

基于特征组合与 CNN 的大坝缺陷识别与分类方法

毛莺池¹ 王 静¹ 陈小丽¹ 徐淑芳¹ 陈 豪²

(河海大学计算机与信息学院 南京 211100)¹ (河海大学水利水电学院 南京 210098)²

摘 要 大坝缺陷识别分类技术是人类智能的基本表现,它是最典型、最困难的模式识别问题之一。由于大坝缺陷图像具有信噪比低、光照分布极度不均匀等特征,分类识别算法的识别率较低。针对这些问题,文中提出一种基于图像 LBP 特征和 Gabor 特征组合与 CNN 相结合(LBP and Gabor feature combination and CNN, LG-CNN)的缺陷图像识别方法,对采集到的大坝图像进行分析,实现对缺陷图像的识别和分类。该方法首先分别提取图像的 LBP 特征与 Gabor 特征;然后将得到的 LBP 特征和 Gabor 特征组合作为 CNN 的输入;最后通过逐层训练网络,实现大坝缺陷类型的分类识别。实验结果表明, LG-CNN 的平均识别准确率为 88.39%,缺陷召回率为 92.75%,与相同参数设置下的 CNN 分类识别算法相比,识别准确率和缺陷召回率分别约提高了 3.1%和 2.5%,具有最优的结果。

关键词 缺陷图像, LBP 特征, Gabor 特征, CNN, 分类

中图法分类号 TP311 文献标识码 A DOI 10.11896/j.issn.1002-137X.2019.03.040

Dam Defect Recognition and Classification Based on Feature Combination and CNN

MAO Ying-chi¹ WANG Jing¹ CHEN Xiao-li¹ XU Shu-fang¹ CHEN Hao²

(College of Computer and Information, Hohai University, Nanjing 211100, China)¹

(College of Water Conservancy and Hydropower, Hohai University, Nanjing 210098, China)²

Abstract Dam defect recognition and classification technology is the basic manifestation of human intelligence. It is one of the most typical and difficult pattern recognition problems. Due to the low signal-to-noise ratio and extremely uneven illumination distribution of dam defects, the recognition rate of classification and recognition algorithm is relatively low. In order to solve these problems, this paper proposed a defect image recognition method based on the combination of image LBP features and image Gabor features combined with CNN(LBP and Gabor feature combination and CNN, LG-CNN), analyzed the collected dam image, and realized the recognition and classification of the defective images. Firstly, the LBP features and the Gabor features of images are extracted respectively. Then, the features of LBP and Gabor are combined to be the input of CNN. Finally, by training the network layer by layer, the classification and recognition of dam defects are realized. The experimental results show that the average recognition accuracy of LG-CNN is 88.39%, as well as the recall rate of defect is 92.75%. Compared with the CNN classification algorithm under the same parameters, the recognition accuracy and the recall rate of defect are increased by 3.1% and 2.5% respectively, and the results is the best results.

Keywords Defect image, LBP feature, Gabor feature, CNN, Classification

1 引言

我国是一个水利大国,目前已拥有水库大坝 9.8 万余座,是世界上水库大坝最多的国家^[1]。但随着时间的推移和坝龄的增长,温度、环境压力等各种因素使得大坝出现了变形、裂缝、渗漏、钙化物析出等一系列表观缺陷。因此,充分利用大坝传感器监测到的图像数据,从中挖掘出有效信息,辅助大坝传感器日常监测数值的数据分析,从而进行更合理、更准确的决策,对后期的险情诊断和修复加固具有重要的意义。

Zhang 等^[2]采用多层感知器比较了 Gabor 特征和几何特征,发现 Gabor 特征具有更好的识别性能,但 Gabor 变换的计算量和特征维数较高,对内存的需求很大。邓洪波等^[3]提出了一种局部 Gabor 滤波器组,其有效地提高了特征的提取速度,减少了存储量,并用 PCA 方法^[4]降维,得到了较好的识别效果。近年来, Gabor 特征提取方法^[5]在特征维数、实时性和准确性上都有所突破。Gabor 小波主要依靠 Gabor 核在频域上对信号进行加窗,描述信号的局部频率信息,可以提取多尺度、多方向的空间频率特征,使得一些局部图像纹理信息

到稿日期:2018-07-03 返修日期:2018-09-15 本文受“十三五”国家重点研发计划项目(2018YFC0407105),华能集团重点研发课题(HNKJ17-21)资助。

毛莺池(1976—),女,博士,教授,硕士生导师,CCF 会员,主要研究方向为分布式数据管理, E-mail:yingchimao@hhu.edu.cn(通信作者);王 静(1993—),女,硕士生,主要研究方向为数据融合、分布式计算;陈小丽(1993—),女,硕士,主要研究方向为数据融合、分布式计算;徐淑芳(1981—),女,博士,讲师,主要研究方向为无线通信网络、信息融合;陈 豪(1982—),男,博士生,高级工程师,主要研究方向为水工结构安全监测。

的特征被放大^[6],更有利于特征识别。而LBP特征能描述图像的全局特征,进行出缺陷图像中亮点、暗点、光照均匀区域以及不同方向边缘等微小特征的分布情况。

综上所述,针对大坝缺陷图像信噪比低、光照分布极度不均匀等特征,亟需一种受光照影响小且能描述缺陷局部纹理特征的图像特征提取方法^[7]。本文综合考虑缺陷图像的全局特征和局部特征,提出了一种基于图像组合特征与卷积神经网络^[8]的缺陷图像分类识别方法(LG-CNN)。该方法将Gabor小波变换和LBP特征提取方法相结合^[9],利用它们各自的优点来获得更高的准确度和更好的鲁棒性。基于特征组合与CNN的缺陷分类识别框架如图1所示。

