

# LIF 和 CNN 的矿井突水水源类型判别

杨 勇<sup>1,3</sup>, 岳建华<sup>1\*</sup>, 李 晶<sup>2</sup>, 张河瑞<sup>1</sup>

1. 中国矿业大学资源与地球科学学院, 江苏 徐州 221008
2. 南京审计大学信息工程学院, 江苏 南京 210029
3. 徐州工业职业技术学院, 江苏 徐州 221140

**摘 要** 矿井进入深部开采过程中, 突水威胁分别来自顶板老空水和底板高压岩溶水。煤矿突水水源类型的在线识别能够预警煤矿水害, 是矿井水害防治关键环节, 对煤矿安全生产具有积极意义。代表离子法作为传统的煤矿突水水源类型识别方法, 需要深入现场采集水样, 密封处理后在实验室检测水样中 7 种典型的无机离子浓度, 计算得到突水评价因子。这种存在检测周期过长、样品易被污染以及预警响应滞后、无法在线判别等不利因素。针对代表离子法方法的不足, 提出了一种基于激光诱导荧光(LIF)和卷积神经网络(CNN)的矿井突水水源判别模型。首先, 针对淮南矿业集团新集二矿的 4 种水体, 2016 年 6 月—2017 年 6 月期间分批次取得 161 组水源样本, 其中采空区积水 46 条, 砂岩水 59 条, 太灰水 42 条和奥灰水 14 条。用 LIFS-405 激光诱导荧光系统发射的 405 nm 激光检测水样, 水体受激后得到突水水样的荧光光谱。主成分分析得到前 10 个主成分累计贡献率不足 85%, 4 种水样无法有效直接辨识。针对该问题和水样荧光光谱中的随机高频波动干扰, 采用一阶滞后滤波方法抑制波动频率较高的周期性干扰; 针对线判别分析对数据更新率的要求, 采用递推平均方法; 在此基础上, 提出了一种改进的递推平均一阶滞后平滑滤波方法, 并对滤波处理后的荧光光谱进行自相关计算, 得到二维自相关荧光光谱特征图。实验表明, 采用改进后的滤波法处理方法, 计算得到的 4 种测试水样的二维荧光光谱图较好的滤除了噪声干扰, 并表现了出了明显的差异性。针对二维自相关荧光光谱特征图, 构建了基于卷积神经网络(CNN)的突水水源类型判别模型, 用于判别突水水源类型。该方法采用深度学习的模型框架, 直接对二维自相关荧光光谱特征图进行识别, 有效避免了 PCA 降维的片面性。理论分析和实验结果表明: 该模型对水源类型的准确识别率达到了 98%, 是一种有效的矿井突水水源类型判别方法, 为在线矿井突水水源类型判别方法提供了新的思路。

**关键词** 突水水源类型判别; 激光诱导荧光技术; 卷积神经网络

**中图分类号:** O657.3   **文献标识码:** A   **DOI:** 10.3964/j.issn.1000-0593(2019)08-2425-06

## 引 言

在我国煤矿重特大事故中, 矿井突水事故在死亡人数上和发生次数上, 仅次于煤矿瓦斯事故, 造成的经济损失一直居各类煤矿灾害之首<sup>[1]</sup>。研究表明, 煤矿各含水层具有不同的水化学及物理特征。对井下突水成分进行实时分析, 快速、准确判别突水水源类型, 是突水预警和灾后抢险决策的重要依据。当前矿井水源识别方法主要是代表离子法<sup>[2]</sup>、水温水位<sup>[3]</sup>和 GIS 法等, 最常用的是以水中 7 种无机离子作为评价因子对突水水源判别分析。该方法具有较好的识别效

果, 但取样工作量较大, 耗时较长, 以及检测样品化学测试容易二次污染, 无法实现动态监测等缺点。针对水化学方法的不足, 王亚等<sup>[4]</sup>采用 LIF 技术对矿井突水水源的特征光谱进行识别。实验结果表明, 该方法能准确地判别水源类型, 且对流场进行连续测量, 不会造成干扰, 可进行水源的在线检测分析, 为矿井突水灾害的预测提供决策依据。

