



中国公路学报  
*China Journal of Highway and Transport*  
ISSN 1001-7372,CN 61-1313/U

## 《中国公路学报》网络首发论文

题目: 基于卷积自编码的沥青路面目标与裂缝智能识别  
作者: 侯越, 陈逸涵, 顾兴宇, 茅荃, 曹丹丹, WANG Linbing, 荆鹏  
网络首发日期: 2020-08-18  
引用格式: 侯越, 陈逸涵, 顾兴宇, 茅荃, 曹丹丹, WANG Linbing, 荆鹏. 基于卷积自编码的沥青路面目标与裂缝智能识别. 中国公路学报.  
<https://kns.cnki.net/kcms/detail/61.1313.U.20200818.1354.002.html>



**网络首发:** 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

**出版确认:** 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

**基金项目：**国家自然科学基金（51708026），科技部重点研发计划（2017YFF0205600），北京工业大学国际科研合作种子基金项目（No. 2018A08），北京市科技创新服务能力建设-基本科研业务费（科研类PXM2019\_014204\_500032）

**作者简介：**侯越（1987-），男，江苏东台人，副教授，工学博士，E-mail: yuehou@bjut.edu.cn

**\*通讯作者：**顾兴宇，教授，博士生导师，东南大学交通学院，南京 211189，guxingyu1976@163.com

**Funding:** National Science Foundation of China (51708026); the National Key R&D Program of China (No. 2017YFF0205600); the International Research Cooperation Seed Fund of Beijing University of Technology (No. 2018A08); the Construction of Service Capability of Scientific and Technological Innovation-Municipal Level of Fundamental Research Funds (Scientific Research Categories) of Beijing City (Funding No. PXM2019\_014204\_500032).

\* Corresponding author. E-mail address: Xingyu Gu, Professor, School of Transportation, Southeast University, Nanjing 211189, China; guxingyu1976@163.com

# 基于卷积自编码的沥青路面目标与裂缝智能识别

侯越<sup>1</sup>，陈逸涵<sup>1</sup>，顾兴宇<sup>\*2</sup>，茅荃<sup>3</sup>，曹丹丹<sup>1</sup>，  
WANG Linbing<sup>4</sup>，荆鹏<sup>1</sup>

(1. 北京工业大学 北京市交通工程重点实验室，北京 100124;

2. 东南大学 交通学院，江苏 南京 210096;

3. 江苏现代路桥有限责任公司，南京 210096;

4. 弗吉尼亚理工大学 土木工程与环境工程系，弗吉尼亚 黑堡 VA24061;)

**摘要：**目前基于深度学习的路面裂缝识别经常面临训练数据集小，以及路面图片标注成本高等问题，本文基于小规模路面图片数据集，利用卷积自编码(CAE)方法进行数据增强，开展包括路面裂缝在内的路面目标智能化识别方法研究。在传统图像几何变换数据增强的基础上，采用CAE 重构图片方法对原始数据集进行两步骤扩增；利用卷积神经网络DenseNet，设置了不同数据扩增方法的对比实验；针对沥青路面裂缝图片背景较黑、裂缝特征不清晰、无监督聚类难度大等问题，采用了一种基于CAE 预训练的深度聚类算法DCEC，对经数据增强的路面图片进行无标注的聚类识别。研究表明：经过DenseNet 网络100代的训练，在同一测试集的测试下，基于原始数据集训练的网络分类准确率为78.43%，利用传统图像处理方法进行扩增后准确率为83.44%，利用本文所使用的图片增强方法进行数据扩增后准确率达87.19%；在保持扩增后数据集样本量大小相同的情况下，与几何变换、像素颜色变换等经典数据增强手段相比，CAE 重构图片的数据扩增方法有较高的路面图片识别精度；CAE 数据扩增方法较受训练数据集样本量的影响，利用传统方法将数据集扩增后进行CAE 特征学习，重构后的图片样本更易被机器识别；相较于传统机器学习聚类算法，本文所采用的DCEC 深度聚类方法

将聚类准确率提升了约 10%，初步实现了无需人工标注的路面目标的端到端智能识别。

**关键词：**道路工程；路面裂缝检测；深度学习；卷积自编码器；深度聚类；数据增强

## Automatic Identification of Pavement Objects and Cracks using the Convolutional Auto-encoder

HOU Yue<sup>1</sup>, CHEN Yi-han<sup>1</sup>, GU Xing-yu<sup>\*2</sup>, MAO Quan<sup>3</sup>, CAO Dan-dan<sup>1</sup>,  
WANG Lin-bing<sup>4</sup>, JING Peng<sup>1</sup>

1. College of Metropolitan Transportation, Beijing University of Technology, Beijing 100124, China; 2. School of Transportation, Southeast University, Nanjing 210096, Jiangsu China; 3. Jiangsu Xiandai Road & Bridge Co., Ltd, Nanjing 210096, Jiangsu China; 4. Department of Civil and Environmental Engineering, Virginia Polytechnic Institute and State University, Blacksburg VA24061, Virginia, USA

**Abstract:** The automatic detection of pavement cracks can significantly improve the efficiency of road maintenance for pavement engineers. At present, the Artificial Intelligence-based pavement crack detection may have the problems of insufficient training dataset or large dataset of pavement images that require high costs for manual classification and labeling. To solve these problems, based on the small-scale pavement image dataset, a study on the applications of the supervised and unsupervised deep learning models using the Convolutional Auto-encoder(CAE) method was conducted to identify different pavement objects, including the pavement cracks. Based on the traditional data augmentation using geometry transformation, comparison tests based on different data augmentation methods were conducted to validate the accuracy of the proposed research. Considering the problems that the background of the asphalt pavement crack image is dark, the crack characteristics are not clear, and the unsupervised clustering is difficult, a deep clustering algorithm DCEC (Deep Convolutional Embedded Clustering) based on CAE pre-training is proposed to study the road images. Test results show that: after 100 iterations of DenseNet network training, under the same test set, the test accuracy of network classification based on the original data set is 78.43%, the test accuracy based on the traditional data augmentation using image transformation method is 83.44%, and the test accuracy based on the method proposed in this study is 87.19%. It can also be found that, under the same dataset sample size, compared with the traditional data augmentation methods such as geometric transformation and pixel color transformation, the data augmentation method using CAE reconstruction has higher recognition accuracy. Results show that the CAE data augmentation method is more easily affected by the quality and sample size of the training data set. After the data set is augmented by the traditional method, CAE learning is then carried out, and the reconstructed image sample is more easily recognized. Compared with the traditional method, the DCEC deep clustering method can improve the accuracy of clustering by about 10%, which preliminarily realizes the end-to-end intelligent recognition of road targets without manual annotation.

**Keywords:** pavement engineering; pavement crack detection; deep learning; convolutional auto-encoder; deep clustering; data augmentation

中图分类号：U416.217

文献标志码：A

## 0 引言

车辆荷载和环境等因素引起的裂缝、坑槽、车辙等病害降低了行车的安全性和舒适性，需要道路工作者定期对路面性能进行检测和评价，采用宏观、微观以及数学分析等方法分析

材料性能与病害成因,从而及时对路面进行养护。传统的路面病害人工检测方法效率低,难以满足现代化智能路面检测的需要,基于路面自动化检测设备进行路面图像采集与病害识别是大势所趋。

受益于人工智能技术的发展,机器学习方法被逐渐应用于路面病害识别领域。机器学习算法的核心在于通过算法学习数据、解析数据、训练模型,从而预测未知的数据。根据输入数据是否带有标签(标注/未被标注)可分为有监督学习和无监督学习。传统有监督学习的机器学习方法包括支持向量机(SVM)、随机森林(RF)等,已被广泛应用于路面图片的分类与检测中。比如 Fujita 等<sup>[1]</sup>利用移动测绘系统获取图像,将图像的多尺度特征用于 SVM 分类器的训练,实现了像素级路面裂缝图像的分类。Pan 等<sup>[2]</sup>利用四种不同的核函数的 SVM 模型对无人机采集的路面光谱图像进行智能检测。

传统机器学习方法的缺点在于需要人工提取图像特征,无法实现路面图片端到端的智能识别,为克服该缺点,机器学习中发展出了以卷积神经网络(Convolutional Neural Network, CNN)为代表的深度学习方法,该算法能够避免复杂的人工图像特征提取过程,自动进行特征学习,因此被广泛应用于图像分类、目标识别等任务中。近几年,深度学习方法逐渐被引入道路领域并取得较多成果。Zhang 等<sup>[3]</sup>基于卷积神经网络提出了 CrackNet 裂缝识别网络,实现了三维沥青路面的像素级裂缝检测。沙爱民等<sup>[4]</sup>利用 3 个 CNN 模型分别完成路面病害识别、路面裂缝特征提取和坑槽特征提取,计算准确度较高。Dorafshan 等<sup>[5]</sup>将 CNN 作为小尺寸图像的分类器,结合滑动窗法并采用迁移学习,将经典的深度卷积神经网络 AlexNet 移植到路面检测应用领域,实现大尺寸水泥混凝土路面破损图像的裂缝检测。