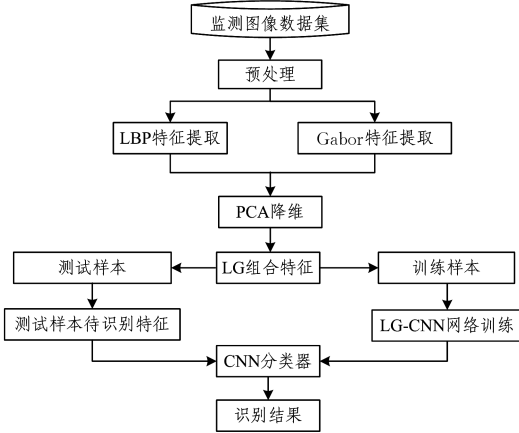


图1 基于特征组合与CNN的缺陷分类识别框架

Fig.1 Framework of defect classification and identification based on feature combination and CNN

本文第2节首先对图像进行LBP特征提取和Gabor特征提取,接着进行LG特征组合,最后进行PCA降维以提取图像特征;第3节将图像的LBP特征与Gabor特征组合后作为CNN的输入,通过逐层训练网络,实现缺陷类型的识别;第4节对比分析了LG-CNN模型的识别效果和运行效率;最后总结本文工作,给出缺陷图像分类识别未来的研究方向。

2 基于LBP-Gabor缺陷图像的特征提取

对从某特高拱坝实际采集的大坝图像数据进行人工视觉层面和检测程序的初步甄别,可以将大坝缺陷图像归类为正常图像、裂缝图像、渗水点图像、渗水面图像以及钙化物析出图像5种,如图2所示。相应地,经LG-CNN模型训练输出的识别分类结果图像中也包括这5类。

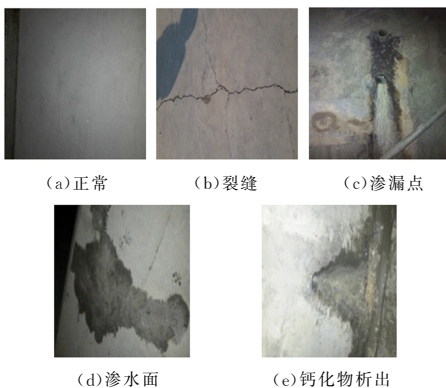


图2 大坝缺陷原图

Fig.2 Original image of dam defect

对输入图像进行预处理,提取图像的LBP+Gabor组合特征并进行PCA降维处理。假设输入图像为 $f(x,y)$,归一化大小为 256×256 。

1) LBP特征提取

$$LBP_{P,R}(g_c) = \sum_{i=0}^{P-1} s(g_i - g_c) 2^i \quad (1)$$

其中, R 为LBP算子覆盖区域的半径, P 为采样点数目, g_c 为图像中心灰度值, g_i 为中心点周围的灰度值。如果 $g_i > g_c$,则 $s(g_i - g_c) = 1$;否则 $s(g_i - g_c) = 0$ 。取 8×8 采样窗口,步长设置为4,依次计算 64×64 采样窗口内的LBP统计特征,得到 $64 \times 64 \times 59$ 维的LBP直方图特征,连接各个区域的LBP直方图即可构成一个LBP特征向量: $\vec{Fet}_{LBP}(x,y) = [b_{x,y,1}, b_{x,y,2}, \dots, b_{x,y,59}]$ 。

2) GABOR特征提取

$$\psi_{\mu,\nu}(z) = \frac{\|k_{\mu,\nu}\|^2}{\delta^2} e^{(-\|k_{\mu,\nu}\|^2 z^2 / 2\delta^2)} [e^{izk_{\mu,\nu}} - e^{\delta^2/2}] \quad (2)$$

其中, $k_{\mu,\nu} = k_\nu e^{i\varphi_\mu}$, $k_\nu = \frac{k_{\max}}{\lambda^\nu}$ 定义频率; $\varphi_\mu = \frac{\mu\pi}{8}$, $\varphi_\mu \in [0, \pi)$ 定义方向; $z = x, y$; $\delta = 2\pi$; $k_{\max} = \pi/2$; $\lambda = \sqrt{2}$ 。

$$G_{\psi f}(x,y,\mu,\nu) = f(x,y) * \psi_{\mu,\nu}(z) \quad (3)$$

其中,*为卷积操作,将输入图片分成 64×64 大小的块,每个块使用5种尺度 $\nu \in \{0, \dots, 4\}$ 和8个方向 $\mu \in \{0, \dots, 7\}$ 的Gabor滤波器 $\psi_{\mu,\nu}(z)$ 。滤波后得到40幅Gabor图像,连接图像对应像素点的Gabor特征可得到 $64 \times 64 \times 40$ 维的Gabor特征向量 $\vec{Fet}_{Gabor}(x,y) = [a_{x,y,1}, a_{x,y,2}, \dots, a_{x,y,40}]$,如图3所示。

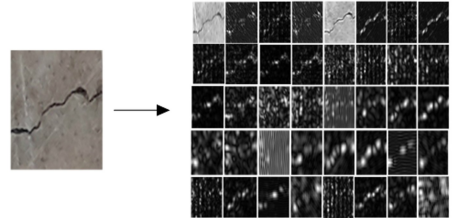


图3 Gabor特征提取图

Fig.3 Gabor feature extraction diagram

3) LG组合特征

大坝缺陷识别是多类识别,有一对一和一对多等识别方法,但通常具有分类器数量较多的缺陷。基于同一类样品之间的联系与不同类别样本之间的联系不同这一原则,可以把多类分类问题转化为二类分类问题。因此,大坝缺陷识别的多类问题也可以转化为大坝缺陷图像之间的类内相似度和类间相似度这一二分类问题。本文中针对LBP特征和Gabor特征的组合,采用求取两向量余弦值计算两向量相似度的方法。

