂的样品预处理, 不仅检测周期相对较长, 而且血液样品

无法回收再利用, 同时均需要专业人员来操作。目前, 已有利用光学手段进行真假血检测的方法, 如: 近红外光谱法、拉曼光谱法、荧光光谱法。有研究利用近红外光谱法对牛羊狗等常见动物全血样品进行了透射光谱测量, 利用支持向量机等算法实现了动物血的分类。Gregory^[3]使用拉曼光谱和化学计量学模型对人血和动物血溯源进行了鉴别。白鹏利等^[4]使用拉曼光谱结合主成分分析方法, 对三种不同动物血和21个人血进行了识别, 结果表明人与动物血样的识别率均高于95%, 动物间的血样识别率可达90%。有报道利用荧光光谱法结合BP神经网络实现了对不同动物血的荧光光谱特征进行提取和识别分类。虽然已有利用纯光谱法来进行血液分类识别的研究报道, 但是, 由于不同动物之间血液光谱数据差异性较小, 且在某些波段会存在不同成分的光谱出现交叠, 导致有效的信号分离和特征提取较难。另外, 光谱采集受环境光照等因素影响, 使得识别准确度有待进一步提高。

收稿日期: 2020-09-25, **修订日期:** 2021-01-20

基金项目: 国家自然科学基金项目(61650402, 51763011), 江西省科技创新杰出青年人才项目(20192BCBL23015), 江西省自然科学基金杰出青年基金项目(2018ACB21006), 南昌市重点实验室(2019-NCZDSY-008), 青年拔尖人才自然科学基金项目(2014QNBjRC004)和校博士启动基金项目(2017BSQD021)资助

作者简介: 任重, 1981年生, 江西科技师范大学光电子与通信重点实验室副教授 e-mail: renzhong0921@163.com

* 通讯作者 e-mail: renzhong0921@163.com

光声技术是基于光致超声机理，采集的信号为超声机械波，从原理上可以较好地避免散射光和光谱混叠的问题。虽然光声技术已经应用于血液、血糖^[5-6]等生物医疗领域，但是光声技术应用于真假血液的分类鉴别却未见研究报道。鉴于光声技术的优势，可提供一种真假血液分类鉴别新方法，本工作采用光声光谱技术对真假血液进行分类识别，以此克服纯光谱法的不足，提高真假血液分类和鉴别准确率。同时，基于采集得到的五种真假血样的光声峰峰值谱，通过研究遗传优化的 WNN 融合 PCA 算法，进一步提高真假血液分类鉴别的准确度。

1 理论

1.1 光声理论

光声技术的基本原理是光致超声效应和热膨胀机制。当一束短脉冲激光以一定的重复频率和脉冲持续时间照射某物质时，由于对入射光的选择性吸收，被照射区域能量沉积(释放)而产生快速升温(降温)，这种温度变化使得物质局部区域产生快速的热膨胀或收缩，进而引发超声机械波，即：光声信号。一般光声信号产生机理可以用波动方程^[7]来进行描述，但由于目前很难得到其具体解析解，通常采用半经验公式^[8]描述光声信号，即

$$p = k \frac{E\mu_a\beta v^n}{C_p r^m R^l} \quad (1)$$

式(1)中， E 为脉冲激光输出能量， μ_a 为被测物质的吸光系数， v 为声速， R 为光源半径， r 为超声探头与光声源的距离， β 为物质热膨胀系数， C_p 为比热容， k 为比例系数。式中 n, m, l 为参数，其满足如下范围： $0.5 < n < 2, 0 < m < 1, 0 < l < 2$ 。从式(1)中可知，光声信号强度与脉冲激光能量、吸光系数、热膨胀系数、比热容、声速、光斑大小和探测距离等多个因素有关。不同物质由于其成分和含量不同，使得物质间的固有属性(光学属性 μ_a 、热学属性 β/C_p 和声学属性 v)存在差异，最终使得光声信号在强度和形态上也存在差异性。通过对光声数据的反演，就可以对被测物质进行定性识别和定量分析等。

1.2 基于遗传优化的小波神经网络

选取小波神经网络(wavelet neural network, WNN)^[9]对真假血液进行分类识别。WNN 由 3 层网络构成，如图 1 所示。

其中输入层有 $m(m=1, 2, \dots, M)$ 个神经元，给出 p 组输入的光声峰峰值数据为 x_m^p ，隐含层有 $k(k=1, 2, \dots, K)$ 个神经元，输出层有 $n(n=1, 2, \dots, N)$ 个神经元。隐含层的神经元激励函数是在 Morlet 小波基函数基础上构造的，即

$$h_k(x) = e^{-\frac{x^2}{2}} \cos(1.75x) \quad (2)$$

WNN 的隐含层输出 O_k^p 为

$$O_k^p = h_k \left(\frac{\sum_{m=1}^M \omega_{km} x_m^p - b_k}{a_k} \right) \quad (3)$$

式(3)中， ω_{km} 为输入层和隐含层的之间的权值， b_k 为小波基函数的平移因子， a_k 为小波基函数的伸缩因子。 m 为输入层节点，下标 k 为隐含层节点。