基于深度学习的算法模型往往含有大量参数,较大样本量的数据集更加有利于训练学习,从而构建出更可靠的模型。当前道路工程领域的研究者们能获得的路面图片数据集样本数量一般比较有限。针对于数据量较小、数据质量较差的数据集,经常采用数据增强(Data Augmentation)的方法对数据集进行扩充。深度学习中的数据增强算法可概括为基本的图像处理方法和基于深度学习的数据增强方法<sup>[6]</sup>:(1)传统的基于图像处理的数据增强方法。主要包括三类<sup>[7]</sup>,其一,图片的几何变换方法,如对图片进行裁剪、旋转、镜像操作;其二,图片的色彩空间变换方法,如调整图片亮度、对图片添加噪声;其三,为解决机器学习中样本不平衡问题而产生的多样本合成算法,如 SMOTE 算法<sup>[8]</sup>。(2)基于深度学习的数据增强方法。其中最为典型的算法是生成式对抗网络(Generative adversary network, GAN)<sup>[9]</sup>,利用生成器与判别器两部分网络对样本图片进行特征学习,并重新生成与训练样本分布一致的较为逼真的图片。该方法受到计算机界的广泛关注,但目前道路桥梁领域应用 GAN 进行数据扩增的研究较少。Zhang 等<sup>[10]</sup>提出一种 CrackGAN 以解决路面裂缝图片分割时的样本不平衡等问题。李良福等基于 GAN 扩充桥梁裂缝图片数据集以缓解裂缝分割时的欠拟合问题<sup>[11]</sup>,利用 GAN 进行图片超分辨率重建提升了桥梁裂缝分割精度<sup>[12]</sup>。生成式对抗网络的训练过程复杂,且计算机计算成本较为庞大<sup>[13-14]</sup>。除了 GAN,另外一种数据增强的工具是卷积自编码器。2017 年 Lore 等<sup>[15]</sup>提出低光网络 LLNet,将自编码器(auto-encoder, AE)的方法引入图像增强领域,该网络通过学习低光网络特征,并利用自编码器网络的去噪能力,重新生成明亮的图像<sup>[15-16]</sup>。同时,在道路工程领域,研究者们多将自编码器的方法直接用于路面图片分类。Shi 等<sup>[17]</sup>提出了一种图像处理与多层自编码器相结合的方法对路面裂缝图片进行识别,钱彬等<sup>[18]</sup>应用稀疏自编码器对图像进行特征学习后利用 softmax 分类器识别路面裂缝图片。到目前为止,将卷积自编码器方法应用于道路图像增强的研究仍然较少。

目前在道路目标识别中,常用的方法是有监督学习。但是即使获得大样本量的数据集,该方法也仍然存在数据集人工标注成本高的缺点,常常要花费数月时间创建一个训练数据集。因此,以聚类 and 降维为主要手段的无监督学习方法受到日益重视。聚类方法根据一定的特征对对象进行分组,而数据降维通过找到共同点来减少数据集的变量,使得无监督学习的



训练样本不需要带有标签,可实现样本的自动分类及标注。典型的聚类算法如 k-means 聚类方法、层次聚类方法、谱聚类方法等,在计算机视觉领域中已有较为深远的研究,也有部分研究者将其应用于道路工程路面目标识别领域,但大多将聚类算法作为图片预处理手段<sup>[22]</sup>或将其作为裂缝图片的语义分割手段<sup>[19-20]</sup>,并未将其直接用于路面图片的分类识别任务中。自编码器也属于无监督深度学习中降维算法的一种,近年来,计算机视觉领域的研究者将以 k-means 代表的传统聚类学习算法与以自编码器为代表的深度学习算法相结合,提出了深度聚类(Deep Clustering)的算法,使得在原始图像上进行大规模端到端的聚类成为可能<sup>[21-25]</sup>。

综上,目前道路病害智能识别中存在如下研究现状:(1) 经常面临训练数据集小的问题,但是基于深度学习的数据增强研究较少;(2) 路面图片标注成本较高,但是低成本无监督学习的研究较少。为此,本文基于小样本路面图像数据集,应用卷积自编码器(Convolutional Auto-encoder, CAE)方法,进行如下研究:首先,对原始图片采用批量图像处理方法进行数据扩增,结合卷积自编码器法进行图像特征学习并重构图像,实现原始数据集的有效扩增。其次,基于卷积神经网络 DenseNet(Dense Convolutional Network) 的有监督学习,测试了不同数据增强方法下的模型分类精度,验证了 CAE 数据扩增方法的有效性。最后,利用一种基于 CAE 预训练的无监督深度聚类算法 DCEC(Deep Convolutional Embedded Clustering),将数据增强后形成的大样本量路面图片删去标签信息进行聚类学习,并将该方法所得结果与传统机器学习方法结果进行对比,初步探讨了无监督深度聚类方法在道路裂缝等目标智能识别中的可行性。

# 1 研究框架及算法

## 1.1 研究框架

本文研究路线如图 1 所示。

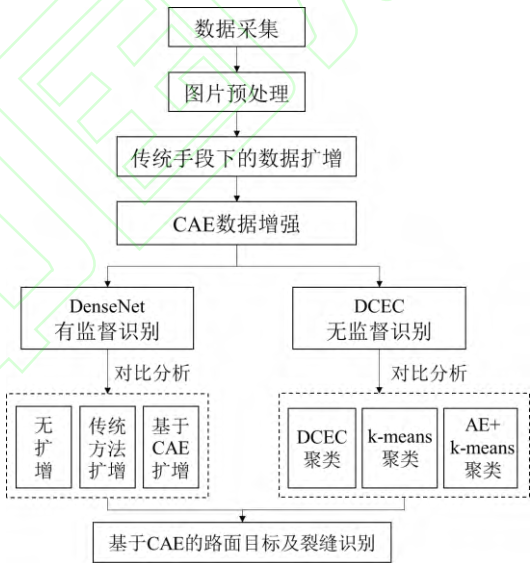


图 1 研究路线

Fig.1 Flowchart of This Study

## 1.2 有监督算法

### 1.2.1 卷积神经网络

卷积神经网络是一种包含卷积计算的深度学习算法,由于其卷积核参数共享和层间局部连接的特点使得卷积神经网络能够以较小的计算代价完成复杂的特征学习任务。卷积神经网络的组成部分包括输入层,隐含层和输出层。

(1) 输入层：在处理图像问题时，输入层为代表图片的长×宽×深的三维矩阵，长和宽即图像的大小，矩阵的深度表示为图像的色彩通道数，黑白图片的深度为 1，RGB 彩色图片的深度为 3，每个像素点以灰度值表示。

(2) 隐含层：通常由卷积层、池化层和全连接层三个部分构筑而成。

卷积层带有权重参数的  $k$  个尺寸为  $l \times l$  的卷积核，卷积核以特定步长(stride) $s$  扫过大小为  $w \times h$  的输入图像，与逐个对应区域的像素值做点积。卷积层可以看做是特征提取的过程，提取结果表示为特征图(feature map)，特征图尺寸为  $((h-l)/(s+1)) \times ((w-l)/(s+1)) \times k$ 。通常利用激励函数对卷积层输出结果进行一次非线性变换，以协助模型表达复杂特征。早期的网络多采用 sigmoid 激励函数，如式(1)，sigmoid 函数具有良好的解释性，函数的输出值在(0,1)之间，但在深层网络中表现不佳。卷积神经网络中通常使用如式(2)所示的 ReLU 函数<sup>[26]</sup>，该函数的导数计算更快，使得训练更易收敛。

$$f(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

$$f(x) = \max(0, x) \quad (2)$$

式中： $x$  表示激活函数的输入信号值。

池化层在一定大小的采样窗口中，在该邻域内对特征图进行特定规则的下采样。通常在池化层采样过后，特征图深度不变，长度和宽度减半，从而降低网络计算量和所需内存。在实际应用过程中，较为典型的池化方法如最大池化与平均池化。最大池化指在池化操作时对邻域内特征点取最大值，平均池化即对邻域内特征点求平均。

全连接层位于隐含层的最后部分，通常  $w \times h \times k$  的特征图在传输到全连接层前将所有特征展平成  $N \times 1$  的向量，其中  $N = k \times w \times h$ 。该向量会与全连接层的  $T \times N$  的权重向量作点积再加上一个偏置，将其传递至激励函数产生该单元的状态，得到  $T \times 1$  的向量，其中  $T$  表示多分类类别数量，最后将信号传递至输出层。在有些网络中，会采取全局平均池化代替全连接层，在上一层的特征图数量往往为类别数量，对每幅特征图取平均值，展平得到  $T \times 1$  的向量，能够大幅减少参数，同时减少过拟合现象<sup>[27]</sup>。

(3) 输出层：在多分类任务中，输出层通常采用 Softmax 的结构输出分类标签。经过如式(3)所示的 Softmax 函数转换，重新得到一个  $T \times 1$  的向量，向量中每个值表示预测结果属于各个类别的概率，将概率最大的类别作为最终的输出结果。

$$f(x)_j = \frac{e^{x_j}}{\sum_{k=1}^K e^{x_k}} \quad (3)$$

式中： $j$  表示第  $j$  个类别； $x_j$  表示该类别的输入信号，该函数分母为所有输入信号的指数和。

常常运用批标准化(Batch Normalization)<sup>[28]</sup>的方法对神经元输入特征进行归一化处理，使得训练过程更易收敛，参与批标准化中每个 Batch 的样本数量称为 Batch size。

在有监督学习中，卷积神经网络使用反向传播的框架进行网络参数更新。在训练过程中，先进行正向传播，得到预测结果，通过损失函数比较预测结果和真值，再从后至前依次调整每一层的参数，不断迭代，直至损失函数收敛，模型训练结束。在训练过程中，更新参数的基本原理为梯度下降，即沿着梯度寻找一个函数的极小值。为使得反向传播算法更快收敛，一般会在网络训练时采用一定的优化算法对该过程进行优化。

### 1.2.2 DenseNet 图片分类网络

DenseNet<sup>[29]</sup> (Dense Convolutional Network)是较为经典的卷积神经网络之一，该网络的核心在于其密集块(Dense Block)结构，即网络中每一层的输入来自该结构中前面所有层的输出，该结构能够在保证模型中层与层之间信息传输最大化的同时，连接网络中的所有层，使

得每层网络学习后特征的传递更加有效，相较于传统的卷积神经网络，该结构更窄，参数更小。

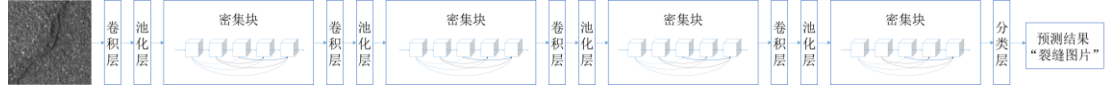


图 2 DenseNet 网络结构图<sup>[29]</sup>

Fig 2. Architecture of DenseNet<sup>[29]</sup>

本文以 DenseNet 作为各类数据增强方法的对比验证网络，所使用的 DenseNet 网络结构框架如图 2 所示。其中，密集块由多次  $1 \times 1$  和  $3 \times 3$  卷积操作构成，在  $3 \times 3$  卷积之前加入  $1 \times 1$  卷积的设计能够减少输入特征映射的数量，从而提高计算效率。两个密集块中间由  $1 \times 1$  卷积层及  $2 \times 2$  平均池化层连接过渡，能够起到降维作用，即经过过渡的卷积和池化层，输出图片尺寸变小，通道数降低。分类层由全局平均池化层和 SoftMax 层组成，获得图片的预测结果<sup>[29]</sup>。