4) PCA降维

本文选用主成分分析法^[5]进行降维处理,在最大程度保留原始数据特征方向的基础上将其投影到较低维空间,最终选取主成分方差占比为97.4%的特征作为提取的LG组合特征,以此作为卷积神经网络的输入。

3 LG-CNN模型训练

3.1 LG-CNN整体结构

基于MatConvNet工具箱设计CNN结构,经图像特征提

取和组合得到单幅图像 1024×1 维的 LG 组合特征向量,通过调整向量行列数将特征向量重构为 32×32 维后作为 CNN 输入进行网络训练过程。表 1 列出了每层网络的操作和具体维度。

表 1 LR-CNN 每层的操作与维度

Table 1 Operation and dimensions of each layer of LR-CNN

级层	LG-CNN 网络层级				卷积核设置				
	高度	宽度	通道	操作	高度	宽度	步长	个数	
输入	32	32	1	C1	7	7	1	8	
L1	26	26	8	P1	7	7	1	—	
L2	20	20	8	C2	3	3	2	16	
L3	9	9	16	P2	5	5	1	—	
L4	5	5	16	C3	5	5	1	96	
L5	1	1	96	RELU	—	—	—	—	
L6	1	1	96	Softmax	—	—	—	—	
L7	1	1	5	—	—	—	—	—	

首先利用一定数量的已标记图像样本建立网络训练样本集,初始化网络参数和权重,然后提取样本集特征作为网络输入进行样本训练。训练过程中通过不断地迭代更新权重来获得分类效果最好的训练网络权重。网络训练结束后,采集到的原图像即可通过训练好的分类器判别出不同类型的大坝缺陷图像。

3.2 CNN 网络参数的设置

CNN 模型架构由许多参数确定,对于不影响 CNN 性能或影响很小的参数,通常按照最优化设置为固定值;而对于能够影响网络收敛和训练准确度的参数,需要在严格的规则下进行调节。影响 CNN 网络的几个关键参数为:学习率(Learning rate)^[10]、权值衰变(Weight Decay)^[11]、动量(Momentum)^[12]、最小批尺寸(mini-Batch size)^[13]、丢失率(Dropout ratio)^[14]。表 2 列出了 LG-CNN 网络的部分参数设置。

表 2 LG-CNN 模型训练的参数设置

Table 2 Parameter settings of LG-CNN model training

参数类型	设定值	属性
权值衰变	0.0001	fixed
动量	0.9	fixed
最小批尺寸(mini-Batch size)	100	—
最大迭代次数(Max iteration)	200000	—
Crop_size	256×256	—
Dropout ratio	0.5	—

3.3 CNN 训练过程和权重更新

由于网络初始权值 ω 是随机分配的,在训练期间,预测值和实际分类通常是不一致的。为了计算两者之间的偏差量,定义 Softmax 损失函数^[15],公式如下:

$$L = \frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^n 1\{y^{(i)} = j\} \log \frac{e^{w_j^T x^{(i)}}}{\sum_{t=1}^n e^{w_t^T x^{(i)}}} \right] + \frac{\lambda}{2} \sum_{j=1}^n W_j^2 \quad (4)$$

其中, L 用来表示 $\sum_{i=1}^m \exp(W_i^T x^{(i)})$ 与 $\sum_{j=1}^n 1\{\cdot\}$ 之间的独立关系; $1\{y^{(i)} = j\}$ 是逻辑表达式,返回值为 0 或 1,若第 i 个输入的预测类相对 j 类是正确的,则返回 1,否则返回 0; λ 是正则项(权重衰变)参数,用于惩罚大的权重,防止出现过度拟合的情况。

为了缩小偏差量,获得期望的结果,更新权重是必要的,

这也是 CNN 训练最重要的步骤。本文选用基于反向传播的 SGD 算法^[16]实验上述目标。标准的梯度下降算法是在整个训练集的基础上更新 ω ;而 SGD 可以在单个或多个训练样本上执行,能最简单、有效地实现最小化偏差。同时为了加快训练速度,在 SGD 算法中引入动量算法(MOBP),具体更新过程如下。

步骤 1 计算出相对于权值 ω 的损失函数梯度值:

$$\nabla_w L(W; x^{(i)}, y^{(i)}) = \frac{1}{m} \sum_{i=1}^m [x^{(i)} \{1(y^{(i)} = j) - p(y^{(i)} = j | x^{(i)}; W)\}] + \lambda W_j \quad (5)$$

步骤 2 超参数引入动量 ϵ 和学习速率 α 来更新速度 v :

$$v \leftarrow \epsilon v - \alpha \nabla_w L(W; x^{(i)}, y^{(i)}), W_j \leftarrow W_j + v \quad (6)$$

在物理学中动量被定义为质量与速度的乘积,在这里考虑的是单元质量。

步骤 3 权重 ω 按式(5)迭代更新,直到收敛,也即 W_j 不再变化为止。式(6)中的上标(i)表示第 i 个训练样本, i 的范围取决于 mini-Batch size(最小批尺寸),即整个数据集中使用的训练样本数量。整个卷积神经网络通过重复上述过程进行偏差量缩小的权重调整工作,最终得到收敛的优化网络权重 ω 。通过保存权重以及之前训练好的网络参数可以得到训练好的 LR-CNN 分类器,用于步骤 3 中测试图像集的分类识别。

