

文章编号: 2095-2163(2021)09-0123-05

中图分类号: U216.3;TP183;TP391.41

文献标志码: A

基于自编码器的零样本扣件检测

申瑞超,柴晓冬,李立明,孙睿,钟倩文

(上海工程技术大学 城市轨道交通学院,上海 201620)

摘要: 基于深度学习的扣件检测需要大量人工标注的扣件图像数据集驱动,然而铁路扣件图像中负样本偏少,不均衡的数据集会使深度学习模型的泛化能力较差,达不到检测扣件状态的效果。针对该问题,本文提出了一种基于自编码器的零样本扣件检测。首先,使用欠完备自编码器、栈式自编码器和卷积自编码器提取扣件正样本图像特征;然后,通过正样本特征向量与基向量的余弦相似度推断出负样本的分布空间;在检测时将各自编码器算法得出的结果利用多数投票法确定样本属性。实验证明,使用本文方法,在只使用正样本训练的情况下,可以有效地检测出扣件图像的负样本,准确率为 95.59%,实现了零样本扣件检测。

关键词: 零样本扣件检测;自编码器;余弦相似度;多数投票法

Zero-shot fastener detection based on Autoencoder

SHEN Ruichao, CHAI Xiaodong, LI Liming, SUN Rui, ZHONG Qianwen

(College of Urban Rail Transit, Shanghai University of Engineering Science, Shanghai 201620, China)

【Abstract】 Fastener detection based on deep learning needs a large number of manually labeled fastener image data sets. However, there are few negative samples in railway fastener images, and the unbalanced data sets make the generalization ability of deep learning model poor, which can not achieve the effect of fastener state detection. To solve this problem, this paper proposes a zero-shot fastener detection method based on Autoencoder. Firstly, the image features of fastener positive samples are extracted by using Under complete Autoencoder, stack Autoencoder and convolution Autoencoder. Then the distribution space of negative samples is deduced by the cosine similarity between the feature vector of positive samples and the basis vector. In the detection process, the results obtained by the respective encoder algorithms are used to determine the sample attributes by voting algorithm. Experimental results show that using this method, in the case of only using positive samples training. The accuracy is 95.59% and realize the zero sample fastener detection, can effectively detect the negative samples of fastener image, can effectively detect the negative samples of fastener image, the accuracy is 95.59%, and realize the zero sample fastener detection.

【Key words】 zero-shot fastener detection; Autoencoder; cosine similarity; voting algorithm

0 引言

铁路运输是一种能源集约式的陆地交通运输,因其运输量大、速度较快、受环境因素影响较小、安全性强等特点受到国家重点发展,截至 2020 年 7 月底,中国铁路运输总营业里程已超过 14 万千米^[1]。在铁路运输中安全尤为重要,扣件作为钢轨与枕木的连接部件,扣件状态的检测在铁路运输安全中扮演着重要角色。扣件处于缺失、位移、断裂等失效状态时,可能会造成列车脱轨等重大安全事故。人工巡检是常用的扣件状态检测手段之一,但绵长的铁路线路和复杂的环境,耗费了巨大的人力和财力,工作人员的安全也得不到保障,利用机器视觉技术检

测铁路扣件状态已成为趋势。如 20 世纪 90 年代,美国的 VIS 综合系统使用人工和相结合的方式对扣件、钢轨和道床等进行检测^[2];2016 年,中国自主研发的综合铁路巡检车在昆明铁路局试用,实现了自动化检测钢轨表面和扣件,其对扣件状态检测的准确率达 96%^[3]。为了提高准确率,机器学习和深度学习应用到扣件检测中,如 Khan R A^[4]等结合脚点检测和模板匹配检测扣件;王强等^[5]利用改进的 LBP 特征对扣件进行检测,提高了在不同环境下的检测的正确率;林菲等^[6]采用深度学习网络 VGG 自动提取扣件特征,使得扣件检测的精度提升,达 97.14%。然而,机器视觉算法的有效性是建立在数据分布大致均衡的假设基础上^[7],采集到的现场扣