判别模型是准确判别突水水源的另一个关键因素。宫凤强等利用 PCA 和距离判别分析法建立突水模型, 对淮南老矿区谢一煤矿的不同含水层的历史样本数据进行验证<sup>[5]</sup>。Sun 提出了运用系统聚类分析的方法识别矿井水源类型<sup>[6]</sup>; 周孟然等采用极限学习机算法, 建立多元分类学习模型预测

收稿日期: 2018-05-22, 修订日期: 2018-10-20

基金项目: 国家“十二五”科技支撑计划项目(2013BAK06B01)和国家自然科学基金项目(41674133)资助

作者简介: 杨 勇, 1981 年生, 中国矿业大学资源与地球科学学院博士研究生   e-mail: yongyang@cumt.edu.cn

\* 通讯联系人   e-mail: yuejh@cumt.edu.cn

矿井水源类型<sup>[7]</sup>。现有的判别模型基于传统的浅层神经网络<sup>[8]</sup>框架,存在一定局限性。首先,在特征选择方面,需要结合大量具有专业背景和工程技能的技术人员提供大量先验知识<sup>[9-10]</sup>,具有一定主观性。其次,在训练多层 BP 网络时,考虑到目标函数为非凸函数,随机的选取网络初始权值,可能造成训练停滞于某个较差的局部最优值。

基于 LIF 技术和卷积神经网络(CNN)<sup>[11]</sup>,提出了一种新的矿井突水水源类型判别模型。首先,采用 LIF 技术获取突水水源荧光光谱;其次,提出了改进的递推平均一阶滞后平滑滤波方法,得到增强的二维自相关荧光光谱特征图;最后,构建基于深度卷积神经网络(CNN)的突水水源类型判别模型,针对光谱特征图判别突水水源类型。判别结果表明:相比传统的递推平均滤波、一阶滞后滤波方法,该方法得到的二维自相关荧光光谱特征图,有效地抑制了荧光噪声,增强了不同水源类型的特征差异性;CNN 判别模型对水源类型正确的识别率达到了 98%,是一种有效的矿井突水水源类型判别方法。

## 1 实验部分

### 1.1 仪器

采用 LIFS-405 激光诱导荧光系统(广东标旗电子有限公司生产),激光入射波长固定为 405 nm,入射激光功率 120 mW,步长设定为 0.5 nm,光谱扫描时间设置为  $0.001 \text{ s} \cdot \text{nm}^{-1}$ 。

### 1.2 材料和光谱数据获取

利用淮南矿业集团的新集二矿的水源作为样本,2016 年 6 月—2017 年 6 月期间分批次取得水样,其中采空区积水 46 条,砂岩水 59 条,太灰水 42 条和奥灰水 14 条。样品置于用 10% HCl 和蒸馏水洗净的聚丙烯样品瓶中,尽快测定其荧光光谱。考虑到数据处理的可操作性,上位机设计光谱接收范围为 340~1 340 nm,每个光谱对应的光谱数据为 2 048 维。为避免因外界干扰等造成随机误差,对每个样本进行 10 次测量,取算术平均值,得到典型水样的光谱如图 1 所示。

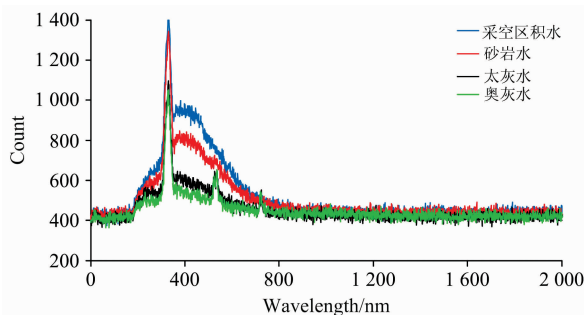


图 1 荧光光谱图

Fig. 1 Fluorescence spectra of mine water

### 1.3 荧光光谱的主成分分析

采用 matlab2016b 软件对获取的 161 组光谱数据进行主成分分析,前 10 个主成分累计贡献率为 80.940%,小于 85%,经过 PCA 降维的光谱数据有较大损失。将提取到的第