步骤 4 缺陷分类识别的输出:经过步骤 2,CNN 分类器训练完成,卷积神经网络的最后一层即 Softmax 层会将多个神经元的输出映射到(0,1)区间内。根据式(7)可以得到第 i 个训练样本属于 m 个训练样本,第 j 类属于 n 类中的概率表达,也即每个输入的独立分类概率。其中函数输入为上层输出的特征向量 $W_n^T x^{(i)}$;输出为一个 5 维的权重向量,代表测试图像隶属于正常、裂缝、渗水点、渗水面、钙化物析出这 5 类缺陷的概率值,总权重值为 1。

$$P(y^{(i)} = n | x^{(i)}; \omega) = \begin{bmatrix} p(y^{(i)} = 1 | x^{(i)}; \omega) \\ p(y^{(i)} = 2 | x^{(i)}; \omega) \\ \vdots \\ p(y^{(i)} = n | x^{(i)}; \omega) \end{bmatrix} = \frac{1}{\sum_{j=1}^n e^{w_j^T x^{(i)}}} \begin{bmatrix} e^{w_1^T x^{(i)}} \\ e^{w_2^T x^{(i)}} \\ \vdots \\ e^{w_n^T x^{(i)}} \end{bmatrix} \quad \text{for } i=1, \dots, m \quad (7)$$

根据测试结果的概率向量,最大概率值所对应的缺陷分类即为测试分类的识别结果。

步骤 5 测试图像结果

CNN 网络权重更新完成,也即分类器训练完成后,可以输入测试图像进行分类识别。本文中选取测试图像的大小为 3264×1836 ,训练样本的大小为 256×256 。这是因为选择较小尺寸的图像来训练网络,可以扫描大于设计尺寸的任何图像,同时对于样本数量的扩充也有帮助;但是如果训练样本图

像尺寸过小,则会造成样本标注困难,一些非缺陷的细节特征也将被关注到,造成模型准确度下降。因此,应选取合适的样本图像大小。

对于缺陷的定位检测,本文通过滑动窗口的方法,在测试图像的不同位置得到滑动窗口大小固定的 LBP 和 Gabor 组合特征,将其输入训练好的 CNN 分类器中进行高层次特征提取和匹配,得到所有的预测值;然后通过筛选,丢弃预测值低的窗口并进行滑窗合并等操作;最终得到输出结果。各类缺陷识别结果如图 4 所示。



图 4 缺陷识别结果图

Fig. 4 Defect recognition result map

算法 1 和算法 2 描述了 LG-CNN 缺陷分类识别算法的过程。首先输入图像数据集;然后分别提取 LBP 特征和 Gabor 特征,经 PCA 降维后组合两种特征向量得到 LG 组合特征,并将其作为模型的输入;通过前向传播和权值更新训练 CNN 分类器,得到结果向量和分类误差。

算法 1 LG-CNN 算法

输入:图像数据集,学习率 α , batchsize, 迭代次数 num, 缺陷结果类型集 $P_{\text{result}} = [P_1, P_2, P_3, P_4, P_5] = [\text{正常}, \text{裂缝}, \text{渗水点}, \text{渗水面}, \text{钙化物析出}]$

输出:分类结果 $P(y^{(i)} = n | x^{(i)}; \omega)$, 误差 Error

1. 提取 LBP 特征向量 $\vec{F}_{\text{etLBP}}(x, y) = [b_{x,y,1} \ b_{x,y,2} \ \dots \ b_{x,y,59}]$
2. 提取 Gabor 特征向量 $\vec{F}_{\text{etGabor}}(x, y) = [a_{x,y,1} \ a_{x,y,2} \ \dots \ a_{x,y,40}]$
3. $\vec{F}_{\text{etLG}}(x, y) \leftarrow \vec{F}_{\text{etLBP}}(x, y) \times \vec{F}_{\text{etGabor}}(x, y)$ // 获得 LG 组合特征
4. 处理 LG 特征,将数据归一化,并使用 PCA 将 LG 特征组合降至 3 维,调整特征行、列数;各类缺陷分别抽取 1200 幅图片样本,提取图片样本的 LG 特征数据作为后续实验的训练数据。
5. 初始化网络参数 r , batchsize, n // 对 CNN 网络参数进行初始化
6. 随机初始化网络, $\text{bias} \leftarrow 0$ // 对 CNN 卷积核与权重进行随机初始化
7. 开始训练网络:
8. for $i=1$ to num do
9. for $j=1$ to numbatchesize do
10. return SGDM($\alpha, \epsilon, \theta, \vartheta, \text{train_size}$) // 方向传输更新权值
11. 验证网络:
12. $P(y^{(i)} = n | x^{(i)}; \omega) \leftarrow \frac{1}{\sum_{j=1}^n e^{w_j^T x^{(i)}}} [e^{w_1^T x^{(i)}} \ e^{w_2^T x^{(i)}} \ \dots \ e^{w_n^T x^{(i)}}]$ // 输出分类概率
13. return Error // 计算分类误差
14. return $P(y^{(i)} = n | x^{(i)}; \omega)$ // 分类结果
15. end

算法 2 SGDM 算法

输入:学习率 α , 动量参数 ϵ , 初始参数 θ , 初始速度 ϑ , 样本数量 train_size

输出:训练好的网络权重

1. 初始化 θ
2. while $m < \text{train_size}$ do // 没有达到停止准则
3. $y^{(i)} \leftarrow \{x^{(1)}, \dots, x^{(m)}\}$ // 包含 m 个样本 $\{x^{(1)}, \dots, x^{(m)}\}$ 的小批量, 对应目标 $y^{(i)}$
4. $g \leftarrow \frac{1}{m} \nabla_{\theta} \sum_i L(f(x^{(i)}; \theta), y^{(i)})$ // 计算梯度估计
5. $\vartheta \leftarrow \epsilon \vartheta \leftarrow \alpha g$ // 计算速度更新
6. $\theta \leftarrow \theta + \vartheta$ // 应用更新
7. end while