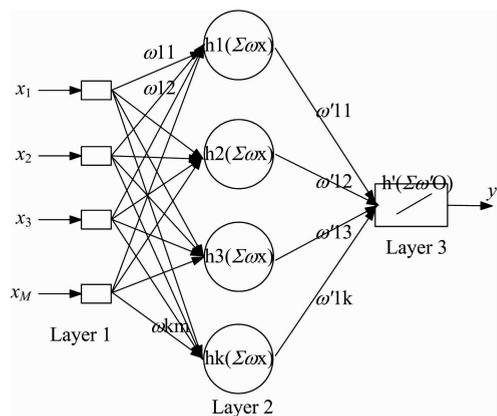


图 1 WNN 网络结构

Fig. 1 Structure of WNN network

WNN 输出层的输出 y_n^p 为

$$y_{pn} = h' \left(\sum_{k=1}^K \omega'_{nk} O_m^p \right) \quad (4)$$

式(4)中， ω'_{nk} 为隐含层与输出层之间的权值。而激励函数 $h'(\cdot)$ ，选择线性函数来作为输出层的激励函数。

WNN 网络的实际输出为 $y(n) = y_n^p$ ，假设网络的期望输出为 d_n^p ，则目标误差函数为

$$E_n^p = \sum_{p=1}^P E^p = \frac{1}{2P} \sum_{p=1}^P \sum_{n=1}^N (d_n^p - y_n^p)^2 \quad (5)$$

式(5)中， d_n^p 为输出层第 n 个结点的期望输出， y_{pn} 为网络实际输出。

采用误差反向传播的方式对权值 ω_{km} 和 ω_{nk} ，以及小波基函数的平移因子 b_k 和伸缩因子 a_k 按照如下方式进行修正，即

$$\begin{cases} \omega_{nk}^{\text{new}} = \omega_{nk}^{\text{old}} + \eta \sum_{k=1}^K \frac{\partial E_n^p}{\partial \omega_{nk}} \\ \omega_{km}^{\text{new}} = \omega_{km}^{\text{old}} + \eta \sum_{k=1}^K \frac{\partial E_n^p}{\partial \omega_{km}} \\ a_k^{\text{new}} = a_k^{\text{old}} + \lambda \sum_{m=1}^M \frac{\partial E_n^p}{\partial a_k} \\ b_k^{\text{new}} = b_k^{\text{old}} + \lambda \sum_{m=1}^M \frac{\partial E_n^p}{\partial b_k} \end{cases} \quad (6)$$

式(6)中， ω_{nk}^{new} ， ω_{km}^{new} ， a_k^{new} 和 b_k^{new} 分别为调整后的权值、小波基函数平移因子和伸缩因子。 ω_{nk}^{old} ， ω_{km}^{old} ， a_k^{old} 和 b_k^{old} 分别为调整前的权值、小波基函数平移因子和伸缩因子。 η 和 λ 为学习率因子。

为了提高 WNN 神经网络对真假血液的识别率，采用遗传算法(genetic algorithm, GA)^[10]对 WNN 的初始权值(ω_{km} ， ω'_{nk})和阈值、小波基函数平移因子 b_k 和伸缩因子 a_k 进行优化。GA 算法流程如图 2 所示。

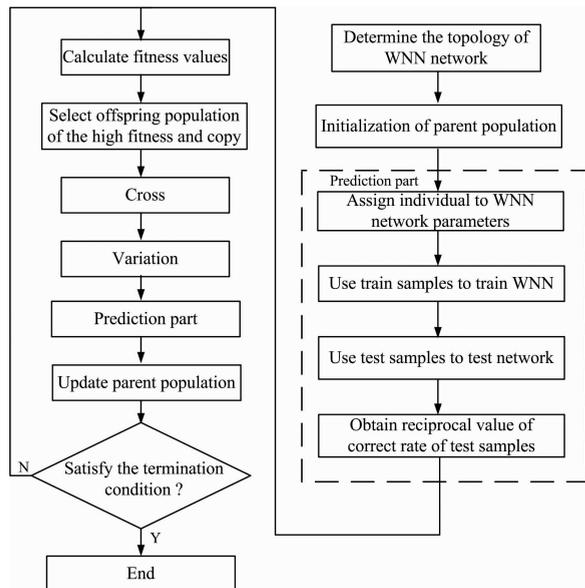


图 2 GA 算法流程图

Fig. 2 Flow chart of GA algorithm

2 实验部分

2.1 装置

构建了一套血液光声检测系统,如图 3 所示。系统中采用波长可调谐 Nd:YAG 532 nm 泵浦源光学参数振荡(optical parametric oscillator, OPO)脉冲激光器(OPOletteTM, 532 II, OPOTEK Inc., USA)作为光源,其重复频率 20 Hz,脉冲持续时间 7 ns,最大输出能量 2 mJ。经过反射镜、光阑和聚焦透镜后将聚焦光斑落入到透明硅胶软管中的血液样品中,激发产生的光声信号,由中心响应频率为 2.5 MHz 的聚焦超声探测器(Doppler Co., China)捕获转换为模拟电压信号后,先经过信号放大器(ATA-5620, Aigtek Co., China)进行 60 dB 增益的信号放大,再由低通滤波器(BLP-7-75+, Mini-Circuits, USA)去除高频噪声信号干扰,然后由带宽为 500 MHz 的双踪数字示波器(54642D, Agilent, USA)进行采集并转换为数字信号,通过 GPIB 接口卡(GPIB-USB-HS, NI, USA)将数据传输至计算机保存和分析处理。另外,构建了一套血液循环子系统便于实验更换血液样品,由小型水