一主成分、第二主成分和第三主成分分别作为 X, Y 和 Z 轴的值,如图 2 所示。PCA 降维不能有效涵盖数据信息,四种水源样本区分不是很明显,有必要建立基于全光谱的 CNN 网络判别模型对于新集二矿水源进行判别预测。

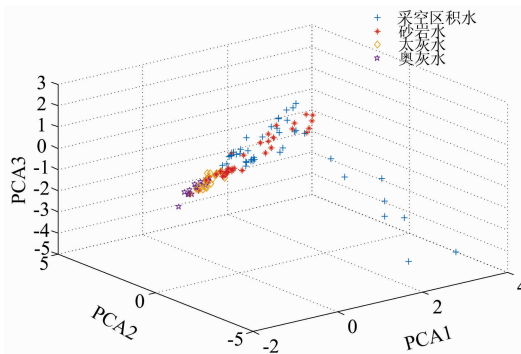


图 2 主成分得分图

Fig. 2 Principal component scores scatter plot (PCA1 × PCA2 × PCA3)

### 1.4 模型结构

CNN 的结构如图 3 所示,由输入层、隐层、全连接层、输出层组成。隐层由多组卷积层和采样层交替组成。卷积层采用特定的卷积核与输入层的特征图卷积再加上一个偏置,通过一个激活函数获得输出特征;池化层对卷积层的输出特征图进行特征筛选,通常采用平均值或最大值形式的池化。全连接层和输出层构成分类器,分类器可以是逻辑接层和输出层构成分类器,分类器可以是逻辑(Logistic)回归、Softmax 回归及 SVM。考虑到 Logistic 回归和 SVM 适用于二分类任务,而 Softmax 回归模型是 Logistic 回归的多分类拓展,用于解决多分类问题的多分类算法。CNN 的训练方式采用批量样本输入的方式,其训练过程如图 4 所示。

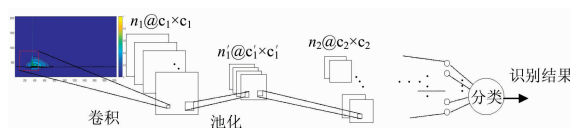


图 3 CNN 结构图

Fig. 3 CNN structure

首先,设定网络的训练参数,初始化网络的权值和偏置,输入特征图经过卷积层、采样层、全连接层处理后,传输至输出层,每一层的输出为下一层的输入;然后,将实际输出与期望输出间的误差通过 BP 算法逐层反向传播,并将误差分配到各层,对网络的权值和偏置进行调整,直至满足收敛条件,以实现网络的有监督训练。