### 1.2.3 有监督识别评价指标

对于各类数据增强算法的有监督分类效果评判，采用准确率(accuracy, ACC)作为评估指标，即分类正确的样本占总样本个数的比例。由于实验中各类别图片数量均衡，利用准确率能够较为简单直观的反应模型的有效性。计算公式如式(4)

$$ACC = \frac{n_c}{n_t} \quad (4)$$

式中： $n_c$ 为预测正确的样本数量； $n_t$ 为总的样本数量。

## 1.3 无监督算法

### 1.3.1 卷积自编码器

自编码器(Autoencoder, AE)是神经网络的一种，属于无监督学习算法，最早于 1986 年被提出，2006 年 Hinton 等<sup>[30]</sup>改进了自编码网络结构并将其用于 BP 神经网络的网络权重预训练阶段。该算法利用反向传播算法训练，使得输入尽可能等于输出。该网络可以由两部分组成：编码器(Encoder)和解码器(Decoder)，如图 3 所示。编码器将输入数据压缩为潜在空间表征，将高维数据压缩为低维数据，用  $f(x)$ 表示。解码器将空间表征重构为输出，相当于数据的解压过程，用函数  $g(x)$ 表示。其优化目标即最小化重构误差，如式(5)。

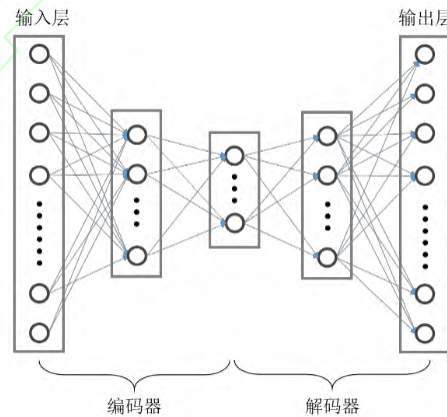


图 3 自编码器网络结构示意图<sup>[30]</sup>

Fig.3 Diagram of Module Structure of Autoencoder<sup>[30]</sup>

$$f, g = \arg \min_{f, g} L(x, f(g(x))) \quad (5)$$

传统的自编码方法使用全连接层，对于二维图像，全连接层会造成空间信息的损失，卷积自编码器(Convolutional Autoencoder, CAE)<sup>[31]</sup>将卷积神经网络结构融入自编码器中，利用卷积层代替全连接层，从而较好的保留图片特征信息<sup>[32]</sup>。其中，编码器中采用与卷积神经网络相似的卷积和池化操作，而在解码器中，采用卷积操作和上采样(Upsampling)操作，使得每一层的输出图片的尺寸大于输入图片的尺寸，将编码器提取的特征图还原到和原图同样尺寸的大小，实现图片重构。本文基于 CAE 方法，利用其编码器结构对路面原始图像进行数据降维并得到图片的特征空间，嵌入聚类网络，实现路面图片的无监督深度聚类；利用其解码器结构将图片重构，获得新的图片，扩充数据集，提高路面病害识别模型鲁棒性。

### 1.3.2 深度卷积嵌入聚类算法

深度卷积嵌入聚类算法 DCEC<sup>[24]</sup>是深度嵌入聚类算法(deep embedded clustering, DEC)<sup>[22]</sup>的改进算法，最大的不同在于 DEC 网络基于 AE，而 DCEC 网络基于 CAE。DCEC 网络由 CAE 网络和一个与 CAE 嵌入层相连接的聚类层(Clustering layer)组成，如图 2 所示。DCEC 中的 CAE 网络部分采用 3 层卷积编码层和 3 层卷积解码层，在编码层后接 Flatten 层，将编码器部分所学得的特征展开，再通过嵌入层 (Embedded layer)中的 10 个神经元，得到 10 维特征，后接全连接层(FC)，利用反卷积操作(DeConvolution)，使得解码器中将特征重构为与原始图片尺寸相同的重构图片。反卷积操作为卷积操作的逆过程，即在卷积操作前，通过填充输入图片的像素(padding)，使得输出图片尺寸大于输入图片尺寸，最终实现图片的重构。编码器和解码器共同训练，使得嵌入层  $h$  有效学习到图片的特征。CAE 特征学习完成后，由嵌入层的潜在表征后接聚类层，并在聚类层利用 k-means 算法进行网络初始化，通过最小化损失函数进行训练，直至得到最优的聚类结果<sup>[33]</sup>。

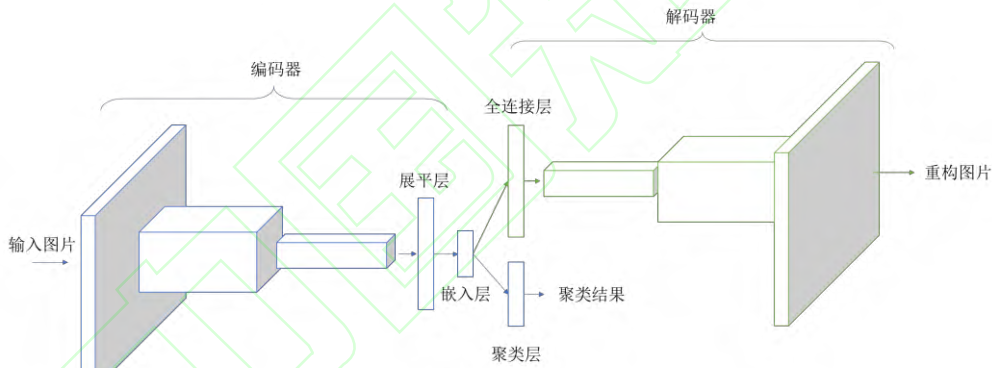


图 4 DCEC 网络框架<sup>[24]</sup>

Fig. 4 DCEC Structure Framework<sup>[24]</sup>

DCEC 网络结构中的损失函数如式(6)所示<sup>[23-24]</sup>，由重建损失  $L_r$  和聚类损失  $L_c$  构成。

$$L = L_r + \gamma L_c \quad (6)$$

式中： $\gamma > 0$ ，表示嵌入空间扭曲程度控制系数。

重建损失表示数据集中的样本在经过 CAE 结构后的均方误差(MSE)，定义为式(7)<sup>[22-24]</sup>

$$L_r = \frac{1}{n} \sum_{i=1}^n \|G_{\omega'}(F_{\omega}(x_i)) - x_i\|_2^2 \quad (7)$$

式中： $n$  是数据集中的图片数量， $x_i$  为第  $i$  张输入模型的原始图片。 $F_{\omega}$ 、 $G_{\omega'}$  分别为编码函数和解码函数， $Z_i = F_{\omega}(x_i)$  表示为第  $i$  张图片经过编码器结构得到的潜在空间特征， $x_i' = G_{\omega'}(Z_i)$  表示为潜在空间特征经过解码器结构得到的重建数据。

聚类损失如式(8)所示，以 KL 散度来表示，用以衡量目标分布  $P$  与拟合分布  $Q$  之间匹配程度。二者差异越大，则 KL 散度越大。若二者完全匹配，则 KL 散度为 0<sup>[22-24]</sup>



$$L_c = KL(P\|Q) = \sum_i \sum_j p_{ij} \lg \frac{p_{ij}}{q_{ij}} \quad (8)$$

式中:  $q_{ij}$  是第  $i$  个样本  $q_i$  属于第  $j$  类的概率, 它表示了原始特征分布经过编码层结构后位于潜在特征空间的嵌入点  $Z_i$  和聚类中心  $\mu_j$  的相似性, 定义为式(9)<sup>[22-24]</sup>

$$q_{ij} = \frac{\left(1 + \|z_i - \mu_j\|^2\right)^{-1}}{\sum_j \left(1 + \|z_i - \mu_j\|^2\right)^{-1}} \quad (9)$$

目标分布  $p_{ij}$  由  $q_{ij}$  决定, 定义为式(10)<sup>[22-24]</sup>

$$p_{ij} = \frac{q_{ij}^2 / \sum_i q_{ij}}{\sum_j \left(q_{ij}^2 / \sum_i q_{ij}\right)} \quad (10)$$

### 1.3.3 无监督识别评价指标

无监督识别结果使用聚类准确率(accuracy, ACC)、标准互信息(Normalized Mutual information, NMI)和调整 Rand 指数(adjusted rand index, ARI)为指标评价聚类模型的效果<sup>[34-35]</sup>。其中, ACC 即正确分类的样本占所有样本的百分比, 取值范围为[0,1], 计算公式与式(4)类似。NMI 与 ARI 都是衡量聚类结果和实际情况的数据分布的匹配程度的指标, 值越大, 表示聚类结果与实际情况更加吻合。计算一组数据聚类结果类别划分和真实类别划分之间 NMI 值的计算公式见式(11)<sup>[36-37]</sup>, 其取值范围为[0,1]。

$$NMI = \frac{2I(X,Y)}{H(X) + H(Y)} \quad (11)$$

式中:  $I(X,Y)$  为聚类结果划分  $X$  与实际划分  $Y$  的互信息量; 用来评价两种划分结果的一致性,  $H(X)$  和  $H(Y)$  分别代表两种划分的熵。

ARI 的取值范围为[-1,1]。ARI 基于  $RI$ <sup>[38]</sup>(Rand Index, 兰德系数)产生, 如式(12)

$$RI = \frac{a+b}{C_2^n} \quad (12)$$

式中: 分子表示聚类结果与实际情况属性一致的样本对数,  $a$  表示聚类结果与实际情况中都为同一类的元素对数,  $b$  表示聚类结果和实际情况都为不同类的元素个数; 分母表示将聚类划分和实际划分的两个数据集的元素进行组合, 共有几种组合方式;  $n$  代表数据集中的样本量。