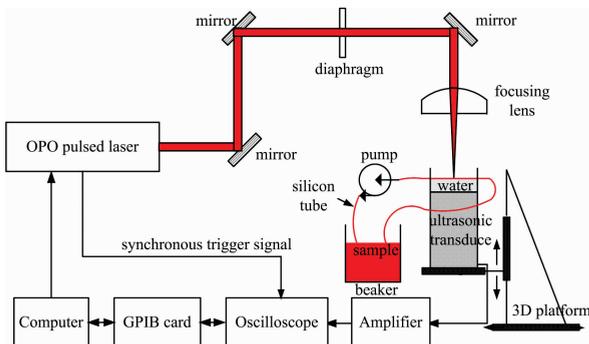


图 3 实验装置图

Fig. 3 Diagram of experimental set-up

泵、硅胶软管和烧杯构成。实验前,OPO 脉冲激光器预热约 30 min。实验中保持室温 20 °C 恒定。

2.2 材料

采用了 5 种不同的血样,即:三种动物真血(马血、牛血和兔血)和两种假血(道具假血和红墨汁)来作为实验真血和假血样品,如图 4 所示。其中马血、牛血和兔血由郑州益康生物工程有限公司提供,为了防止动物真血凝结,在出厂前已按照 $1.6 \text{ mg} \cdot \text{mL}^{-1}$ 比例添加了乙二胺四乙酸(ethylenediamine tetraacetic acid, EDTA)抗凝血剂。实验血样总样本数 125 个,即:每种血样校正集样本 20 个,测试集样本 5 个。每次实验,将循环子系统烧杯和硅胶软管用医用酒精消毒,然后再用开水煮沸进行二次消毒,以防止血液在实验过程中的污染。



图 4 实验样品

Fig. 4 Experimental samples

- (a): Horse blood; (b): Cow blood; (c): Rabbit blood;
(d): Props fake blood; (e): Red ink

3 结果与讨论

3.1 光声检测结果及分析

使用 2.1 节实验系统,对所有 125 组真假血液样品进行了光声检测。实验中,将约 10 mL 的血样装入烧杯中,开启血液循环子系统模拟血液流动,并将测试血样导入到硅胶软管中。开启 OPO 脉冲激光器激光输出,经过数字示波器将数据采集平均 128 次,得到了在激发波长 740 nm 下的五种不同血样的实时光声信号如图 5 所示。

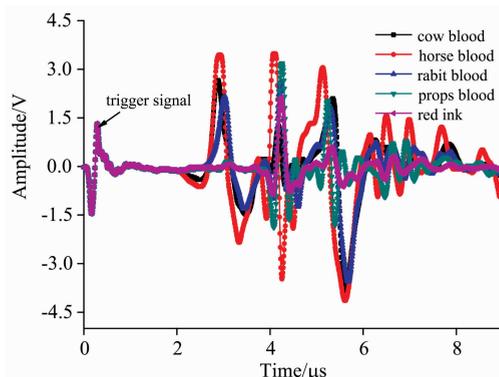


图 5 740 nm 波长下不同血样的时间分辨光声信号

Fig. 5 Time-resolved photoacoustic signals of different blood samples at 740 nm

从图 5 中可知,动物真血与假血的时间分辨光声信号,不仅动物真血的光声信号幅值大于假血,而且在信号轮廓上

也存在差异。三种动物真血的时间分辨光声信号分别在 3, 4.2 和 5.2 μs 附近有明显的波峰信号, 且幅值较大。而对于假血而言, 在 3 μs 处的波峰信号与另外两个时间点的波峰信号相比较不明显, 且整体幅值均小于真血。另外, 从峰值的时间位置上来看, 不同血样的光声信号峰值的时间位置均不同, 对于三种动物真血而言, 牛血的峰值时间点最小, 依次是马血和兔血。而假血的峰值时间均大于真血。

为了分析不同血样在不同激发波长下的光声信号强度变化情况, 得到了在 700~1 064 nm 波段下的五种不同血样光声峰峰值。将每种血样的各 25 组全波段的光声峰峰值进行平均, 获得了平均光声峰峰值谱, 如图 6 所示。

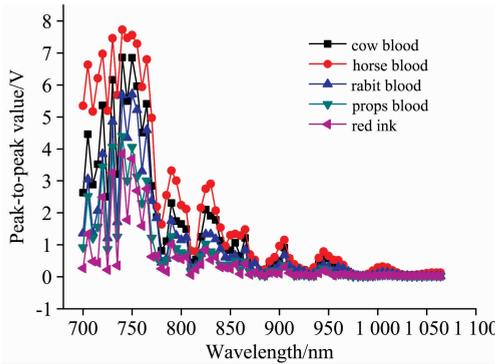


图 6 五种不同血样在 700~1 064 nm 波长范围内的平均光声峰峰值谱

Fig. 6 Photoacoustic peak-to-peak mean spectral of five different blood samples at 700~1 064 nm