基金项目: 国家自然科学基金(51975347);上海市科委重点支撑项目(18030501300)。

作者简介: 申瑞超(1994-),男,硕士研究生,主要研究方向:铁路扣件状态检测;柴晓冬(1962-),男,博士,教授,主要研究方向:通信与信号工程及故障诊断。

通讯作者: 柴晓冬 Email: cxdyj@163.com

收稿日期: 2021-03-18

件图像正负样本往往不均衡^[8],这样的数据训练出的模型就会造成将负样本检测为正样本的情况。如Gibert^[9]和李永波^[10]在利用机器视觉算法检测扣件时,正负样本失衡,虽然在其各自提出的算法中做了验证,但是负样本较小的情况下的算法指标不能代表模型的泛化能力^[11],在现实应用中可能会造成漏报率较高的情况。为解决机器学习中样本不均衡问题,Nekooeimehr^[12]等提出了一种加权训练方式,通过对少数样本的采样次数和权重进行自适应,实现不均衡样本的机器学习^[12]。本文提出了一种基于

自编码器的零样本扣件检测,只需要提供正样本,就可以达到检测扣件状态的目的。

算法流程如图1所示。首先,使用欠完备自编码器、栈式自编码器和卷积自编码器提取扣件正样本图像特征;通过正样本特征向量与基向量的余弦相似度推断出负样本的分布空间;在检测时将各自编码器算法得出的结果利用多数投票法确定样本属性。实验证明,使用本文方法,在只使用正样本训练的情况下,可以有效地检测出扣件图像的负样本,实现了零样本扣件检测。

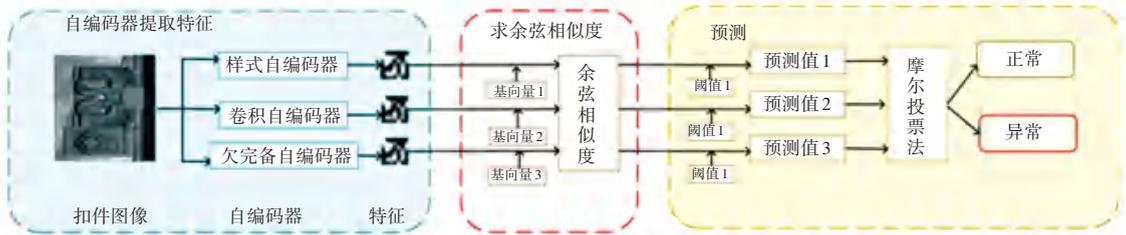


图1 算法流程图

Fig. 1 Algorithm flow chart

1 自编码器

自编码器是前馈非循环神经网络,不需要借助事先标注的标签,通过无监督学习,学习到数据的高效表示^[13],具有非常好的提取数据特征的能力,在图像重构、聚类、机器翻译等方面有着广泛应用。自编码器的通用结构如图2所示,有输入层、隐含层和输出层,可以分为编码器部分和解码器部分^[14]。假设自编码器的输入层输入为 v ,隐含层的输出为 y ,输出层的输出为 s ,则 v 和 s 有相同的维度,在训练时,通过调整自编码器内部的参数,尽可能的让 v 和 s 的值相同,训练完成后,移除解码器,新的数据经过编码器前馈运算输出的值就是编码结果。

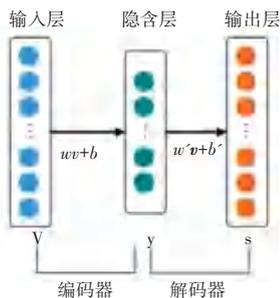


图2 自编码器通用结构

Fig. 2 General structure diagram of Autoencoder

编码器通过式(1)将 v 映射为 y ,实现了原始数据的编码,解码器通过式(2)将编码数据 y 重构为 s ,实现数据解码。

$$y = f(v) = \xi_1(w_1v + b_1) \quad (1)$$

$$s = g(v) = \xi_2(w_2y + b_2) \quad (2)$$

其中, w_1 和 b_1 是编码器的参数矩阵和偏置量, w_2 和 b_2 为解码器的参数矩阵和偏置量,通过对 w_1 、 b_1 、 w_2 和 b_2 的调整,使 v 和 s 相同; ξ_1 和 ξ_2 为映射函数。本文使用3种自编码器:欠完备自编码器、栈式自编码器和稀疏自编码器。