ARI 对  $RI$  进行了调整, 使得聚类结果所得分数具有更高的区分度, ARI 的计算公式如式(13)<sup>[39]</sup>

$$ARI = \frac{RI - E[RI]}{\max(RI) - E[RI]} \quad (13)$$

## 2 数据采集与图片处理

### 2.1 数据采集与预处理

本文所采用的路面图像集来源于江苏省部分沥青路面的巡检结果, 采集的原始数据为  $4,096 \times 2,000$  像素的 RGB 图像。本研究的计算平台为一台移动计算工作站, 其处理器为 Intel Core i7-7700 CPU, 显卡为 NVIDIA Quadro P3200 with Max-Q Design GPU, 内存 32.0GB。算

法模型基于 Python 语言实现。有研究表明深度学习的不同架构会对病害识别结果产生影响<sup>[40]</sup>，为方便计算，本文所使用的深度学习框架统一采用以 Tensorflow 为后端的 Keras 框架。

为使数据集满足训练要求，对原始图像进行预处理。首先，由于受到计算机内存限制，需要对原始图片进行批量剪裁，解决图片过大难以训练的问题。但过小的图片不足以包含路面目标的特征，因此对图片进行先裁剪，后缩放的处理。本文通过双线性插值进行横向缩放至  $4,000 \times 2,000$  像素，对缩放后的图片进行连续剪裁，横纵向图片裁剪步长分别为  $step_x=400$ ,  $step_y=400$ , 得到  $400 \times 400$  像素的子图像，并将子图像缩放至  $200 \times 200$  像素。为平衡数据集各类别样本图片数量，筛选出 400 张路面背景图像，400 张路面裂缝图像，400 张路面标线图像和 400 张路面修补裂缝图像，共 1600 张图片，删去图片的彩色信息，并将图片转换为单通道的灰度图，作为原始数据集，各类别示例图片如图 5 所示。受采集条件、光照条件等额外因素影响，原始数据集中包含有图片曝光程度良好(如图 5 中第一排样本图片)和曝光程度较差(如图 5 中第二排样本图片)的路面图片。

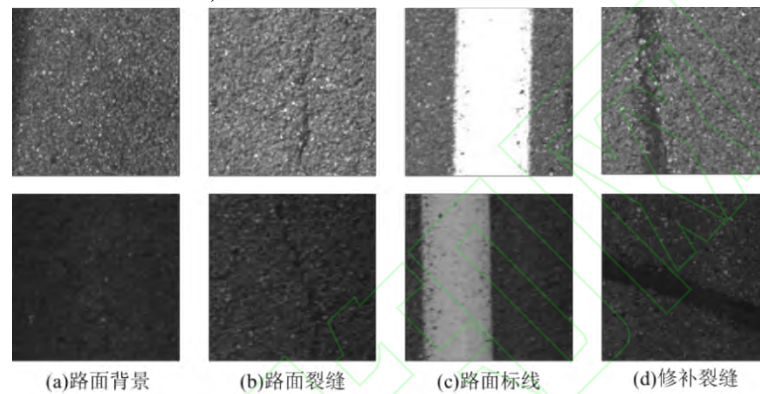


图 5 原始数据集中各类别路面图片

Fig. 5 Different Classes of Pavement Surfaces Images in Original Dataset

## 2.2 基于图像处理的数据扩增

由于神经网络参数量十分庞大，使得这些参数的训练往往需要基于大量数据。数据增强是深度学习中的常用提升数据量的步骤，目的在于通过扩大数据集的规模，获得更加稳健的网络，提高模型的泛化能力。本文的原始数据集样本量较小，直接用于训练模型将导致模型鲁棒性较差，泛化能力较弱。本文将原始数据集的 80%，共 1280 张图片作为训练集，原始数据集的 20%，共 320 张作为测试集。对于训练集中的 1280 张图片，首先采用传统数据扩增方法，运用图像批量处理的手段对原始数据集进行数据增强：第一步，将每张图片进行垂直镜像变换，如图 6(a)所示，第二步，对于原始图片及镜像图片，分别进行  $90^\circ$ 、 $180^\circ$  和  $270^\circ$  旋转变换，如图 6(b)(c)(d)所示。

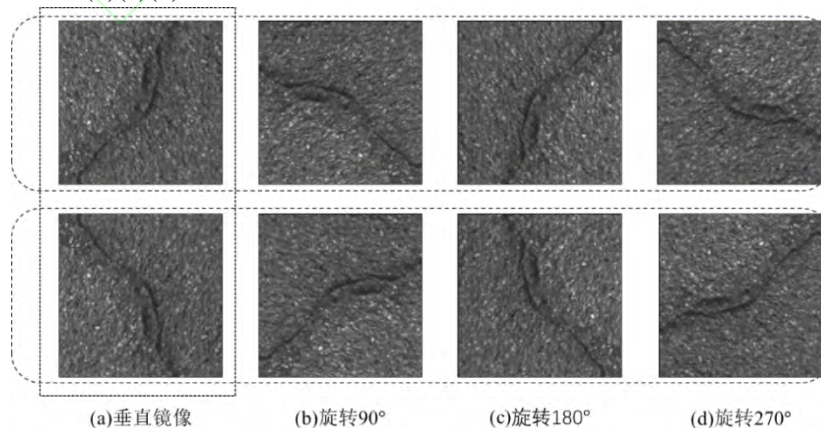


图 6 基于图像处理的数据扩增示意图

Fig. 6 Diagram of Data Augmentation Based on Image Processing Methods

数据增强后，数据总量扩大至原始数据训练集的 8 倍，共 10240 张图片，记为 Geometric 数据集，包括路面背景、路面裂缝、路面标线、路面修补裂缝图片各 2560 张，并以此为基础，利用 CAE 图片重构方法生成新的数据集。

### 3 CAE 图片重构

#### 3.1 网络结构与参数设置

本文用于数据扩增的 CAE 网络如图 7 所示，由 2 层编码层和 2 层解码层构成。当图片输入时，先经过编码器，进行如下操作：1) 编码层 1：首先进行卷积核尺寸为  $3 \times 3$ 、个数为 16 的卷积操作，以 ReLU 作为激活函数，后接一个步幅为 2 的最大池化层；2) 编码层 2：进行卷积核尺寸为  $3 \times 3$ 、个数为 32 的卷积操作，同样以 ReLU 作为激活函数，再次经过一个步幅为 2 的最大池化层。此时，编码器已学得图片特征，表示在尺寸为  $50 \times 50 \times 8$  特征图中，随后在解码器中进行图片重构：1) 解码层 1：首先将特征图进行卷积核尺寸为  $3 \times 3$ 、个数为 8 的卷积操作，经过 ReLU 激活函数后，经过一个步幅为 2 的上采样层(UpSampling)。2) 解码层 2：进行卷积核尺寸为  $3 \times 3$ 、个数为 16 的卷积操作，经过 ReLU 激活函数后，经过一个步幅为 2 的上采样层。此时，特征图尺寸还原为  $200 \times 200 \times 16$ ，后经过输出层。在输出层中仍采用尺寸为  $3 \times 3$  的卷积核，为了使得输出图片获得深度为 1 的灰度图，卷积核个数为 1，并且当输入小于 0 时，ReLU 激活函数的输出值为 0，在图像重构时可能造成部分像素损失，因此输出层采用 sigmoid 激活函数，最终获得尺寸为  $200 \times 200 \times 1$  的路面图片。

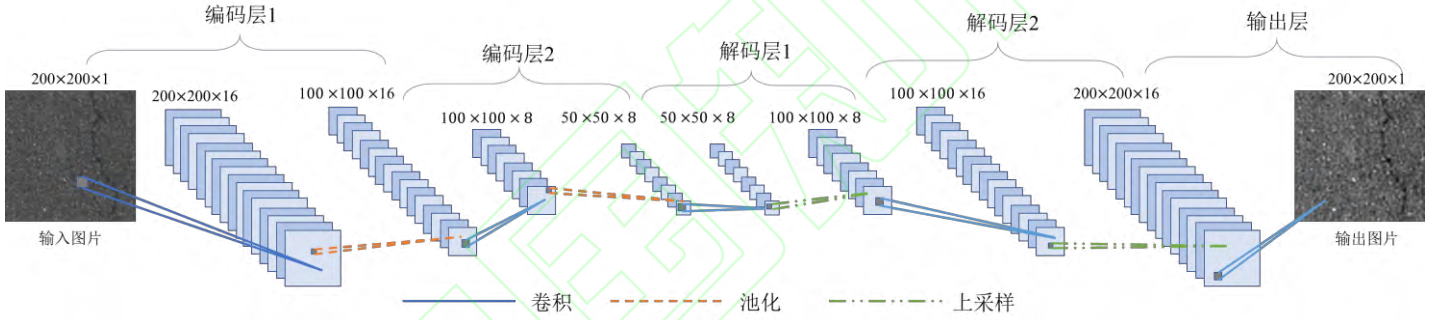


图 7 本研究所采用的 CAE 结构

Fig. 7 CAE Structure Used in This Study

在卷积自编码器的训练过程中，编码器和解码器共同训练，优化目标为最小化重构误差。以自适应学习率调整算法 Adadelata (Adaptive Learning Rate Method)<sup>[41]</sup>作为反向传播过程中梯度下降的优化算法，该算法的优势在于能够自适应调整梯度下降时的学习速率，而不需要手动设置。注意到 Adadelata 和 Adam<sup>[42]</sup>等算法均可采用，本文在图片重构部分采用 Adadelata 优化算法是基于初步试算的结果，而在 DenseNet 和 DCEC 部分采用了 Adam 算法。在未来的研究中将会对优化算法的选取进行深入的研究。