## 2 光谱数据预处理

在矿井水源荧光光谱数据的采集过程中,由于光谱接收仪器存在光谱幅值抖动、随机误差及光线干扰等问题,导致接收到的光谱曲线在整个波段范围内出现不同程度的抖动,

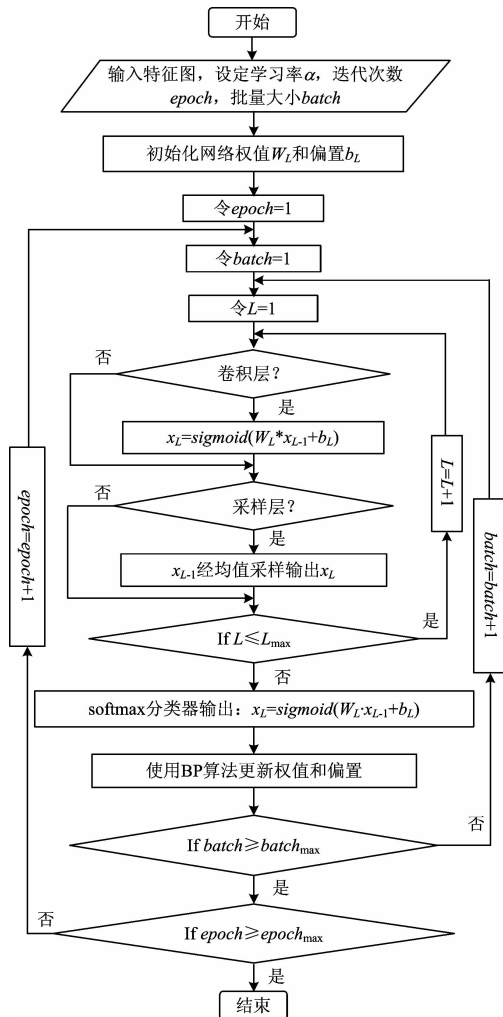


图 4 CNN 的训练过程

Fig. 4 Training process of CNN

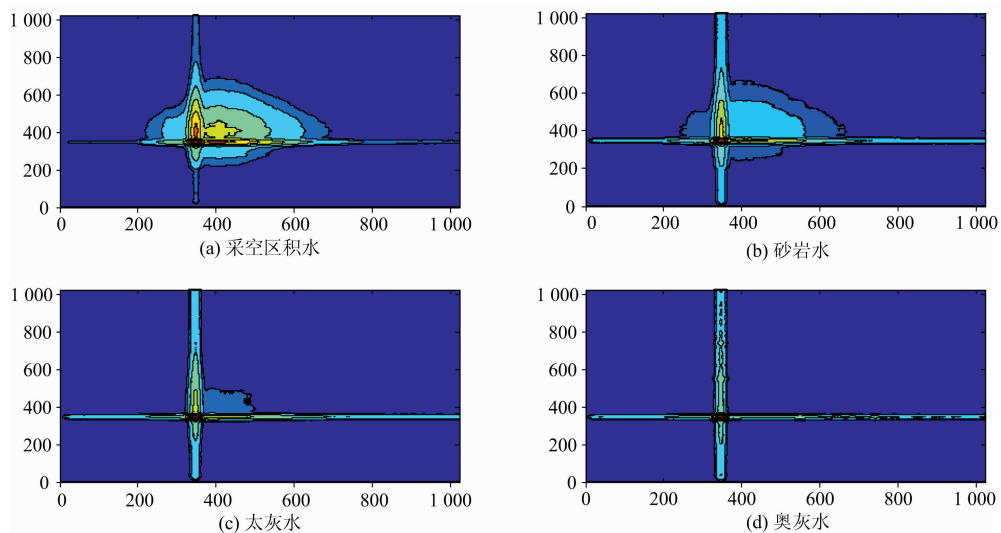


图 5 递推平均滤波的三维荧光光谱图

(a): 采空区积水; (b): 砂岩水; (c): 太灰水; (d): 奥灰水

Fig. 5 Three-dimensional fluorescence spectra of Recursive average filtering

(a): Goaf hydrips; (b): Sandstone water; (c): Limestone water; (d): Ordovician Limestone water

直接影响光谱计算的准确性, 无法反应水质的真实情况, 如图 1 所示。

### 2.1 递推平均滤波和一阶滞后滤波法

图 5 和图 6 分别给出了递推平均滤波和一阶滞后滤波后的自相关二维特征光谱图。较递推平均一阶滞后滤波法, 该方法仍不能有效滤除干扰, 经过自相关处理, 不仅放大了光谱曲线中的有效参数, 而且噪声的能量随之亦被放大, 不利于对光谱曲线的进一步分析。

### 2.2 改进的递推平均一阶滞后滤波法

提出了一种新的滤波算法, 融合了递推平均和一阶滞后滤波对所测水样的荧光光谱进行滤波处理。首先, 对光谱数据进行一阶滞后滤波处理, 如式(1)所示

$$y = y_2(1 - a) + ay_1 \quad (1)$$

其中,  $y$  为光谱滤波后的幅值,  $y_2$  为荧光光谱幅值,  $y_1$  为滤波后荧光光谱  $y_2$  的前一波长的幅值,  $a$  取 0.5。其次, 把连续取  $N$  个采样值看成一个队列, 队列的长度固定为  $N$ , 先进先出原则把队列中的  $N$  个数据进行算术平均运算,  $N$  取 14。对于改进后的递推平均一阶滞后滤波法结果如图 7 所示。对平滑处理后的一维光谱数据求自相关及归一化处理, 得到二维特征光谱图如图 8 所示, 采用改进后的递推平均一阶滞后滤波法处理一维荧光光谱数据, 得到了四种测试水样的三维荧光光谱图并表现了出了明显的差异性。