1.1 欠完备自编码器

欠完备自编码器的隐含层神经元数少于输入层,即 s 的维度小于 v 的维度。初始状态时, w_1 、 b_1 、 w_2 和 b_2 都是随机初始化,通过最小化损失函数 L 实现 s 与 v 的值相同。假设 v 和 s 的维度为 n ,本文欠完备自编码器使用的均方误差(Mean Square Error, MSE)损失函数的表达式为:

$$L(v, s) = \frac{1}{n} \sum_i^n (s_i - v_i)^2 \quad (3)$$

其中, s_i 表示输入图像的第 i 个像素值, v_i 表示解码器的第 i 个输出值。自编码器的参数用 w 表示,参数的更新如式(4):

$$w = w - \eta \frac{\partial L}{\partial w} \quad (4)$$

其中, η 是学习率,通过链式求导法则更新自编码器中的每个参数使 $L(v, s)$ 趋近于0,此时 y 就是 v 的低维度表示。

本文使用的欠完备自编码器如图3所示,隐含层使用49个神经元,输出层与图像维度相同,使用

式(5) *ReLU* 作为映射函数。

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (5)$$

在训练时将扣件图像的每个像素依次排列, 拉伸为 211 600 维的向量作为自编码器的输入, 将输出值带入式(3)求得损失, 并通过式(4)更新自编码器参数。

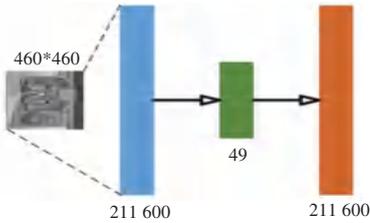


图 3 欠完备自编码器

Fig. 3 Undercomplete Autoencoder

1.2 栈式自编码器

栈式自编码是一个由多层自编码器堆叠组成的神经网络, 其前一层自编码器的输出作为其后一层自编码器的输入。对于一个 k 层栈式自编码的编码过程就是按照从前向后的顺序执行每一层自编码器的编码步骤。假设 H_i 表示栈式自编码器的第 i 层 ($0 \leq i \leq k$), 自编码器的隐含层 H_i 会作为 H_{i+1} 层的输入层, 第一个输入层就是整个神经网络的输入层, 利用贪心算法训练每一层的步骤如下:

(1) 首先采用自编码网络, 训练从输入层到 H_1 层的参数, 训练完毕后, 去除解码层, 只留下从输入层到隐藏层的编码阶段;

(2) 接着训练从 H_1 到 H_2 的参数, 把无标签数据的 H_1 层神经元的输出值作为 H_2 层的输入, 然后再进行自编码训练, 训练完毕后, 再去除 H_2 层的解码层;