#### 3.2 图片重构结果

将 Geometric 数据集输入 CAE 网络，经过 1000 代的充分训练后，重新生成 10240 张图片，记为 GeoCAE。部分重构图片如图 8 所示。



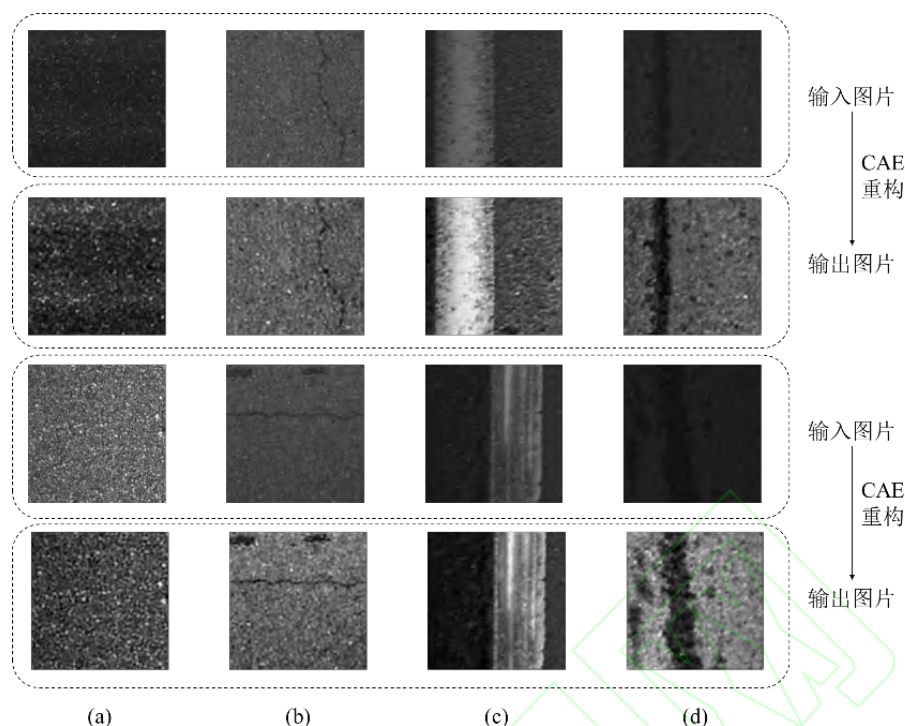


图 8 部分 CAE 输入图片与重构图片:(a)路面背景, (b)裂缝, (c) 路面标线, (d) 修补裂缝  
Fig. 8 CAE Input Images and Generated Images :(a) Pavement Background, (b) Crack, (c) Pavement Marking, and (d) Sealed Crack

对比图 8 中的输入图片和经 CAE 重构后输出的图片, 由于沥青路面为黑色, 受路面纹理本身因素以及采集时间、天气、光照条件等外部因素影响, 路面图片的背景噪声较大, 数据集中部分路面图片难以辨认。而经 CAE 重构, 输出图片的图片特征相对更加明显, 本文所使用的 CAE 方法起到了较好的路面图片去噪和增光效果。

并且, 在本文的数据集中, 部分图片曝光条件不良, 同时裂缝占整张图片的比例较小, 导致这些图片中路面的裂缝特征相对不明显, 如图 9(a)(c)所示。当裂缝图片经过 CAE 模型的特征学习与重构增强, 重新生成的图片中路面的裂缝相对更加清晰可见, 如图 9(b)(d)所示。

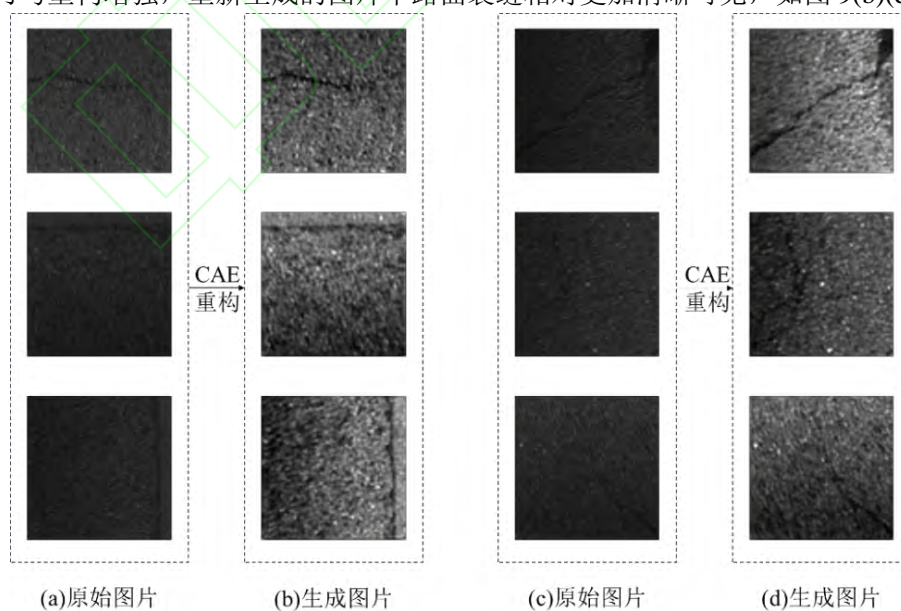


图 9 光照较差条件下部分沥青路面裂缝图片重构

Fig. 9 Generated Crack Images of Asphalt Pavement under Poor Illumination Conditions



## 4 有监督分类对比试验

### 4.1 对比实验设置

为验证 CAE 重构图片的数据扩增方法在有监督分类任务中的有效性，本文设计了如表 1 所示的对比实验。

表 1 对比实验中的数据集成

Table 1 Datasets for Comparison

对比实验	训练集	数据扩增方法	训练集样本量	测试集
1	Original	无扩增	1280	原始数据测试集
	Geometric	几何变换方法	10240	
	GeoCAE	先几何变换扩增，后 CAE 重构扩增	20480	
2	CAE2560	基于原始数据的 100 代 CAE 重构扩增	2560	
	Mirror	镜像变换		
	Rotation	图片旋转 90°、180°、270° 后随机选取其中每类别图片 640 张为训练集		
	Noise	增加高斯噪声		
	Histogram	图片自适应直方图均衡化方法		
3	CAE1280	与原始数据集一一对应的 CAE 重构图片	1280	CAE-test1
	Mirrorbased CAE1280	基于镜像变换扩增后数据集进行 CAE 特征学习，将模型迁移至原始数据集重构图片		CAE-test2
	Geobased CAE1280	将 GeoCAE 中训练完成的 CAE 模型，迁移至原始数据集重构图片		CAE-test3

(1) 对比实验 1 用以测试本文所使用的两步骤数据扩增手段的有效性。Original 中各样本图片不经过任何数据扩增手段；Geometric 为基于 Geometric 为利用图像处理方法扩增后的训练集；GeoCAE 为经过图像处理和 CAE 重构图片两个步骤扩增后的数据集。

(2) 对比实验 2 用以对比 CAE 数据增强方法与传统数据增强方法。传统数据增强方法选择了镜像、旋转两种基于空间几何变换的数据增强方法，和对图片增加噪声、图像自适应直方图均衡化两种基于像素颜色变换的数据增强方法。各实验均在原始数据集的基础上扩增，扩增后的训练集包含各数据增强手段增强后的图片与原始图片，并且保证扩增后各训练集的数据量相等，测试经各类方法数据增强后的模型在同一原始数据构成的测试集中的表现。

(3) 对比实验 3 用以测试经 CAE 重构后的图片是否更利于被深度学习模型识别，并测试 CAE 重构图片效果受特征学习时数据质量及数据量大小等因素的影响。CAE1280 组：基于原始数据集图片进行 CAE 特征学习与重构，该组实验的 1280 张训练集图片与测试集中的 320 张图片与原始数据集的训练集和测试集一一对应；MirrorbasedCAE1280 组：首先将原始数据集进行镜像变换扩增，CAE 特征学习时数据量扩充到原始数据集的 2 倍，利用学习完成的模型对原始数据集进行重构，同样生成与原始训练集与测试集一一对应的图片，包括训练集 1280 张与测试集 320 张图片；GeobasedCAE1280 组实验基于 Geometric 数据集进行 CAE 图片特征学习，即通过旋转和镜像的传统几何手段扩增，CAE 特征学习时数据量扩充到原始数据集的 8 倍，将学习训练完成的 CAE 模型基于原始数据集生成图片，同样得到与原始数据集的训练集和测试集一一对应的 CAE 重构图片，包括训练集 1280 张与测试集 320 张图片。

各组对比实验以 DenseNet 为统一的网络训练 100 代，训练时 Batch Size 取值为 16，即一次训练 16 个样本，以提高内存的利用率。使用 Adam<sup>[42]</sup>作为反向传播时的梯度下降的优化算法。Adam 算法中的超参数的配置如下： $\alpha=1\times10^{-5}$ ，表示学习率，对权重的更新比率进行控制，较小的值能够使得模型在收敛时获得更好的性能； $\beta_1=0.9$ ， $\beta_2=0.999$  分别控制一阶矩和二阶矩均值的衰减率； $\epsilon=1\times10^{-8}$ ，是为了防止在算法实现时除以 0 而设置的参数，其取值相对固定。

#### 4.2 对比试验结果分析

经过 100 代的训练，各组对比试验测试结果如表 2 所示。

表 2 采用不同数据集训练后模型的测试结果

Table 2 Test Results of DenseNet model after Training with Different Data Sets

对比实验	训练集	测试集	各类别测试准确率 (%)				
			总体	背景	裂缝	修补裂缝	路面标线
1	Original	原始数据测试集	78.43	61.25	76.25	90.00	86.25
	Geometric		83.44	72.50	72.50	97.50	91.25
	GeoCAE		87.19	88.75	75.00	97.50	87.50
2	CAE2560	原始数据测试集	82.50	72.50	81.25	87.50	88.75
	Mirror		80.63	77.50	72.50	80.00	92.50
	Rotation		80.00	70.00	70.00	93.75	86.25
	Noise		79.69	80.00	75.00	77.50	86.50
	Histogram		76.56	68.75	62.50	87.50	87.50
3	Orgbased CAE1280	CAE-test1	79.06	67.50	78.75	88.75	81.25
	Mirrorbased CAE1280	CAE-test2	89.38	95.00	88.75	86.25	87.50
	Geobased CAE1280	CAE-test3	83.13	90.00	71.25	80.00	91.25