## 3 结果与讨论

本实验验证基于 CNN 的矿井突水水源识别模型对采空区积水、砂岩水、太灰水和奥灰水 4 种测试水样的判别性能。对采集的光谱数据作为样本训练 CNN 网络, 对一维光谱数据采用自相关和归一化的计算得到特征二维光谱图。采用 5 折交叉验证, 随机抽取 20% 个样本用作测试, 剩余 80% 个样本用于训练 CNN 模型。CNN 网络结构为: 卷积层、池

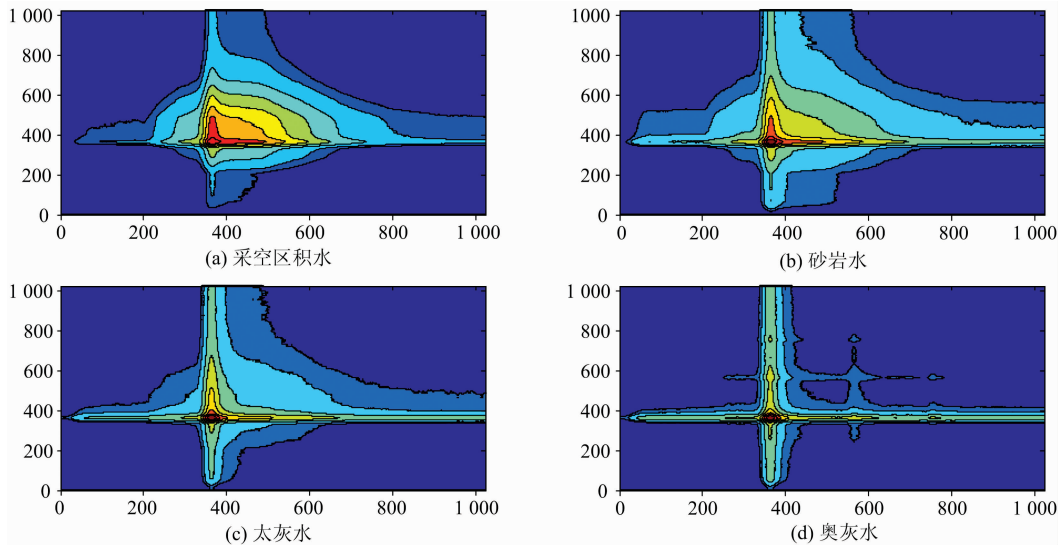


图 6 一阶滞后滤波的三维荧光光谱图

(a): 采空区积水; (b): 砂岩水; (c): 太灰水; (d): 奥灰水

Fig. 6 Three-dimensional fluorescence spectra of one order lag filtering

(a): Goaf hydrogyps; (b): Sandstone water; (c): Limestone water; (d): Ordovician Limestone water

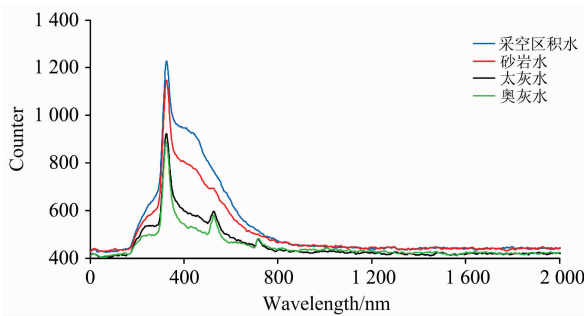


图 7 递推平均一阶滞后滤波的荧光光谱图

Fig. 7 Fluorescence spectra obtained by of one order lag filtering mixed with recursive average