4 实验验证

4.1 实验环境

实验数据来源于某特高拱坝的实测图像资料,由放在在每个采集点的监测仪器按时间序列采集。选取 6640 张原始采集图像,每幅图像的像素大小均为 3264×1836 。其中,5540 张图像用于训练和验证过程,1100 张图像用于测试过程。

如图 5 所示,图像从中心位置向外衍射分割成 256×256 像素的图像,用来建立训练样本集和验证样本集。样本集图像已标注缺陷的类型,每个图像都经过重采样。表 3 列出了用于 LG-CNN 训练和测试的分类图像数目,每次训练随机从样本库选择训练样本集,其中用于训练和验证的数量之比为 4:1。

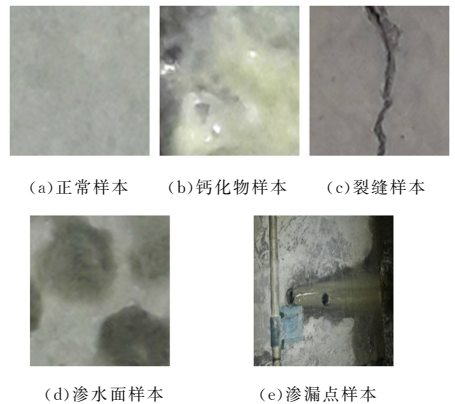


图 5 各类缺陷样本图

Fig. 5 Various types of defect samples

表 3 LG-CNN 实验数据集信息

Table 3 Information of LG-CNN experimental data set

序号	类别	训练集	验证集
1	正常	1200	2800
2	裂缝	1200	2076
3	渗水点	1200	820
4	渗水面	1200	1500
5	钙化物	1200	2380
6	总	6000	9576

为了对本文提出的 LG-CNN 混合模型进行性能检验,根

据以下两个条件选择对比算法。

1) 与不同特征输入的识别算法比较:选用相同的 CNN 分类器,根据输入训练网络特征的不同,将本文基于特征组合与 CNN 的缺陷识别方法(LG-CNN)与基于 LBP 特征的 CNN 算法(L-CNN)、基于 Gabor 特征的 CNN 算法(G-CNN)进行对比分析。

2) 与不同分类器的性能比较:用经典 CNN, SVM^[17], LDA^[18], AdaBoost^[19] 等分类器作为对比算法进行特征训练实验,通过比较分析算法的多种评价指标来验证本文算法的有效性。其中, CNN 分类器的参数设置与本文 LG-CNN 算法的参数设置一致,输入为图像训练集的灰度图像; SVM 和 LDA 分类器的特征仍为 LBP+Gabor 组合特征, SVM 采用常用的 LIBSVM, 其核函数采用径向基函数(Radial Basis Function, RBF)^[20]: $K(x, x_i) = \exp\{-|x - x_i|^2 / 2\delta^2\}$, δ 为 RBF 核函数参数。

实验实现了基于特征组合与 CNN 识别算法(LG-CNN)的参数验证过程;与 L-CNN, G-CNN, CNN 等算法进行了对比实验,从不同特征输入的识别算法和不同分类器的性能两个条件出发进行了实验对比分析。

4.2 实验评价标准

实验结果主要通过表 4 所列的 3 个指标进行评估。

表 4 实验结果评价标准

Table 4 Evaluation criteria of experimental results

指标	说明
缺陷识别准确度	识别效果评价指标包括准确率、召回率和综合评价指标
特征提取时间	针对不同的特征输入,卷积神经网络提取特征的平均时间
分类识别时间	算法从输入到识别出缺陷类型所需要的平均时间

1) 缺陷识别准确度

目前用于评价图像分类识别算法的指标^[21]主要有误差率、准确率、精准率、召回率、综合指标。这些指标通过混淆矩阵来衡量。混淆矩阵可以表示模型分类,以缺陷类为例,将裂缝的分类结果分为 4 类。

① 真正类 TP(True Positive):实际图像为裂缝,模型识别结果为裂缝;

② 假正类 FP(False Positive):实际图像为非裂缝,模型识别结果为裂缝;

③ 真负类 TN(True Negative):实际图像为非裂缝,模型识别结果为非裂缝;

④ 假负类 FN(False Negative):实际图像为裂缝,模型识别结果非裂缝。

4 种情况的记号定义如表 5 所列。

表 5 4 种情况的记号定义

Table 5 Notation definitions under four conditions

实际值/预测值	分类结果为裂缝	分类结果为非裂缝
实际裂缝	TP	FN
实际非裂缝	FP	TN

基于表 5 的统计和表 4 中对缺陷识别准确度的说明,以缺陷类为例,将准确率、召回率和综合评价指标分别做如下定义。

准确率(Precision, P):分类正确的图像数目占预测结果为正类的图像数目的百分比。

$$P = \frac{TP}{TP + FP} \quad (8)$$

召回率(Recall, R):分类正确的图像数目占测试图像中实际为正类的图像数目的百分比。

$$R = \frac{TP}{TP + FN} = \frac{TP}{P} \quad (9)$$

综合评价指标(F-measure):在 P 和 R 指标出现矛盾的情况下,需要综合考虑两者。

$$F = \frac{(\alpha^2 + 1)P * R}{\alpha^2(P + R)} \quad (10)$$

准确率与召回率是相互影响的,一般情况下,两者的增长趋势是相反的,此时可以利用综合指数 F 进行评价, F 越大,性能越好。本文在进行缺陷检测与识别实验时,优先考虑提高准确率,在保证准确率的情况下提升召回率。