通过对比实验 1 可知，不经过任何数据增强手段下，DenseNet 在原始数据集的总体分类精度为 78.43%，其中背景图片的分类准确率较低。融合旋转和镜像的基于几何手段的传统数据扩增方法 (83.44%)能够小幅提升路面图片总体分类准确率，精度得到主要提升的类别为路面背景图片和修补裂缝图片，但对于裂缝图片的分类准确率反而有所降低。而经过图像处理与 CAE 重构两步骤数据扩增的方法 (87.19%)，相较于无扩增手段，原始数据测试集的总体分类准确率提升，其中路面背景图片的分类正确率提升最为明显，并且缓解了几何变换扩增手段造成路面裂缝图片识别正确率降低的情况。

对比实验 2 对比了 CAE 数据增强方法与其他传统典型的数据增强算法，根据实验结果可知，当保持扩增后训练集样本量相等时，除图片直方图均衡化变换以外，其他各类传统数据增强方法都能一定程度上提升图片总体识别精度。其中，CAE 图像增强算法相较于不使用数据扩增手段，模型图片识别精度得到提高。对比其他传统数据增强算法，CAE 数据扩增方法有更好的表现。

对比实验 3 运用 3 种不同方法获得原始数据的 CAE 重构后的图片，并保持重构后实验训练集及测试集样本与原始数据集样本一一对应。根据实验结果可知，将 CAE 重构后的图片代替原始图片能够整体提升 DenseNet 对图片的识别精度，证明了 CAE 重构后的图片比原始图片更加有利于机器识别。经过几何扩增后的数据训练 CAE 模型，再就训练完成的模型对原始数据集进行重构，其重构后的数据质量与 CAE 训练时数据量大小有关。其原因可能在于，CAE 本身是一种深度学习方法，更多的训练数据更有利于 CAE 模型学习到图片特征。通过几何变换等传统手段预扩充数据集，再利用 CAE 对预扩充后的数据集进行特征学习的方法，有利于 CAE 模型性能的提升，从而提高 DenseNet 分类准确率。此外，注意到本文各

类别样本量较为均衡，因此未将精确度、召回率以及 F1 分数等作为评判指标，而只采用准确率来进行初步衡量，未来的深入研究可考虑更多的评判指标。

总体来说，相比于传统的数据扩增手段，本文所使用的 CAE 数据扩增手段有一定的优越性，CAE 重构后的图片能够一定程度提升机器对路面图片的识别精度。但 CAE 方法受特征学习时数据集样本量的影响，在未来应用中，当原始数据集较小时，采用融合传统数据扩增与 CAE 数据增强的方法，能够使得识别精度有更明显的提升。

## 5 DCEC 无监督聚类

### 5.1 网络结构与参数设置

为解决实际工程应用过程中大样本量数据集人工标注成本高的问题，本文继续探究利用深度聚类算法，实现大规模路面图片端到端无标签聚类的可能性。本文基于扩充后的样本量较大的数据集，对图片进行无监督聚类学习。

研究中所使用的无监督学习模型 DCEC 网络参考框架<sup>[24]</sup>如图 10 所示。DCEC 网络中所使用的 CAE 结构为：在编码器部分，依次进行卷积核尺寸为  $5 \times 5$ 、个数为 32 的卷积操作(卷积层 1)，卷积核尺寸为  $5 \times 5$ 、个数为 64 的卷积操作(卷积层 2)以及卷积核尺寸为  $3 \times 3$ 、个数为 128 的卷积操作(卷积层 3)。每层卷积编码层步长为 2，并以 ReLU 作为激活函数，每次输出特征图数量翻倍，每张特征图大小减半。将编码器后接维度为 80000 的展平层，再利用全连接操作，将所学得的特征在嵌入层降维至 10 维，后接解码器结构。解码器首先通过一个全连接层(FC)，将 10 维特征嵌入层全连接至与 Flatten 层相对称的 80000 个神经元中。接着经过重塑层，将 80000 维向量重塑为  $25 \times 25 \times 128$  的特征图。其后依次进行卷积核尺寸为  $3 \times 3$ 、个数为 32 的反卷积操作(反卷积层 1)，卷积核尺寸为  $5 \times 5$ 、个数为 32 的反卷积操作(反卷积层 2)，以及卷积核尺寸为  $5 \times 5$ 、个数为 1 的反卷积操作(反卷积层 3)，每次反卷积操作的步长为 2，除了输出层使用 sigmoid 激活函数以外，其余层仍然以 ReLU 作为激活函数。经过各层编码层的反卷积操作，图片尺寸不断增加，直至图片重构为与原始路面图片同样大小的灰度图片<sup>[24,33]</sup>。

DCEC 中的聚类层结构即将 10 维特征的嵌入层后接聚类层 (Clustering layer)，使得图片特征参数大小大大降低<sup>[24,33]</sup>。

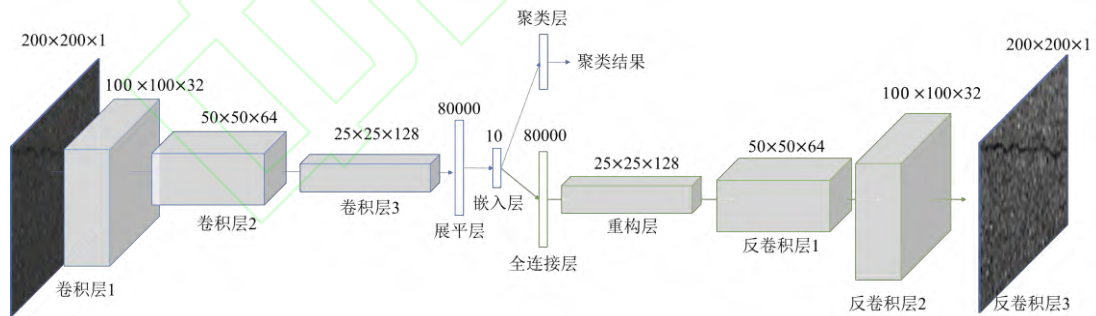


图 10 本文中 DCEC 网络框架图<sup>[24]</sup>

Fig. 10 DCEC Network Structure Used in This Study<sup>[24]</sup>

DCEC 分两阶段训练，第一阶段由模型中 CAE 部分预训练完成特征提取，使得预训练所得的 10 维特征值能够较好的拟合输入数据集的特征；第二阶段由聚类模型完成，在 DCEC 算法中嵌入 k-means 聚类层，利用损失函数继续迭代，微调网络，间隔一定迭代次数更新聚类结果，比较前后两次更新后标签的分配情况，若更改变动小于一定阈值，则停止训练，最终完成无监督聚类过程。本文使用的 DECE 模型中的各个超参数的设置参考了 DEC 网络及后续改进算法<sup>[22-24]</sup>，各超参数取值如表 3 所示：

表 3 DCEC 模型各超参数设置<sup>[22-24]</sup>  
Table 3 Hyper-parameters in DCEC<sup>[22-24]</sup>

阶段	超参数	取值	含义
预训练	优化算法	Adam <sup>[42]</sup>	$\alpha=0.001, \beta_1=0.9, \beta_2=0.999, \varepsilon=1 \times 10^{-8}$
	迭代次数	200	预训练阶段迭代次数为 200 次
	Batch size <sup>[28]</sup>	256	预训练阶段每次训练 256 张图片
聚类	优化算法	Adam <sup>[42]</sup>	$\alpha=0.001, \beta_1=0.9, \beta_2=0.999, \varepsilon=1 \times 10^{-8}$
	$\gamma$	0.1	决定式(6)中两部分 Loss 的权重
	Batch size <sup>[28]</sup>	256	每次训练 256 张图片
	更新间隔	140	间隔 140 代更新聚类结果
	tol	0.001	当两次连续更新之间标签分配更改小于阈值时停止训练
	maxiter	$2 \times 10^4$	允许迭代的最大次数

## 5.2 训练结果对比分析

### 5.2.1 算法对比

在无监督学习部分, 选取 Geometric 训练集中路面背景、裂缝及路面标线三类样本图片各 2000 张, 以及其对应的每类 2000 张 CAE 重构后的图片, 共 12000 张图片作为数据集进行测试。在数据输入无监督训练模型时, 删去所有数据的标签类别信息, 仅在计算准确率时使用。为验证 DCEC 算法的有效性, 将聚类学习结果与以下两种方法进行对比实验:

(1) k-means<sup>[43]</sup>: 利用 python 语言中的 sklearn 机器学习算法工具中的 k-means 算法包, 根据集群图片之间的的距离度量, 当质心移动小于公差时, 迭代停止, 完成聚类。

(2) AE+k-means<sup>[44]</sup>: 利用全连接的 Auto-encoder (AE) 结构对图片进行特征学习和降维, 在 AE 结构充分学习后, 将 AE 编码器部分的输出后接 k-means 聚类层, 实现聚类。

表 4 不同无监督机器学习模型的测试结果

Table 4 Test Results of Different Unsupervised Machine Learning Model

算法	ACC	NMI	ARI
DCEC	0.5798	0.2211	0.1637
k-means	0.4700	0.0708	0.0394
AE+k-means	0.4785	0.0772	0.0442

各方法训练结果如表 4 所示。在大样本量的沥青路面图片无监督聚类中, 相较于普通机器学习无监督算法, DCEC 算法将聚类准确度提升了约 10%, 同时 NMI 以及 ARI 两项指标得分远高于普通机器学习算法。表明, 相较于传统算法, 利用基于 DCEC 的路面目标与裂缝无监督识别方法所得聚类结果与原始数据匹配程度更高, 聚类效果更好。

### 5.2.2 数据集对比

本文进一步探究是否使用 CAE 算法扩充数据集对无监督聚类精度提升的影响, 建立如表 5 所示的以三组数据集构成的对比实验。

表 5 无监督学习数据集组成

Table 5 Data Set Composition in Unsupervised Learning

数据集 编号	说明	各类别图片数量			总数量
		路面背景	裂缝	路面标线	
数据集 A	由未经 CAE 重构的图片组成	2000	2000	2000	6000
数据集 B	由经 CAE 重构后的图片组成	2000	2000	2000	6000
数据集 C	为数据集 A 和 B 的混合	4000	4000	4000	12000

在预训练阶段, 各数据集中经过 200 代的预训练, Loss 较快达到收敛, 表明 DCEC 中的 CAE 网络已完成数据集的特征学习。



在聚类阶段继续训练，并调整网络。当两次连续更新之间的标签分配更改率小于设定的阈值 0.001 时停止训练。训练过程中，随迭代次数增加，不同训练数据集下 DCEC 网络各指标变化曲线如图 11。