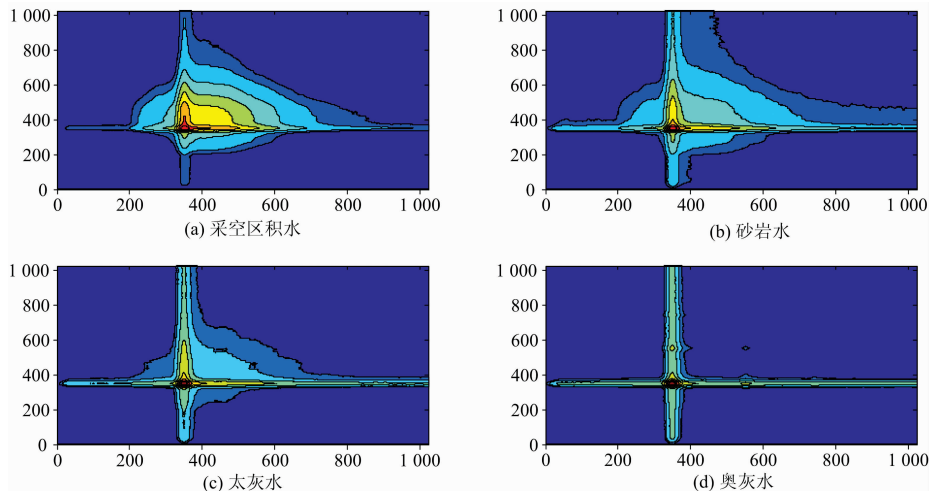


图 8 递推平均一阶滞后滤波法的三维荧光光谱图

(a): 采空区积水; (b): 砂岩水; (c): 太灰水; (d): 奥灰水

Fig. 8 Three-dimensional fluorescence spectra of one order lag filtering mixed with recursive average

(a): Goaf hydrogyps; (b): Sandstone water; (c): Limestone water; (d): Ordovician Limestone water

化层、卷积层、池化层、卷积层、池化层、全连接层、分类器。第一个卷积层采用卷积核  $64@4 \times 4$ ，第二个卷积层采用卷积核，第三个卷积层采用卷积核，池化层均采用均值池化大小为。网络训练参数设定如表 1 所示。

图 9 给出了采用 3 种方法正确识别四种水样的平均结果。其中，蓝线表示三维荧光光谱图识别率变化趋势，随着迭代次数的增加，正确识别的概率表现出了逐渐提升，识别性能有所改善。图中采用一阶滞后滤波法的识别结果的变化趋势，在迭代初期表现出较快速的提升，在 300 次迭代之后呈现出稳定的波动。其平均识别率为 85.19%，最高识别率达到了 91.8%，相比前 2 种方法的识别率有所下降。该结果的产生归因于一阶滞后滤波法得到的三维荧光光谱图对噪声的

表 1 超参数设置  
Table 1 Hyperparameters

参数名称	数值
Hyper parameters	Values
Dropout ratio	0.2
Moment	0.9
Learn rate	0.002 5
Regularization type	L2
Minibatch	64