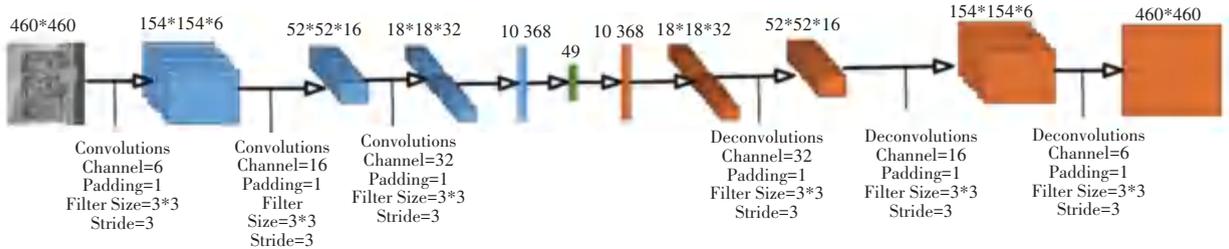


图 5 卷积自编码器

Fig. 5 Convolution Autoencoder

2 余弦相似度和多数投票法

扣件图像经自编码器提取的特征都为 49 维的高维特征, 本文使用余弦相似度作为度量这些高维特征之间相似度的尺度。余弦相似度是用空间中两

(3) 对所有层重复(1)和(2), 即移除前面自编码器的输出层, 用另一个自编码器代替, 再用反向传播进行训练, 直到栈式自编码器的层数达到 k 。

本文所用栈式自编码器结构如图 4 所示。编码器部分神经元采用递减式结构, 减少了重要信息的丢失; 输出层的维度为 49; 在解码器部分, 各层神经元数与解码器对称, 使用式(5)作为映射函数。

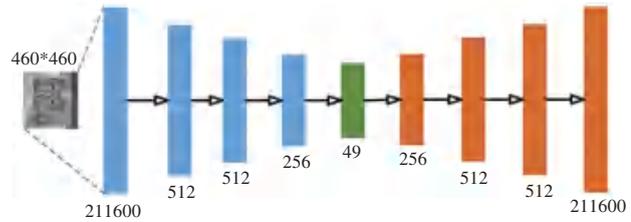


图 4 栈式自编码器

Fig. 4 Stack Autoencoder

1.3 卷积自编码器

含有卷积操作层 (Convolutional layers) 的自编码器称为卷积自编码器。卷积层含有一系列滤波器, 在特征图上按照一定的步长滑动, 将卷积核参数与像素值对应相乘后相加输出。卷积层作用于图像的局部区域进而获得图像的局部特征, 多层卷积层的应用, 可以扩大这个局部区域的范围, 使提取到的特征更具备空间特性, 因此卷积操作更适合用于图像领域^[15]。

本文使用的卷积自编码器结构如图 5 所示。在编码器部分, 卷积层数量为 3, 每层卷积核数量递增, 卷积核的大小为 3×3 , 步长为 3; 在输出层与卷积层之间有一个全连接层作为过渡; 输出层维度为 49; 在解码器部分有与编码器部分对称的全连接层和反卷积层, 使用式(5)作为映射函数。

个向量的夹角来判断这两个向量的相似程度, 两个向量夹角越大, 距离越远, 最大距离就是两个向量夹角 180° ; 夹角越小, 距离越近, 最小距离就是两个向量夹角 0° , 完全重合。假设有 \mathbf{A} 、 \mathbf{B} 两个 n 维向量, 余弦相似度用式(6)求解。

$$\cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n \mathbf{A}_i \times \mathbf{B}_i}{\sqrt{\sum_{i=1}^n (\mathbf{A}_i)^2} \times \sqrt{\sum_{i=1}^n (\mathbf{B}_i)^2}} \quad (6)$$

其中, \mathbf{A}_i 和 \mathbf{B}_i 代表向量的第 i 维的值。使用 3 种自编码器对扣件图像特征进行编码,并通过余弦相似度确定扣件状态,当 3 种方法识别的扣件状态不一致时,使用多数投票法确定结果。假设扣件的正常状态用 1 表示,异常状态用 0 表示,其真值见表 1。

表 1 投票法真值表

Tab. 1 Truth table of voting method

欠完备自编码器	栈式自编码器	卷积自编码器	最终结果
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

在多数投票法中通过统计 3 种检测结果中的扣件状态,选择结果中的多数作为最终结果。表 1 为 3 种自编码器可能的预测值和最终结果,预测时通过真值表确定扣件最终状态。

3 实验结果与分析

本文实验涉及的软硬件平台及版本信息见表 2。

表 2 实验软硬件平台及版本信息

Tab. 2 Experimental software and hardware platform and version information

软件或硬件平台	类型或版本
GPU	Quadro P5000
CPU	Intel(R) Xeon(R) W-2125
RAM	128GB
Python	3.5
Ubuntu	16.04
Tensorflow	1.40
Keras	2.1.3