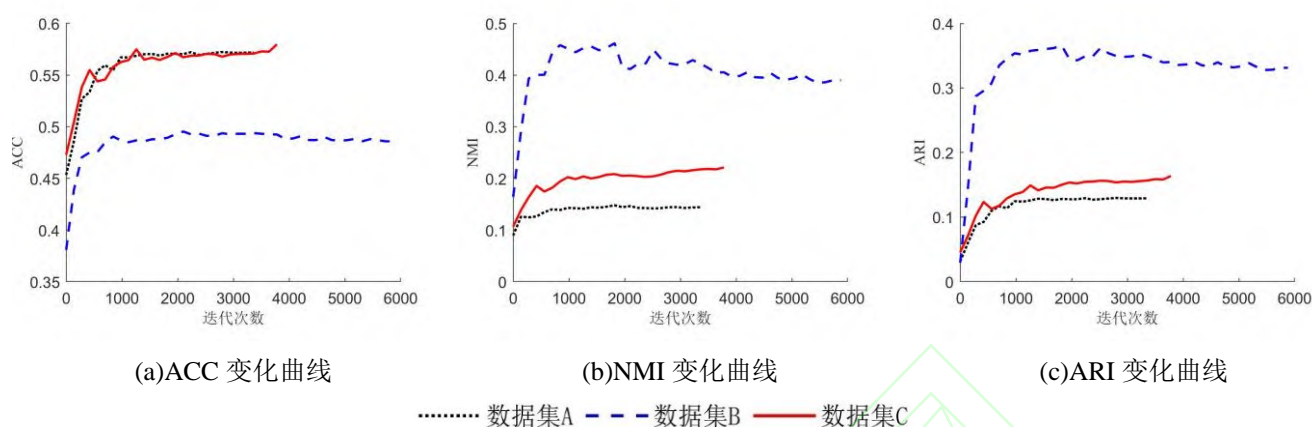


图 11 DCEC 训练过程各指标变化曲线

Fig. 11 Changing Curve of DCEC Training Process Parameter

结束训练后，各数据集经 DCEC 网络无监督训练结果见表 6。由结果可知，融合了未经重构及重构后图片的数据集 C 的聚类准确率与仅包含未经重构的数据集 A 所得的准确率相近，但数据集 C 的 NMI 值和 ARI 值较高。而仅由重构后图片组成的数据集 B 所得准确率较低，但所得 NMI 值和 ARI 值高于数据集 A 与数据集 C。由此可见，CAE 数据增强方法能够一定程度使得聚类结果与数据真实分布情况更加匹配，但总的来说对路面图片的无监督聚类识别整体提升效果较为有限。其可能的原因在于，路面背景图片噪声较大，造成数据集类内差距较小，最终导致整体无监督识别精度较低。

表 6 各数据集训练结果

Table 6 Training Results of Different Data Sets

数据集	迭代次数	ACC	NMI	ARI
数据集 A	3360	0.5718	0.1441	0.1291
数据集 B	5580	0.4858	0.3901	0.3312
数据集 C	3780	0.5798	0.2211	0.1637

当前，深度聚类算法仍处于发展阶段，大规模数据集下的无监督聚类算法尚未发展成熟。目前的深度聚类算法大多停留在理论研究阶段，距离真正的工程应用还需进一步研究完善。在计算机视觉领域，研究者们大多将其用于对黑底白字，图片尺寸较小(28×28 像素)，图片特征较为明显的 MNIST<sup>[45]</sup>或图片尺寸更小(16×16 像素)的 USPS<sup>[23]</sup>两种手写数字数据集进行试验，聚类准确率较有监督学习分类准确率低。也有学者将深度聚类在稍复杂的图片尺寸为 96×96 像素 STL-10<sup>[46]</sup>数据集上进行实验，其所得准确率则更低。例如 DEC 网络提出者 Xie<sup>[22]</sup>利用 DEC 对各数据集进行实验，尽管在 MNIST 数据上取得了 84.3% 的实验精度，但当使用 DEC 对经 8×8Hog 特征提取后的 STL-10 数据集进行测试，准确率仅为 35.9%。

综上所述，将深度聚类方法真正应用于图片特征不够明显的道路裂缝等病害图片的识别上还需要进一步研究，但本文所使用的 DCEC 网络是一种较有前景的无监督分类方法。

## 6 结语

本文利用传统图片几何变换和 CAE 网络重构增强图片两种方法对小样本路面图片数据集进行扩充，利用卷积神经网络对原始数据集及两组增强数据集进行有监督分类对比研究，并基于深度聚类模型 DCEC 探究了无监督学习在路面裂缝识别方面的应用，得到以下结论：

(1) 针对路面裂缝图片的小样本数据集,运用常规图片几何变换和 CAE 网络重构增强等数据扩增方法能够扩充样本数量,较好地提升了神经网络模型对路面图片分类的精度。

(2) 本文所使用的 CAE 重构路面图片更具图片特征,在沥青黑色路面背景下较为清晰,在有监督学习中,该数据扩增方法能够一定程度提升模型的可靠性。对比传统的数据增强方法,该方法可以较好地提升识别精度。并且,将传统方法与 CAE 方法相结合的方式,能够提升深度学习模型对图片的识别准确率。

(3) 深度无监督聚类在人工智能领域仍处于发展状态,本文所使用的基于 CAE 图像特征预训练的 DCEC 深度聚类网络方法对路面图片的识别精度相较于传统无监督聚类算法提升了约 10%,且聚类结果与真实图片分布匹配程度更佳,因此具有潜在的应用前景,但将该算法进行真正的工程应用还有待深入研究。

(4) 本文在设置 CAE 算法与其他各类数据增强算法的对比实验中,没有讨论生成式对抗网络方法 GAN。在后续的工作中将对本文算法与 GAN 等算法进行对比,并进一步探究本文所使用的算法在不同场景下路面病害识别的适用性。

参数解释表

序号	参数	含义
1	$k$	卷积核个数
2	$l$	卷积核长宽
3	$s$	卷积操作步长
4	$w$	卷积层输入图像宽度
5	$h$	卷积层输入图像长度
6	$x$	激活函数输入信号值
7	$N$	全连接特征层展平向量元素个数
8	$T$	多分类类别数量
9	$j$	多分类任务中的类别编号
10	$x_j$	$j$ 类别的输入信号
11	$ACC$	准确率
12	$n_c$	预测正确的样本数量
13	$n_t$	总的样本数量
14	$L_r$	DCEC 网络损失函数中的重建损失
15	$L_c$	DCEC 网络损失函数中的聚类损失
16	$\gamma$	嵌入空间扭曲程度控制系数
17	$n$	数据集中的图片数量
18	$x_i$	第 $i$ 张输入模型的原始图片
19	$F_\theta$	编码函数
20	$G_{\theta'}$	解码函数
21	$Z_i$	第 $i$ 张图片图片经过编码器结构得到的潜在特征空间
22	$x'_i$	潜在空间特征经过解码器结构得到的重建数据
23	$P$	目标分布
24	$Q$	拟合分布
25	$q_{ij}$	第 $i$ 个样本 $q_i$ 属于第 $j$ 类的概率
26	$\mu_j$	第 $j$ 类样本的聚类中心
27	$p_{ij}$	目标分布

28	NMI	标准互信息
29	ARI	调整 Rand 指数
30	RI	兰德系数
31	X	聚类结果划分 X
32	Y	实际划分 Y
33	I(X,Y)	聚类结果划分 S 与实际划分 P 的互信息量
34	H(X)	聚类结果划分的熵
35	H(Y)	实际划分的熵
36	$a$	聚类结果与实际情况中都为同一类的元素对数
37	$b$	聚类结果和实际情况都为不同类的元素个数
38	$step_x$	横向裁剪图片步长
39	$step_y$	纵向裁剪图片步长
40	Adam	反向传播时的梯度下降的优化算法
41	$\alpha$	学习率, 优化算法中由学习率对权重的更新比率进行控制, 较小的值能够使得模型在收敛时获得更好的性能
42	$\beta_1$	Adam 算法中控制一阶矩均值的衰减率的参数
43	$\beta_2$	Adam 算法中控制二阶矩均值的衰减率的参数
44	$\varepsilon$	Adam 算法中为了防止在算法实现时除以 0 而设置的参数
45	Adadelata	自适应学习率调整算法
46	Batch size	每次训练图片张数
47	$tol$	当两次连续更新之间的标签分配更改小于阈值时停止训练

## 参考文献:

## References:

- [1] FUJITA Y, SHIMADA K, ICHIHARA M, et al. A Method Based on Machine Learning using Hand-crafted Features for Crack Detection from Asphalt Pavement Surface Images [C] // SPIE. Thirteenth International Conference on Quality Control by Artificial Vision 2017. Bellingham: SPIE, 2017: 103381.
- [2] PAN Y, ZHANG X, CERVONE G, et al. Detection of Asphalt Pavement Potholes and Cracks Based on the Unmanned Aerial Vehicle Multispectral Imagery [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11 (10): 3701-3712.
- [3] ZHANG A, WANG K C P, LI B, et al. Automated Pixel-Level Pavement Crack Detection on 3D Asphalt Surfaces Using a Deep-Learning Network [J]. Computer-aided Civil and Infrastructure Engineering, 2017, 32: 805-819.
- [4] 沙爱民, 童 峥, 高 杰. 基于卷积神经网络的路表病害识别与测量[J]. 中国公路学报, 2018, 31(1):1-10.  
SHA Ai-min, TONG Zheng, GAO Jie. Recognition and Measurement of Pavement Disasters Based on Convolutional Neural Networks [J]. China Journal of Highway and Transport, 2018, 31 (1):1-10.
- [5] DORAFSHAN S, THOMAS R J, MAGUIRE M. Comparison of Deep Convolutional Neural Networks and Edge Detectors for Image-based Crack detection in Concrete [J]. Construction and Building Materials, 2018, 186: 1031-1045.