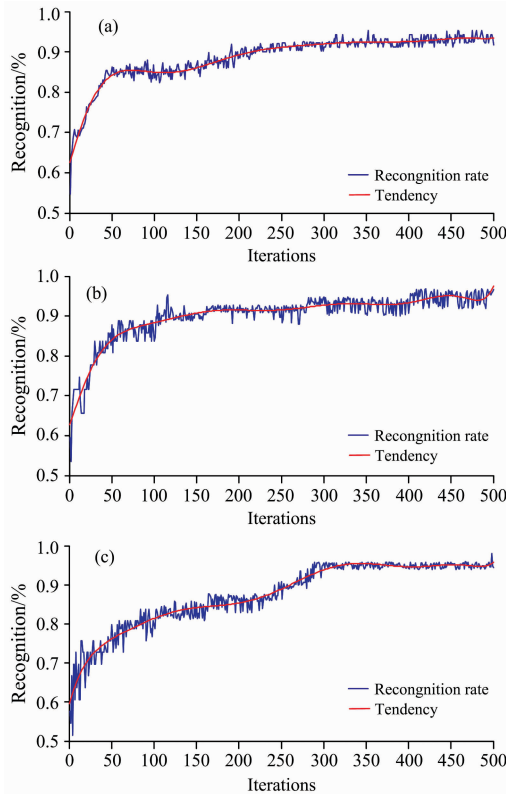


图 9 水源类型的判别结果

(a): 递推平均滤波; (b): 一阶滞后滤波;  
(c) 递推平均一阶滞后滤波

Fig. 9 Recognition results of 4 filters

(a): Recursive average filtering; (b): First-order lag filtering;  
(c): First-order lag and recursive average filtering

抑制效果较差。相比本文所提出的递推平均一阶滞后滤波法, 平均识别率达到了 87.56%, 趋于稳定之后的平均识别率达到了 94.95%, 最高识别率达到了 98%。递推平均一阶滞后滤波法呈现出最优的变化趋势。结果表明, 采用递推平均一阶滞后滤波法运用自相关函数得到的三维荧光光谱图有效抑制了荧光噪声, 有效的扩大了不同含水层水源类型特征差异。

考虑到验证不同识别模型的对矿井突水水源类型判别的准确度, 比较方法包括本文的方法和 PCA-BP 模型。模型参数设置, BP 模型采用三层网络模型, 隐藏节点的个数决定模型判别效率, 激活函数采用 tansig 作为传递函数。实验中, 先对光谱数据进行 PCA 处理, PCA 维度值分别设置为 3, 4, 6 和 8, 全部样本作为训练集, 系统随机选取样本作为测试集判定 PCA-BP 模型的效果。

PCA-BP 模型的实验结果如表 2 所示。光谱主成分数选取为 4 时, 判别正确率最高达到了 85%, 随着主成分数的增加, 判别率呈下降趋势, 当主成分数为 8 时判别率降低至 68.75%。基于卷积神经网络的判别模型判别率优于 PCA-BP 模型。

表 2 基于 PCA-BP 算法的突水水源识别率 (%)

Table 2 Recognition rate of different inrush water Based on PCA-BP algorithm

	PCA=3	PCA=4	PCA=6	PCA=8
判别率	83.75	85	80	68.75

### 4 结 论

为了解决矿井突水水源类型的在线判别分析问题, 从判别模型分析入手, 提出了一种基于 LIF 技术和卷积神经网络(CNN)的矿井突水水源判别模型, 并将该算法应用到 4 种含水层水样水源判别分析中。算法特点包括: (1) 针对水样光谱数据的高频随机噪声, 提出了递推平均一阶滞后滤波法, 生成二维自相关荧光光谱特征图。(2) 针对二维自相关荧光光谱特征, 构建了基于 CNN 的矿井突水水源类型判别模型。实验结果表明, 该方法对水源类型的准确识别率达到了 98%, 是一种有效的矿井突水水源类型判别方法, 为矿井突水水源类型判别方法提供了新的思路。目前的工作仅对 4 种水样进行了辨识, 下一步将扩大样本数量, 进一步探究 4 种矿井水的差异, 为真正的工程应用提供理论依据。

### References

[ 1 ] WU Qiang, CUI Fang-peng, ZHAO Su-qi, et al(武 强, 崔芳鹏, 赵苏启, 等). Journal of China Coal Society(煤炭学报), 2013, 38(4): 561.

[ 2 ] XU Xing, TIAN Kun-yun, WANG Gong-zhong, et al(徐 星, 田坤云, 王公忠, 等). Journal of Safety & Environment(安全与环境学报), 2017.

[ 3 ] LIU Jian-min, WANG Ji-ren, LIU Yin-peng, et al(刘剑民, 王继仁, 刘银朋, 等). Journal of Safety & Environment(安全与环境学报), 2015, (1): 38.