从图 6 中可知, 在 700~1 064 nm, 不同血样的峰峰值变化趋势大致相同, 且峰峰值较大的波长均出现在 740, 790, 830, 865, 905, 945 和 1 010 nm 处。但是, 不同血样的光声峰峰值却存在差异。相比而言, 在 700~875 nm 内, 真血的光声峰峰值比假血的大, 而在 875~1 064 nm, 除了在 905, 950 和 1 000 nm 三个波长附近, 其他波长的光声峰峰值的差别不明显。

3.2 血液分类识别结果及分析

3.2.1 基于 WNN-GA 分类识别结果

先后采用了 WNN 和 WNN-GA 算法对 5 种血样进行分类识别, 训练样本为 100 组, 测试样本为 25 组。将 700~1 064 nm 全波段的真假血的光声峰峰值谱作为 WNN 网络的输入数据, 输入神经元个数为 74。对于隐含层的神经元个数, 根据公式 $\sqrt{m+n}+a^{[11]}$ 来确定, 其中 m 为输入层神经元个数, n 为输出层神经元个数, 即: 1, a 取值范围为 $[0, 10]$, 故隐含层神经元个数的取值范围为 $[8, 18]$ 。利用 GA 算法对 WNN 网络的权值和阈值、小波基平移因子和伸缩因子进行了优化。通过调整两个学习率因子, 在学习率因子 $\eta=0.001$, $\gamma=0.06$ 的情况下, 对比了 WNN 和 WNN-GA 两种算法下, 不同隐含层神经元个数对真假血液的分类识别准确率, 如图 7(a) 所示。

在图 7(a) 中, 对于 WNN 而言, 隐含层神经元个数为 13 时, 其对真假血液的分类识别准确率最优为 64%; 而对于

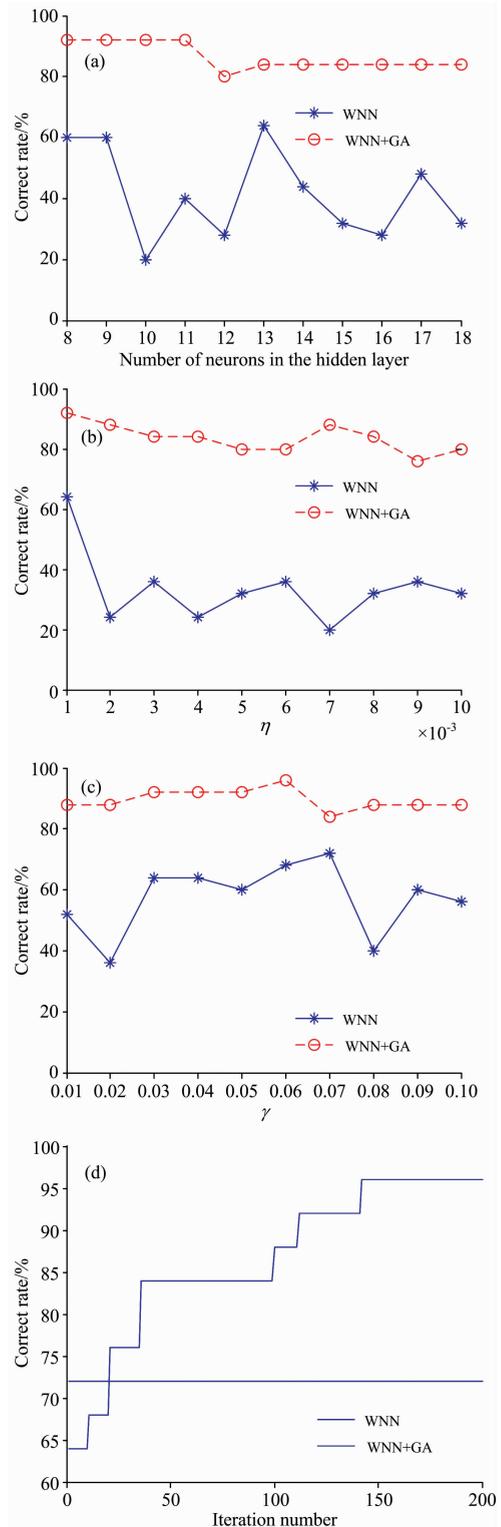


图 7 WNN 和 WNN-GA 各参数对真假血液分类识别的影响

(a): 隐含层神经元个数的影响; (b): 学习率因子 η 的影响; (c): 学习率因子 γ 的影响; (d): 迭代次数的影响

Fig. 7 Effects of parameters on the classification and identification of real or fake blood for WNN and WNN-GA

(a): Effect of neuron numbers in hidden layer; (b): Effect of learning factor η ; (c): Effect of learning factor γ ; (d): Effect of iteration number

WNN-GA, 隐含层神经元个数为 8—11 时, 对真假血液的分类识别准确率最优可达 92%。

学习率因子 η 和 γ 对 WNN 和 WNN-GA 两者对真假血液分类识别率的影响, 分别如图 7(b) 和 (c) 所示。从 7(b) 可知, 对于 WNN 和 WNN-GA 而言, 当学习率因子 $\eta=0.001$ 时, 两者的分类识别准确率均最高。随着学习率因子 η 增大, 两者的识别准确率均呈下降趋势, 但是 WNN 识别准确率振荡不稳定现象较严重。对于学习率因子 γ , WNN 的识别准确率依然出现不稳定振荡现象, 当 $\gamma=0.02$ 时, WNN 的识别准确率只有 36%, 当 $\gamma=0.07$ 时, WNN 的识别准确率最大为 72%。而对于 WNN-GA, 其识别准确率变化趋势较平坦, 且当 $\gamma=0.06$ 时, WNN-GA 对真假血液的分类识别准确率最优达到 96%。