以下实验都基于相同的软硬件条件。本文将扣件分为正常状态和异常状态两类,如图 6 所示。异常状态扣件包括断裂、弹出等情况。数据集中图像

共 600 张,其中正常状态扣件 430 张,异常状态扣件 170 张。随机选取 260 张正常图像作为训练集,340 张作为测试集。测试集中正常状态扣件 170 张,异常状态扣件 170 张。

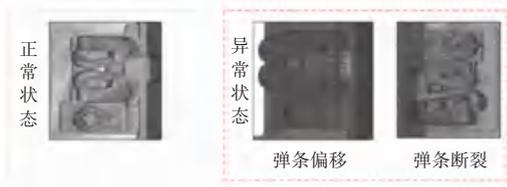


图 6 扣件状态

Fig. 6 The state of the fastener

本文用向量余弦值作为扣件状态的判定标准。在训练集中,任意选取一个图像作为参照,经过自编码器得到的特征 T 定义为基向量,训练集中其余图像经过自编码器得到的向量与基向量求余弦值,该余弦值为两张图像的相似度,范围为 $-1 \sim 1$,相似度越接近 1,向量代表的图像与基向量越相关。在不同算法中特征向量的提取不尽相同,对 3 个自编码器进行上述操作,得到的相似度最小值见表 3。

表 3 各算法的最小相似度

Tab. 3 Minimum similarity of each algorithm

算法	相似度最小值
栈式自编码器	0.975 6
卷积自编码器	0.983 5
欠完备自编码器	0.976 1

本文将测试图像的相似度阈值设置为训练集中相似度的最小值,即测试图像的自编码向量在相似度最小值和 1 之间时扣件为正样本,否则测试图像中的扣件为负样本。为进一步提高算法的正确率,本文使用投票法对 3 种算法给出的结果使用多数投票法进行投票。各个算法实验结果见表 4。

表 4 自编码器实验结果

Tab. 4 Experimental results of the Autoencoder

算法	实验结果	
	正样本检出数量	负样本检出数量
栈式自编码器	143	158
卷积自编码器	147	162
欠完备自编码器	135	164
多数投票法	159	166

实验证明,自编码器可以有效地提取扣件特征,将扣件图像信息映射到高维空间中,利用特征之间的相似度实现了在只训练正样本的情况下检测扣件状态。从负样本的检出数量来看,因为卷积编码器的卷积操作能更好地提取到扣件的空间特征,使

得3种自编码器中,卷积自编码器表现最好。使用多数投票法后,正样本的检出数量和负样本的检出数量都有所增加。

代先星使用VGG深度学习网络对扣件进行分类^[8],李永波使用HOG特征提取与SVM分类器结合方式检测扣件^[10],这两种算法的指标在各自的数据集中都较为优秀,为证明本文算法的优越性,在相同的软硬件条件和数据集下将本文方法与以上两种算法对比,并使用总体精度(Overall Accuracy, OA)、精确率(Precision)、虚警率(False alarm, FPR)和漏警率(Missing alarm, FNR)评估扣件状态检测算法,式(7)~(10)。

$$OA = \frac{TP}{N} \times 100\% \quad (7)$$

$$Precision = \frac{N_{TP}}{N_{TP} + N_{FP}} \times 100\% \quad (8)$$

$$FPR = \frac{N_{FP}}{N_{TN} + N_{FP}} \times 100\% \quad (9)$$

$$FNR = \frac{N_{FN}}{N_{TP} + N_{FN}} \times 100\% \quad (10)$$

其中, TP 表示算法被正确分类的个数; N 代表所有测试集的数量; N_{TP} 表示数据集中正样本数量; N_{FP} 表示负样本被检测为正样本的数量; N_{TN} 表示数据集中负样本的数量; N_{FN} 表示正样本被检测为负样本的数量。