[ 4 ] WANG Ya, ZHOU Meng-ran, YAN Peng-cheng, et al(王 亚, 周孟然, 闫鹏程, 等). Spectroscopy and Spectral Analysis(光谱学与光

- 谱分析), 2017, 37(3): 978.
- [ 5 ] GONG Feng-qiang, LU Jin-tao(宫凤强, 鲁金涛). Journal of Mining & Safety Engineering(采矿与安全工程学报), 2014, 31(2): 236.
- [ 6 ] Sun L H. Arabian Journal of Geosciences, 2014, 7(9): 3417.
- [ 7 ] WANG Ya, ZHOU Meng-ran, YAN Peng-cheng, et al(王 亚, 周孟然, 闫鹏程, 等). Journal of the China Coal Society(煤炭学报), 2017, 42(9): 2427.
- [ 8 ] Tirumala S S, Narayanan A. Hierarchical Data Classification Using Deep Neural Networks. International Conference on Neural Information Processing, 2015.
- [ 9 ] Ali J B, Saidi L, Mouelhi A, et al. Engineering Applications of Artificial Intelligence, 2015, 42(C): 67.
- [10] Poswar F D O, Farias L C, Fraga C A D C, et al. Journal of Endodontics, 2015, 41(6): 877.
- [11] Dabove P, Manzino A M. GPS Solutions, 2017, 21(3): 1213.

## Online Discrimination Model for Mine Water Inrush Source Based CNN and Fluorescence Spectrum

YANG Yong<sup>1,3</sup>, YUE Jian-hua<sup>1\*</sup>, LI Jing<sup>2</sup>, ZHANG He-rui<sup>1</sup>

1. School of Resources and Geosciences, China University of Mine and Technology, Xuzhou 221008, China

2. School of Information Engineering, Nanjing Audit University, Nanjing 210029, China

3. School of Information and Electrical Engineering, Xuzhou College of Industrial Technology, Xuzhou 221140, China

**Abstract** As deep mining goes, the water inrush threat is from the roof goaf water and the bottom pressure karst water. Coal mines water inrush water types on-line discrimination, serving as an effective monitoring method to predict mine water hazards, is an important step in Mine water disaster prevention and control work to ensure coal mine safety production. Representative ion method, as a traditional method to discriminate mine water inrush sources, must collect and seal water samples on-site, test samples in laboratory using 7 typical inorganic ion concentrations, and calculate water bursting evaluation factor. The method has disadvantages of too long detection time, easy contamination for samples, delayed warning response and misjudgment. Due to above reasons, the paper proposes a mine water inrush sources discrimination model based on Laser Induced Fluorescence (LIF) and Convolutional Neural Network (CNN). First, based on 4 types of water sources, 161 samples were collected from Xinji Second Mine of Huainan mining group during June 2016 to June 2017, including oaf water 46 items, Sandstone water 59 items, Limestone water 42 items and Ordovician limestone water 14 items. In the experiment, samples were stimulated by 405 nm laser using LIFS-405 Laser Induced Fluorescence System, and the fluorescence spectra of four kinds of 161 groups of water inrush samples were obtained. During principal component analysis, the cumulative contribution rate of the top ten components was less than 85%, making 4 types of water samples almost indistinguishable. Second, considering the random high frequency fluctuations in water fluorescence spectra, first-order lags filtering method should be used to reduce periodic high frequency fluctuations. Considering data update rate, recursive averaging method should be adopted. The paper proposes an improved recursive average first-order lag smoothing filtering method further to calculate autocorrelation processing to get enhanced two-dimensional autocorrelation characteristic fluorescence spectra. The experimental results show that calculated autocorrelation characteristic fluorescence spectra have excellent performance on interference elimination and discrimination. Finally, based on autocorrelation characteristic fluorescence spectra, mine water inrush sources discrimination model using CNN was constructed to discriminate water inrush types. The method adopts deep learning framework using autocorrelation characteristic fluorescence spectra to avoid selecting features in subjective ways. Theoretical analysis and experimental results show that the correct recognition rate of water source type can reach 98%. It is an effective way to discriminate the source of water inrush from mines and provides a new idea to discriminate the types of mine water inrush sources.

**Keywords** Discriminate inrush water source types; Laser-induced fluorescence; Convolution neural network

(Received May 22, 2018; accepted Oct. 20, 2018)