对于 WNN-GA, 利用 GA 对 WNN 的参数进行迭代优化。图 7(d) 为 WNN-GA 算法在不同迭代次数时的真假血液分类识别准确率。从图中可知, 随着迭代次数的增加, WNN-GA 对真假血液的分类识别准确率逐渐提高, 当迭代次数达到 150 时, WNN-GA 的识别准确率达到 96%。而对于 WNN 而言, 获得的识别准确率最大只有 72%。

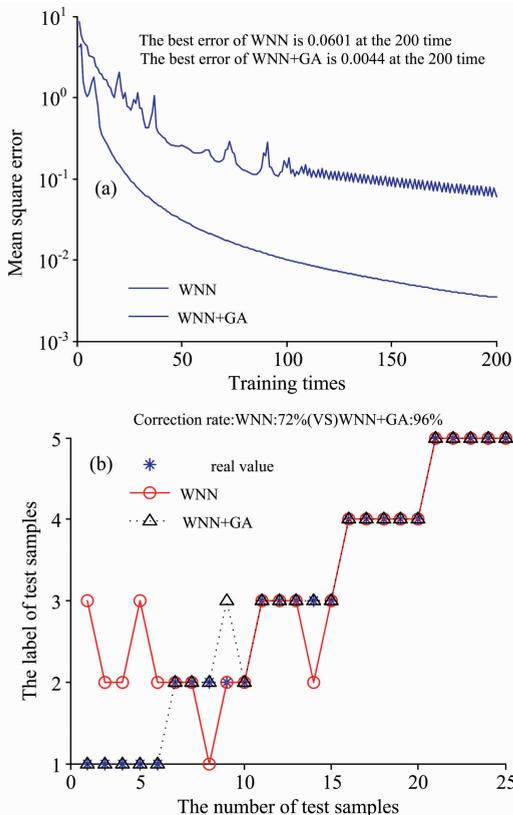


图 8 WNN 和 WNN-GA 对真假测试血样的分类识别结果

(a): 训练次数对测试血样均方误差的影响;

(b) 测试血液的分类识别结果

Fig. 8 Results of classification and identification of real or fake blood for WNN and WNN-GA

(a): Effect of training time on the mean square error;

(b): Results of classification and identification of test bloods

图 8(a) 为 WNN 和 WNN-GA 不同训练次数对 25 组测试血样识别结果的均方误差。从图中可知, 随着训练次数的增加, 两者对测试血样的均方误差均逐渐减小。但是, WNN 出现了振荡不稳定现象, 而对于 WNN-GA, 在训练次数超过 10 之后, 其识别均方误差呈平稳下降趋势。且 WNN-GA 的均方误差均明显小于 WNN 的均方误差。图 8(b) 给出了 WNN 和 WNN-GA 对 25 组测试血样的分类识别结果。由于 WNN 和 WNN-GA 均为监督学习分类识别算法, 利用数字 1, 2, 3, 4 和 5 来对马血、牛血、兔血、道具假血和红墨汁进行标记。从图 8(b) 可知, WNN 对两种假血的分类识别较准确, 而对三种真血的鉴别准确度不高。但是, 对于 WNN-GA, 除了牛血有 1 个测试样本的结果误判为兔血之外, 其余血样的分类识别均正确。经过计算, WNN-GA 的分类识别率为 96%。因此, 与 WNN 相比, WNN-GA 大大提高了真假血液的分类识别率。

3.2.2 基于 PCA-WNN-GA 分类识别结果

由于血样的全波段光声峰峰值, 不仅包含了具有区分度的特征波长下的光声信息, 还包含了不具备区分度波长下的信息, 使得真假血液的分类识别准确度受到影响。同时, 输入血样的数据维数较大, 会造成 WNN 内部参数过多且计算复杂度高。为了进一步提高 WNN-GA 对真假血样的分类识别准确率, 采用主成分分析 (principle components analysis, PCA) 算法^[12] 提取真假血液样本的主要特征信息。

PCA 的基本原理是数据降维, 即: 将原数据进行某个维度的投影, 再用较少的变量信息来代替原始众多变量的信息, 而又最大限度地不丢失原先的特征信息, 即

$$\begin{cases} Z_1 = k_{11}x_1 + k_{12}x_2 + \dots + k_{1p}x_p \\ Z_2 = k_{21}x_1 + k_{22}x_2 + \dots + k_{2p}x_p \\ \dots \\ Z_m = k_{m1}x_1 + k_{m2}x_2 + \dots + k_{mp}x_p \end{cases} \quad (7)$$

式(7)中, x_1, x_2, \dots, x_p 为原始变量 (下标 p 表示变量的个数), Z_1, Z_2, \dots, Z_m 为新变量, 即主成分 (下标 m 表示新变量的个数), k_{mp} 为载荷系数, 且 $m \leq p$ 。