2) 特征提取时间

特征提取时间是针对不同的特征输入,卷积神经网络提取特征的平均时间,时间越短,算法的运行效率越高。

3) 分类识别时间

分类识别时间是算法从输入到识别出缺陷类型所需要的平均时间,时间越短,算法的运行效率越高。

4.3 算法参数设置的对比分析

在卷积神经网络中,选取不同的参数时实验结果也会有所不同,因此,通过对不同卷积核(k_x, k_y)、批尺寸(batchsize)进行研究学习,得出最后网络训练的最优结果。

1) 卷积核参数的影响

卷积核大小决定了提取的 Feature Map 的大小,是提高 CNN 性能的关键。表 6 反映了卷积核参数的设置对分类误差的影响。其中 k_1 指第一层卷积核, k_2 指第二层卷积核, $error_1, error_2, error_3$ 是卷积层特征图个数分别为 6 和 12、7 和 14、8 和 16 时实验的分类误差。

表 6 卷积核参数的设置对 CNN 网络训练误差的影响

Table 6 Effect of convolution kernel parameter setting on training error of CNN network

k_1	k_2	$error_1/\%$	$error_2/\%$	$error_3/\%$
3	4	13.31	12.26	11.97
3	6	11.53	11.58	11.19
3	8	11.74	12.66	12.63
3	10	40.70	13.14	13.04
5	3	13.51	12.67	14.50
5	5	11.13	10.57	11.16
7	2	12.98	14.75	30.70
7	4	10.50	11.04	10.43
7	6	10.83	10.36	11.35

由表 6 可知,对同一个数据集进行实验时,卷积层的特征图个数的变化会影响实验误差,误差最小时对应的卷积层的卷积核大小不变。理论上,卷积核越大越好,但如果太大,特征图过于复杂,超过卷积核的表示能力时,将很容易出现过拟合的情况,因此,卷积核大小需谨慎选择。

2) batchsize 的影响

实验中,保持其他参数不变,改变 batchsize 的大小,结果如表 7 所列。

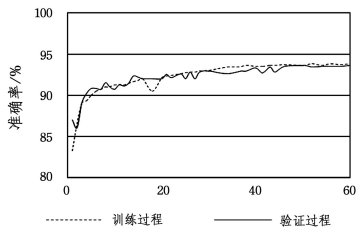
表 7 不同 batchsize 对训练结果的影响

Table 7 Effects of different batchsize on training results

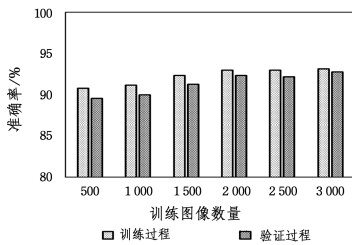
batchsize	20	40	80	160
error/%	3.53	4.86	7.80	11.13

由表 7 可知,其他参数相同时,随着 batchsize 的减小,误差减小。实验中每次选取 batchsize 个样本进行训练,训练完成后,网络模型参数会被更新。因而,在实验数据量较小的情况下,减小 batchsize 的值可能会增加运行时间,使模型跳出局部最优。

确定各参数的大小后,将得到一个完整的 LG-CNN 模型架构。当一个完整的数据集通过了神经网络一次并且返回了一次,这个过程称为一个 epoch。训练过程中,每次权重更新(epoch)的准确度变化如图 6 所示。



(a) 训练过程中每个 epoch 的准确度



(b) 不同样本数量下的准确度

图 6 LG-CNN 训练过程中准确度的变化

Fig. 6 Accuracy changes in LG-CNN training process

如图 6(a)所示,在 CNN 训练过程中,每更新一次权重,网络就会自动对输入的图像特征再次进行学习以提取深层特征,此时准确度也会上升,直至网络收敛。在第 51 个 epoch 的准确率达到 93.81% 和第 49 个 epoch 的准确率达到 93.62 时,训练集和验证集分别获得最高识别准确度。

为了找到理想的训练样本数量,图 6(b)给出了选用不同样本数量进行训练时训练集和验证集的准确度,实验中其他参数的设置不变。可以看出,样本数量达到 2000 时验证集的准确度达到 92.4%。因此,为了训练一个性能较好的 CNN 分类器,应至少选取样本数量为 2000 及以上的样本集。

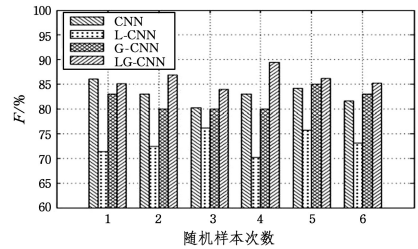
4.4 分类识别的对比分析

初步提取的 LBP 和 Gabor 图像特征经过 LG-CNN 算法进一步得到了细化^[22],然后由模型最后一层分类输出得到一个 5 维的概率向量,最大概率值对应的分类即为算法的识别

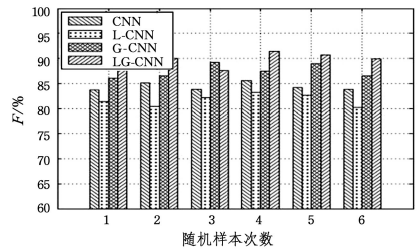
结果。在上述实验的基础上,将 LG-CNN 算法与其他分类识别算法进行比较。根据不同评价指标对所提算法进行对比分析,实验分两部分:1)对输入的分类识别算法的性能进行比较;2)对不同分类器的性能进行比较。