因为扣件状态的检测关乎铁路安全,所以在保证整体精度的情况下,虚警率越低越好,算法分类实验结果见表5。由实验数据可知,传统的机器学习在没有负样本时,整体表现较差,不能达到扣件状态检测的效果;基于深度学习的VGG网络只通过学习正样本数据,总体精度不够,虚警率过高,不适用于扣件检测这种负样本较少的情况;本文方法不管是在总体精度还是虚警率方面比前两种方法都大有改观,综合各个评价指标,本文提出的扣件检测算法最好,实现了零样本检测铁路扣件状态。

表5 算法分类实验结果

算法	评价指标			
	总体精度	精确度	FPR	FPN
HOG+SVM[11]	59.71	66.93	33.07	23.77
VGG[8]	72.65	64.63	35.36	0
本文方法	95.59	97.70	2.30	6.47

4 结束语

本文提出了一种基于卷积自编码器的零样本扣

件检测方法,首先通过欠完备自编码器、栈式自编码器和卷积自编码器分别对扣件提取特征,再通过特征与基向量的相似度判定扣件的状态,最后通过多数投票法最终确定扣件状态,为铁路扣件检测提供了一种新的解决思路。本文算法在测试集中的整体精度为95.59%,精确度为97.70%,FPR为2.30%,FPN为6.47%,比传统方法和深度学习方法在评价指标方面有所提升,本文方法解决了利用机器视觉算法检测扣件状态时数据集不均衡的问题。

参考文献

- [1] 樊曦. 全国铁路营业里程突破14万公里[EB/OL]. http://www.xinhuanet.com/2020-08/08/c_1126342560.htm, 2020-8-8.
- [2] SHOLL H, AMMAR R, GREENSHIELDS I, et al. Application of Computing Analysis to Real-Time Railroad Track Inspection [C]// Automation Congress, 2006. WAC'06. World. IEEE, 2006: 1-6.
- [3] 王留军,张鹰. 基于机器视觉的车载式铁路轨道智能巡检系统研究[J]. 成都信息工程学院学报, 2016, 31(2): 185-189.
- [4] KHAN R A, ISLAM S, BISWA S R. Automatic detection of defective rail anchors [C]// IEEE International Conference on Intelligent Transportation Systems, IEEE, 2014.
- [5] 王强,李柏林,侯云,等. 一种改进的LBP特征实现铁路扣件识别[J]. 西南交通大学学报, 2018, 53(5): 893-899.
- [6] 林菲,杨子明,李永光,等. 钢轨扣件检测技术研究[J]. 中国铁路, 2019(6): 103-110.
- [7] 李哲. 基于GAN的不平衡数据增强算法及其应用研究[D]. 北京:北京交通大学, 2019.
- [8] 代先星,阳恩慧, WANG Kelvin Chenping. 铁路扣件缺陷自动检测研究进展[J]. 铁道科学与工程学报, 2017, 14(11): 2452-2459.
- [9] GIBERT X, PATEL V M, CHELLAPPA R. Deep Multi-task Learning for Railway Track Inspection[J]. IEEE Transactions on Intelligent Transportation Systems, 2017: 1-12.
- [10] 李永波,李柏林,熊鹰. 基于HOG特征的铁路扣件状态检测[J]. 传感器与微系统, 2013, 32(10): 110-113.
- [11] 路扬. 面向小样本不平衡数据的生物医学事件抽取方法研究[D]. 吉林:吉林大学, 2019.
- [12] NEKOOEIMEHR I, LAI - YUEN S K. Adaptive semi-supervised weighted oversampling (A-SUWO) for Imbalanced Datasets[J]. Expert Systems with Applications, 2016, 46(3): 405-416.
- [13] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors[J]. Nature, 1986, 32(6088): 533-536.
- [14] 赵洋,余翔湛,郝科委. 基于半监督学习的网络应用流识别研究[J]. 智能计算机与应用, 2020, 10(2): 355-359, 364.
- [15] 朱凯利,朱海龙,刘靖宇,等. 基于图卷积神经网络的交通流量预测[J]. 智能计算机与应用, 2019, 9(6): 168-170, 177.