根据特征值累计贡献率大于 80%, 分别选取了输入样本的 2, 4, 6 和 8 个主成分, 再利用 WNN-GA 算法进行识别。根据隐含层神经元个数公式可以知道 PCA 降维后的隐含层神经元范围为 [2, 13]。当学习率因子 $\eta=0.001$, $\gamma=0.04$ 时, 获得了在不同隐含层神经元个数情况下, 不同主成分对真假血液的识别准确率, 如图 9(a) 所示。

从图 9(a) 可知, 当 2 个主成分时, 神经元个数为 8 所对应的识别准确率只有 84%。当 4 个主成分时, 神经元个数为 5 所对应的识别准确率为 92%。随着主成分个数增大到 6, 神经元个数为 11 时, 其识别准确率可达 100%。当继续增加主成分个数时, 最优识别准确率依然为 100%。图 9(b) 和 (c) 分别给出了不同主成分, 学习率因子 η 和 γ 在不同取值对真假血液的分类识别准确率的影响。从图中可知, 随着主成分个数的增加, 识别准确率逐渐增大; 当 $\eta=0.001$ 时, 对于所有的主成分均得到最优识别准确率, 且对于 6 个主成分而言, $\eta=0.001$ 和 $\gamma=0.04$ 所对应的分类识别准确率可达 100%。图 9(d) 为 PC-WNN-GA 算法中不同主成分下, 遗传迭代次

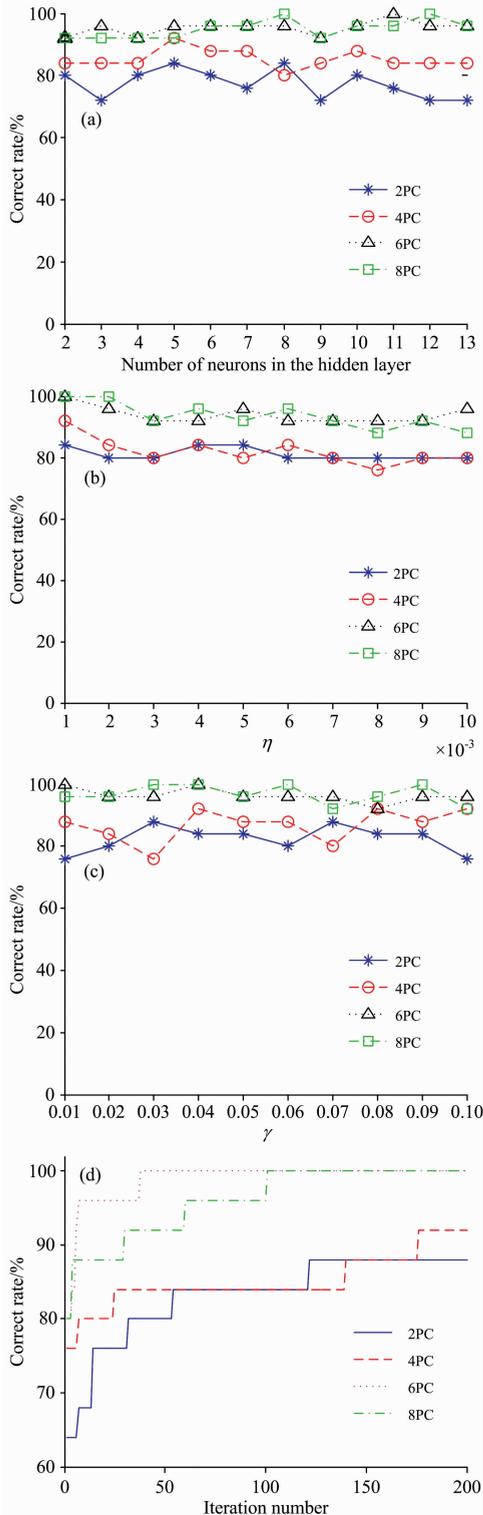


图 9 PCA-WNN-GA 算法中各参数对真假血液分类识别的影响

(a): 隐含层神经元个数的影响; (b): 学习率因子 η 的影响;
(c): 学习率因子 γ 的影响; (d): 迭代次数的影响

Fig. 9 Effects of parameters on the classification and identification of real or fake blood for PCA-WNN-GA

(a): Effect of neuron numbers in hidden layer; (b): Effect of learning factor η ; (c): Effect of learning factor γ ; (d): Effect of iteration number

数对真假血分类识别准确率的影响。从图 9(d)可知, 对于所有主成分而言, 随着遗传迭代次数的增加, 真假血的分类识别准确率逐渐提高; 并且, 随着主成分的增加, 对真假血的分类识别准确率也有较大提高, 当主成分个数增加到 6 时最终可以获得 100% 的识别准确率。

图 10(a)给出了 PCA-WNN-GA 算法, 在四个不同主成分个数下, 样本训练次数对测试血样分类识别均方误差的影响曲线。从图中可知, 随着样本训练次数的增加, 对测试血样的分类识别均方误差均逐渐减小。并且, 分类识别均方误差随着主成分个数的增加也减小了。经过计算当样本训练次数为 200 次时, 主成分个数为 6 所对应的分类识别均方误差约为 0.004 2。