1) 与不同特征输入的识别算法的比较

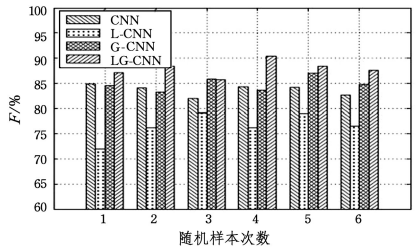
实验选取 2000 个随机样本进行 6 次网络训练,统计结果如图 7 所示。



(a) 4 种算法的精确度比较



(b) 4 种算法的召回率比较



(c) 4 种算法综合指标 F 的比较

图 7 不同输入算法的比较

Fig. 7 Comparison of different input algorithms

由图 7(a)可以看出,识别算法的准确度与样本的选取有关,训练样本的不同也会造成准确度的差异。但 LG-CNN 算法的准确度高于 L-CNN 算法和 G-CNN 算法,最高能达到 89.39%,这是因为 LG-CNN 算法综合了缺陷图像中关键的纹理特征,即 LBP 和 Gabor 特征,并通过 CNN 提取更深层次的特征,相较于单特征的分类器输入训练能够提取更多的细节信息,因而识别准确度高。从图 7(b)来看,L-CNN 算法的召回率偶尔低于 L-CNN 和 G-CNN 算法,这是因为后者基于单个特征输入的识别算法,考虑的细节因素不够全面,受一些干扰因素的影响,将最终分类结果错误识别成某一类缺陷,导致召回率提高。从图 7(c)可以看出,L-CNN 算法的综合指标 F 值也比其他几种算法高。综上所述,LG-CNN 算法的缺陷识别精确度比较理想,相比其他几种单特征输入的算法有明显优势。

2) 与不同分类器的性能比较

实验选用 1000 个随机样本,80% 的样本数据用作模型训

练,其中模型训练数据采用 20% 的数据进行交叉验证,20% 的样本数据用作模型结果测试。表 8 统计了各种算法的平均识别准确率、平均召回率、最差效果、最好效果以及综合指标 F 的实验数据。表 9 给出了不同分类器对于不同缺陷的分类识别率。

表 8 各种分类识别算法的准确率和召回率对比

Table 8 Comparison of accuracy rates and recall rates of various classification algorithms

分类识别算法	平均识别准确率和召回率	最差	最好	平均综合指标 F
LG-CNN	0.8839	0.8067	0.9263	0.9052
	0.9275	1	1	
CNN	0.8524	0.8316	0.9325	0.8766
	0.9023	0.8293	1	
SVM	0.8479	0.7195	0.9263	0.8720
	0.8975	0.9318	0.9329	
AdaBoost	0.8059	0.7317	0.9053	0.8139
	0.8220	0.7500	0.9250	
LDA	0.8337	0.7788	0.9100	0.8498
	0.8666	0.8537	1	

实验分析:基于特征组合和 CNN 模型的分类识别算法在 1000 次随机样本实验下的平均识别准确率为 88.39%,缺陷召回率为 92.75%,具有最优的结果,说明该模型在保证准确率的同时具有非常高的召回率,这样能够减少程序漏检的情况。

表 9 给出了实验中各分类识别算法对 5 种缺陷的平均分类识别率。基于组合特征的 LG-CNN 模型在不同缺陷上的分类准确率总体上优于其他分类器,尤其是在裂缝和渗漏点两种缺陷上,平均识别率分别达到 90.6% 和 89.98%。从表中可以看出,5 种方法对裂缝的识别率均最高,即裂缝容易与其他 3 种缺陷(渗漏点、渗水面和钙化物)区分;渗水和钙化物析出容易导致错误分类的主要原因是两者形状相似,均为不规则形状,再加上渗水后期也会有钙化物析出,因而识别难度较大。

表 9 不同识别算法对缺陷的平均分类识别率

Table 9 Average recognition rate of different recognition algorithms on defects

分类识别算法	(单位:%)				
	正常	裂缝	渗漏点	渗水面	钙化物
LG-CNN	87.65	90.60	89.98	83.21	84.75
CNN	87.50	92.83	84.30	80.20	84.16
SVM	88.13	91.40	79.18	80.34	85.62
LDA	85.43	88.20	86.50	78.20	80.10
AdaBoost	87.75	90.70	82.30	81.21	90.50

从上述实验可以看出,在参数设置合理的情况下,基于特征组合和 CNN 模型的分类识别算法相比传统的分类识别算法在准确率和召回率上都有明显提高,在缺陷分类识别上能达到较好的分类效果,且分类器训练时间较短,能在保证准确率的同时提高算法运行效率。

结束语 本文从大坝缺陷图像^[23]的特征分析入手,根据检测图像的时间和空间分布特点^[24],首先分别提取训练样本图像的 LBP 特征和 Gabor 特征图;接着,基于特征向量余弦

特征组合两种特征得到 LG 组合特征;然后将 LG 组合特征作为深度学习模型 CNN 的输入,通过逐层训练神经网络,提取更高层次的特征;在提取 LG 特征后将其输入到训练好的 CNN 分类器中,最终分类识别出缺陷类别。本文从图像特征提取入手,研究了监测图像缺陷的检测和识别技术,但后续还有很多需要研究和改善的地方,如提出的大坝缺陷图像特征提取与融合研究^[25-26]并没有考虑到缺陷检测和评判的实时性^[27],后期研究可以重点关注这方面。