- [6] SHORTEN C, KHOSHGOFTAAR T M. A survey on Image Data Augmentation for Deep Learning [J]. Journal of Big Data, 2019, 6(1): 1-48.
- [7] 邱 果. 基于高分辨率图像的纺织物缺陷检测算法研究[D]. 哈尔滨:哈尔滨工业大学,2019.
- QIU Guo. Research on Textile Defect Detection Algorithm Based on High Resolution Image [D]. Harbin: Harbin Institute of Technology.
- [8] CHAWLA N V, BOWYER K W, HALL L O, et al. SMOTE: Synthetic Minority Over-sampling Technique [J]. Journal of Artificial Intelligence Research, 2002, 16: 321-357.
- [9] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative Adversarial Nets [C] // Curran Associates. Proceedings of Advances in Neural Information Processing Systems 27 (NIPS 2014). San Francisco: Curran Associates, 2014: 2672-2680.
- [10] Zhang K, Zhang Y, Cheng H, et al. CrackGAN: A Labor-Light Crack Detection Approach Using Industrial Pavement Images Based on Generative Adversarial Learning [J]. arXiv: Computer Vision and Pattern Recognition. arXiv:1909.08216 [cs.CV]. 2019.
- [11] 李良福,孙瑞赟. 复杂背景下基于图像处理的桥梁裂缝检测算法[J]. 激光与光电子学进展,2019,56 (6): 112-122.
- LI Liang-fu, SUN Rui-yun. Bridge Crack Detection Algorithm Based on Image Processing under Complex Background [J]. Laser and Optoelectronics Progress. 2019, 56 (6): 112-122.
- [12] 李良福,胡 敏. 基于生成式对抗网络的细小桥梁裂缝分割方法 [J]. 激光与光电子学进展, 2019, 56(10): 102-112.
- LI Liang-fu, HU Min. Method for Small-Bridge-Crack Segmentation Based on Generative Adversarial Network [J]. Laser and Optoelectronics Progress. 2019, 56(10): 102-112.
- [13] WANG J, PEREZ L. The Effectiveness of Data Augmentation in Image Classification Using Deep Learning [J]. arXiv: Computer Vision and Pattern Recognition. arXiv:1712.04621 [cs.CV], 2017.
- [14] TANAKA F H K S, ARANHA C. Data Augmentation Using GANs [J]. arXiv: Learning, 2019.
- [15] LORE, KIN G, AKINTAYO, et al. LLNet: A Deep Autoencoder Approach to Natural Low-light Image Enhancement [J]. Pattern Recognition, 2015, 61:650-662.
- [16] 王万良,杨小涵,赵燕伟,等. 采用卷积自编码器网络的图像增强算法[J]. 浙江大学学报(工学版), 2019, 53(9):1728-1740.
- WANG Wan-liang, YANG Xiao-han, ZHAO Yan-wei, et al. Image Enhancement Algorithm with Covolutional Auto-encoder Network [J]. Journal of Zhejiang University (Engineering Science), 2019, 53 (9): 1728-1740.
- [17] SHI L, GAO C, ZHANG J. Pavement Distress Image Recognition Based on Multilayer Autoencoders [C] // Springer. Artificial Intelligence and Computational Intelligence. Berlin: Springer, 2012: 666-673.
- [18] 钱 彬,唐振民,徐 威. 基于稀疏自编码的路面裂缝检测[J]. 北京理工大学学报,2015,35(8):800-804,809.
- QIAN Bin, TANG Zhen-min, XU Wei. Pavement Crack Detection Based on Sparse Autoencoder [J]. Transactions of Beijing Institute of Technology, 2015, 35(8): 800-804,809.
- [19] 董安国,张仙艳,薛宏智,等. 混凝土表面裂缝检测的多级聚类算法[J]. 交通运输工程学报, 2013,13(6):7-13.



DONG An-guo, ZHANG Xian-yan, XUE Hong-zhi, et al. Multi-level Clustering Algorithm for Crack Detection of Concrete Surface [J]. Journal of Traffic and Transportation Engineering, 2013, 13(6):7-13.

- [20] 王德方,曾卫明,王倪传. 基于改进 K-means 算法的不均匀光照下道路裂缝检测[J]. 计算机应用与软件,2015,32(7):244-247,300.

WANG De-fang, ZENG Wei-ming, WANG Ni-zhuan. Road Crack Detection under Uneven Illumination using Improved K-means Algorithm [J]. Computer Applications and Software. 2015, 32 (7): 244-247,300.

- [21] 徐港,赵恬悦,蒋赏,等. 基于多种连通域特征的结构表面裂缝提取方法[J]. 华中科技大学学报(自然科学版), 2019, 47(10): 52-55, 68.

XU Gang, ZHAO Tian-yue, JIANG Shang, et al. Extraction Method of Structural Surface Cracks Based on Multiple Connected Domain Features [J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2019, 47 (10): 52-55, 68.

- [22] XIE J, GIRSHICK R, FARHADI A. Unsupervised Deep Embedding for Clustering Analysis [C] // Curran Associates. Proceedings of the 33rd International Conference on Machine Learning. New York: Curran Associates, 2016: 478-487.
- [23] GUO X, GAO L, LIU X, et al. Improved Deep Embedded Clustering with Local Structure Preservation [C] // AAAI Press: Proceedings of the 26th International Joint Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2017: 1753-1759.
- [24] GUO X, LIU X, ZHU E, et al. Deep Clustering with Convolutional Autoencoders [C] // Springer. Neural Information Processing: 24th International Conference, ICONIP 2017, Proceedings, Part II. Berlin: Springer, 2017: 373-382.
- [25] CARON M, BOJANOWSKI P, JOULIN A, et al. Deep Clustering for Unsupervised Learning of Visual Features [C] // Springer. Proceedings of the European Conference on Computer Vision (ECCV). Berlin: Springer, 2018: 132-149.
- [26] NAIR V, HINTON G E. Rectified Linear Units Improve Restricted Boltzmann Machines [C] // Omnipress. Proceedings of the 27th International Conference on Machine Learning (ICML-10). Omnipress, 2010: 807-814.
- [27] LIN M, CHEN Q, YAN S. Network in Network [J]. arXiv: Neural and Evolutionary Computing, 2013.
- [28] IOFFE S, SZEGEDY C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift [J]. arXiv: Learning, 2015.
- [29] HUANG G, LIU Z, VAN D M L, et al. Densely Connected Convolutional Networks [C] // IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2017: 4700-4708.
- [30] HINTON G E, SALAKHUTDINOV R R. Reducing the Dimensionality of Data with Neural Networks [J]. Science, 2006, 313 (5786): 504-507.
- [31] MASCI J, MEIER U, CIRESAN D, et al. Stacked Convolutional Auto-encoders for Hierarchical Feature Extraction [C] // Springer. Artificial Neural Networks and Machine Learning – ICANN 2011. Berlin: Springer, 2011: 52-59.
- [32] 袁非牛,章琳,史劲亭,等. 自编码神经网络理论及应用综述[J]. 计算机学报,2019,42(1):203-230.

YUAN Fei-niu, ZHANG Lin, SHI Jin-ting, et al. Theories and Applications of Auto-encoder Neural Networks: A Literature Survey [J]. Chinese Journal of Computers, 2019, 42(01): 203-230.

- [33] 谢娟英,侯 琦,曹嘉文. 深度卷积自编码图像聚类算法[J]. 计算机科学与探索,2019,13(4):586-595.
- XIE Juan-ying, Hou Qi, CAO Jia-wen. Image Clustering Algorithms by Deep Convolutional Autoencoders [J]. Journal of Frontiers of Computer Science and Technology, 2019, 13(4): 586-595.
- [34] VINH N X, EPPS J, BAILEY J. Information Theoretic Measures for Clusterings Comparison: Is a Correction for Chance Necessary? [C] // Association for Computing Machinery. Proceedings of the 26th Annual International Conference on Machine Learning. New York: Association for Computing Machinery, 2009: 1073-1080.
- [35] VINH N X, EPPS J, BAILEY J. Information Theoretic Measures for Clusterings Comparison: Variants, Properties, Normalization and Correction for Chance [J]. Journal of Machine Learning Research, 2010, 11: 2837-2854.
- [36] 陈梅. 面向复杂数据的聚类算法研究[D]. 兰州: 兰州大学,2016.
- CHEN Mei. Research of Clustering Algorithm for Complex Data [D]. Lan-zhou: Lanzhou University, 2016.
- [37] LNF Ana, Anil K Jain. Robust Data Clustering [C] // IEEE. Computer Vision and Pattern Recognition, 2003. Proceedings. New York: IEEE, 2003: 128-136.
- [38] RAND W M. Objective Criteria for the Evaluation of Clustering Methods [J]. Journal of the American Statistical Association, 1971, 66(336): 846-850.
- [39] HUBERT L, ARABIE P. Comparing Partitions [J]. Journal of Classification, 1985, 2(1):193-218.
- [40] 薛亚东,李宜城. 基于深度学习的盾构隧道衬砌病害识别方法[J]. 湖南大学学报, 2018, 45(3): 100-109.
- XUE Ya-dong, LI Yi-cheng. A Method of Disease Recognition for Shield Tunnel Lining Based on Deep Learning [J]. Journal of Hunan University (Natural Science), 2018, 45(3): 100-109.
- [41] ZEILER M D. ADADELTA: An Adaptive Learning Rate Method [J]. arXiv: Learning, 2012.
- [42] KINGMA D P, BA J. Adam: A Method for Stochastic Optimization [J]. arXiv: Learning, 2014.
- [43] WAGSTAFF K, CARDIE C, ROGERS S, et al. Constrained k-means Clustering with Background Knowledge [C] // BRODLEY C E, DANYLUK A P. Proceedings of the 18th International Conference on Machine Learning. San Francisco: Morgan Kaufmann Publishers Inc., 2001: 577-584.
- [44] SONG C, LIU F, HUANG Y, et al. Auto-encoder Based Data Clustering [C] // Springer. Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. Berlin: Springer, 2013: 117-124.
- [45] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based Learning Applied to Document Recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [46] COATES A, NG A, LEE H. An Analysis of Single-layer Networks in Unsupervised Feature Learning [C] // JMLR. Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics. Fort Lauderdale: JMLR, 2011: 215-223.