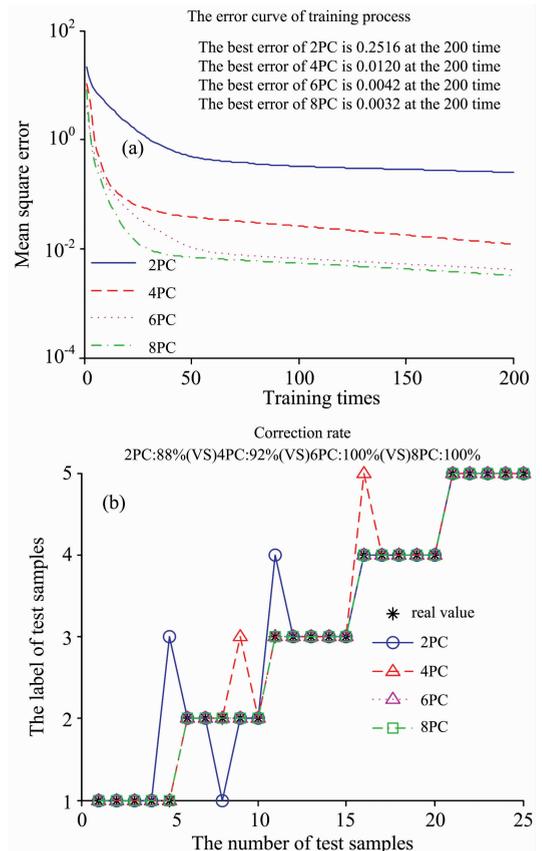


图 10 不同主成分个数下 PCA-WNN-GA 对真假测试血样的分类识别结果

(a): 训练次数对测试血样均方误差的影响;
(b): 测试血液的分类识别结果

Fig. 10 Results of classification and identification of real or fake blood for PCA-WNN-GA under different principle components

(a): Effect of training time on the mean square error;
(b): Results of classification and identification of test bloods

图 10(b)为 PCA-WNN-GA 算法选取四个不同主成分时, 对五种不同真假测试血样的分类识别结果。从图中可知, 选取 2 个主成分时, 对三个真血的分类识别存在偏差, 此时分类识别准确度只有 88%。当选取 4 个主成分时, 由于

累计贡献率超过了 90%，所能表征区分血样的特征增多，使得对测试血样的分类识别率提高到了 92%。当选取 6 个主成分时，PAC-WNN-GA 算法对测试血样的分类识别率提高到了 100%。

3.2.3 不同算法的分类识别结果对比

最后，对比了 7 种不同算法对 25 个测试血样进行分类识别的效果。其中，选取了 5 种监督型分类识别算法，即：极限学习机(extreme learning machine, ELM)、支持向量机(support vector machine, SVM)、WNN、WNN-GA 和 PCA-WNN-GA，另外选取 2 种无监督型分类识别算法，即：竞争神经网络(competitive neural networks, CNN)和自组织映射神经网络(self organizing map neural network, SOM)。对于 ELM 而言，由于隐含层传递函数采用的是 BP 神经网络的 sigmoid 函数，且由于 ELM 网络隐含层到输出层的权值和阈值参数没有给出，采用的是对隐含层的输出结果和训练样本的标签求伪逆的思想，它的目标误差为 0，故无需进行样本训练。得到的 7 种不同算法对真假血液的分类识别准确率如图 11 所示。

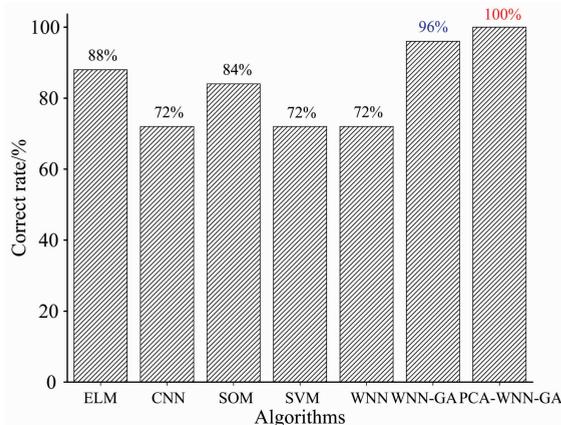


图 11 七种不同算法的真假血液识别率对比

Fig. 11 Correction rate of classification and identification of real or fake bloods for seven different algorithms

从图 11 中可知，单纯的 WNN 分类识别准确度与 CNN 和 SVM 算法的准确度相当，且低于 ELM 和 SOM 算法。其原因是由于在全波段下的血液样本维数较大，且 WNN 识别准确度受初始化权值和阈值、小波基平移因子和伸缩因子，以及学习率因子等影响，使得识别正确率不高。在用 GA 算法对 WNN 网络的权值和阈值、小波基平移因子和伸缩因子进行迭代优化后，使得真假血液分类识别率提高了 24%。并且，在利用 PCA 算法对血样的全波段光声峰峰值提取 6 个主成分的基础上，利用 PCA-WNN-GA 算法，最终可以使得对真假血液的分类识别准确率提高到 100%。