参 考 文 献

- [1] 苏南. 我国 200 米级高坝密集,安全风险不容忽视[EB/OL]. http://www.thepaper.cn/newsDetail_forward_1858088. 2017.
- [2] ZHANG S B, LIU N, HUO H, et al. Image classification based on hierarchical shape feature extraction model [J]. High-tech communication, 2016, 26 (1): 81-88. (in Chinese)
张盛博, 刘娜, 霍宏, 等. 基于层次形状特征提取模型的图像分类[J]. 高技术通讯, 2016, 26(1): 81-88.
- [3] DENG H B, JIN L W. Facial Expression Recognition Based on Local Gabor Filter Bank and PCA+LDA[J]. Journal of Image & Graphics, 2007, 12(2): 322-329. (in Chinese)
邓洪波, 金连文. 一种基于局部 Gabor 滤波器组及 PCA+LDA 的人脸表情识别方法[J]. 中国图象图形学报, 2007, 12(2): 322-329.
- [4] BELARBI M A, MAHMOUDI S, BELALEM G. PCA as Dimensionality Reduction for Large-Scale Image Retrieval Systems [J]. International Journal of Ambient Computing and Intelligence (IJACI), 2017, 8(4): 45-58.
- [5] ZHANG R, WANG Q, ZHANG P. Image Texture Feature Extraction Based on Gabor Transform [J]. Revista Tecnica de la Facultad de Ingenieria Universidad del Zulia, 2016, 39(1): 237-244.
- [6] RAHMAN S, RAHMAN M M, HUSSAIN K, et al. Image enhancement in spatial domain: A comprehensive study [C] // 2014 17th International Conference on Computer and Information Technology (ICCIT). IEEE, 2014: 368-373.
- [7] XIE F, CHEN L T, QIU H. Research on image classification method based on texture feature extraction and system implementation [J]. Computer Application Research, 2009, 26 (7): 2767-2770. (in Chinese)
谢菲, 陈雷霆, 邱航. 基于纹理特征提取的图像分类方法研究及系统实现 [J]. 计算机应用研究, 2009, 26(7): 2767-2770.
- [8] ZARÁNDY Á, REKECZKY C, SZOLGAY P, et al. Overview of CNN research: 25 years history and the current trends [C] // 2015 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE, 2015: 401-404.
- [9] ZHAO J, TONG W Q. Near infrared face recognition based on LBP and Gabor hybrid features [J]. Journal of East China Normal University (Natural Science Edition), 2016 (4): 77-85. (in Chinese)
赵骥, 童卫青. 基于 LBP 和 Gabor 混合特征的近红外人脸识别 [J]. 华东师范大学学报(自然科学版), 2016(4): 77-85.

- [10] JIANG Z, WU Q, TAN Y, et al. Analysis Range of Coefficients in Learning Rate Methods of Convolution Neural Network[C]// International Symposium on Distributed Computing & Applications for Business Engineering & Science, 2016.
- [11] GAO Y, MA J, ZHAO M, et al. NDDR-CNN: Layer-wise Feature Fusing in Multi-Task CNN by Neural Discriminative Dimensionality Reduction [J/OL]. <https://arxiv.org/pdf/1801.08297.pdf>.
- [12] TAN W X, ZHAO C, WU H, et al. Intelligent alerting for fruit-melon lesion image based on momentum deep learning[J]. *Multimedia Tools & Applications*, 2016, 75(24):1-21.
- [13] PENG C, XIAO T, LI Z, et al. MegDet: A Large Mini-Batch Object Detector[C]// IEEE Conference on Computer Vision and Pattern Recognition, 2018:6181-6189.
- [14] ELLEUCH M, MAALEJ R, KHERALLAH M. A New Design Based-SVM of the CNN Classifier Architecture with Dropout for Offline Arabic Handwritten Recognition[J]. *Procedia Computer Science*, 2016, 80:1712-1723.
- [15] LIU W, WEN Y, YU Z, et al. Large-margin softmax loss for convolutional neural networks[C]// International Conference on International Conference on Machine Learning, 2016.
- [16] NIE X, CHENG A, SHEN D, et al. New Algorithm SGD of Surface Modification and Its Application in Springback Compensation[J]. *China Mechanical Engineering*, 2008, 24(19):2972-2976.
- [17] CHERKASSKY V, MA Y. Practical selection of SVM parameters and noise estimation for SVM regression[J]. *Neural Networks*, 2004, 17(1):113-126.
- [18] ZHANG Y, ZHOU X, WITT R M, et al. Dendritic spine detection using curvilinear structure detector and LDA classifier[J]. *Neuroimage*, 2007, 36(2):346-360.
- [19] RÄTSCCH G, ONODA T, MÜLLER K R. Soft Margins for AdaBoost[J]. *Machine Learning*, 2001, 42(3):287-320.
- [20] CHEN S, COWAN C N, GRANT P M. Orthogonal least squares learning algorithm for radial basis function networks[J]. *IEEE Transactions on Neural Netw*, 1991, 64(5):829-837.
- [21] STOEAN C. In Search of the Optimal Set of Indicators when Classifying Histopathological Images[C]// International Symposium on Symbolic & Numeric Algorithms for Scientific Computing, 2017.
- [22] BAG S, HARIT G. A medial axis based thinning strategy and structural feature extraction of character images[C]// IEEE International Conference on Image Processing, 2010.
- [23] FAN X, WU J, SHI P, et al. A novel automatic dam crack detection algorithm based on local-global clustering[J]. *Multimedia Tools & Applications*, 2018, 77(20):26581-26599.
- [24] KAWASAKI Y, MATSUSHIMA K, ZHONG Z. Image-based pavement crack detection by percolation theory[C]// 2017 9th International Conference on Information Technology and Electrical Engineering (ICITEE), IEEE, 2017:1-6.
- [25] YU N, QIU T, FENG B, et al. Image Features Extraction and Fusion Based on Joint Sparse Representation[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2011, 5(5):1074-1082.
- [26] WANG Z, CUI P, LI F, et al. A data-driven study of image feature extraction and fusion [J]. *Information Sciences*, 2014, 281(4):536-558.
- [27] TYCHSEN-SMITH L, PETERSSON L. DeNet: Scalable Real-time Object Detection with Directed Sparse Sampling[C]// 2017 IEEE International Conference on Computer Vision (ICCV). IEEE Computer Society, 2017.