4 结 论

真假血液的快速准确分类识别对维护社会治安、保障食品安全和医疗卫生等方面至关重要。通过利用光声光谱技术联合 PCA-WNN-GA 算法实现了对真假血液的高准确度分类识别，同时对血液样本未造成任何破坏，可以实现回收利用。其中构建了一套血液光声检测系统，获取了五种血样的光声信号和光声峰峰值谱。从光声信号幅度、轮廓、峰值时间点、光声峰峰值等角度宏观对比了三种动物真血和两种假血之间的信号差异。为了实现真假血液的分类识别，通过三层神经网络和类 Morlet 小波基函数构建了 WNN 神经网络，将五种血样的 100 组样本在全波段下的光声峰峰值谱作为输入数据进行训练，为了提高 WNN 网络的识别率，采用 GA 算法对 WNN 网络的权值、阈值和小波基函数的平移因子和伸缩因子进行了优化。同时，鉴于输入的全波段血样的数据维数较大，采用了 PCA 算法提取了所有血样的特征主成分，再利用 WNN-GA 算法来对提取的血样主成分进行样本训练和测试血样的分类识别，结果表明，当选取主成分个数为 6 时，PCA-WNN-GA 算法对真假血液的分类识别准确率可以达到 100%。因此，光声光谱技术联合 PCA-WNN-GA 算法在真假血液的分类鉴别方面有较好的应用价值。

References

- [1] Archana N A N, Rita C, Manjunath N, et al. *Clinica Chimica Acta*, 2018, 485: 305.
- [2] Xu G., Hu B., Chen G., et al. *Biological Trace Element Research*, 2015, 164(2): 192.
- [3] Mclaughlin G, Doty K C, Lednev I K. *Forensics International*, 2014, 238(5): 91.
- [4] BAI Peng-li, WANG Jun, YIN Huan-cai, et al(白鹏利, 王 均, 尹焕才, 等). *The Journal of Light Scattering(光散射学报)*, 2016, 28(2): 163.
- [5] Ren Z, Liu G D, Huang Z, et al. *International Journal of Optomechanics*, 2015, 9(3): 221.
- [6] LÜ Peng-fei, LU Zhi-qian, HE Qiao-zhi, et al(吕鹏飞, 陆志谦, 何巧芝, 等). *Optics and Precision Engineering(光学精密工程)*, 2019, 27(6): 1301.
- [7] Mohamed F, Ali H. *Discrete & Continuous Dynamical Systems*, 2017, 22(2): 491.
- [8] Ren Z, Liu G D, Huang Z, et al. *Chinese Optics Letters*, 2013, 11(S21701): 1.
- [9] Juarez-Guerra E, Alarcon-Aquino V, Gomez-Gil P, et al. *Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology*, 2020, 92(2): 187.
- [10] Yu W B, Zhao F Y. *International Journal of Green Energy*, 2019, 16(12): 938.
- [11] CHEN Ming(陈 明). *Matlabneural Network Principle and Example Precise Solution (Matlab)(神经网络原理与实例精解)*. Beijing:

Tsinghua University Press(北京: 清华大学出版社), 2012. 164.

[12] Teymoori F, Asghari G, Mirmiran P, et al. *Scientific Reports*, 2017, 7(1): 16838.

Classification and Identification of Real or Fake Blood Based on OPO Pulsed Laser Induced Photoacoustic Spectroscopy

REN Zhong^{1,2*}, LIU Tao¹, LIU Guo-dong^{1,2}

1. Key Laboratory of Optic-Electronics and Communication, Jiangxi Science and Technology Normal University, Nanchang 330038, China
2. Key Laboratory of Optic-Electronic Detection and Information Processing of Nanchang City, Jiangxi Science and Technology Normal University, Nanchang 330038, China

Abstract In order to rapidly and accurately achieve the identification of the real or fake blood, as well as recycled usage of blood, the photoacoustic spectroscopy was used in this work to establish a set of blood photoacoustic detection systems and to capture the photoacoustic signal of blood samples. Three kinds of animal blood (horse blood, cow blood, and rabbit blood), two kinds of fake blood (props blood and red ink), the total number of blood samples are 125 groups, were used as the experimental samples. The photoacoustic signals and photoacoustic peak-to-peak spectral of all blood samples at 700 ~ 1 064 nm were obtained. Photoacoustic experimental results show that the amplitude, profile, peak-point time, and peak-to-peak values of real and fake blood samples are different. To achieve the classification and identification of the real and fake blood with high precision, we used the wavelet neural network optimized by a genetic algorithm (WNN-GA) to train the 100 groups of samples for five kinds of blood in full wavelengths. Moreover, a kind of Morlet-like wavelet basis function was built. Then, 25 groups of blood samples were tested. Meanwhile, the GA algorithm was used to optimize the weights and thresholds of WNN network and the shift factor and stretch factor of wavelet basis function, and two learn factors can be adjusted. Compared with WNN, the correction rate of classification and identification for real and fake blood based on WNN-GA improved by 24%. Then, the principle components analysis (PCA) algorithm was used to extract the characteristic information of real or fake blood from the photoacoustic peak-to-peak full spectral. After that, the chosen principle components were trained and test by the WNN-GA algorithm. Results show that under 6 principle components, the algorithm of PCA-WNN-GA algorithm improves the correction rate of classification and identification for real and fake blood to 100%. Finally, compared with other the six algorithms, the correction rate of classification and identification for PCA-WNN-GA was superior to others. Therefore, the classification and identification of the real and fake blood can be well achieved via photoacoustic spectroscopy combined with the PCA-WNN-GA algorithm.

Keywords Photoacoustic spectroscopy; Pulsed laser; Classification and identification; Blood

(Received Sep. 25, 2020; accepted Jan. 20, 2021)

* Corresponding author