

# 基于 CNN 的高速铁路侵限异物特征快速提取算法<sup>\*</sup>

王 洋,余祖俊,朱力强,郭保青

(北京交通大学机械与电子控制工程学院 北京 100044)

**摘 要:**高速铁路异物侵限检测系统用来检测是否有异物侵入高速铁路安全限界。为增加系统的可靠性,提出了一种基于卷积神经网络(CNN)的特征快速提取算法。针对特征计算速度缓慢的问题,提出简化的全连接网络结构;针对准确率因简化网络结构而下降的问题,提出将卷积层的卷积核进行预先训练;最后为防止因全连接而导致的对称性特征提取,提出加入稀疏性参数的快速特征提取算法。改进后的卷积神经网络,在保证准确率的基础上加快了计算速度,同时满足了实时性和高准确率的要求。实验表明处理单幅图像的速度为0.15 s,准确率为99.5%。

**关键词:**异物识别;卷积神经网络;预先训练卷积核;快速特征提取;稀疏编码

**中图分类号:** TP391.4 TH701 U215.8 **文献标识码:** A **国家标准学科分类代码:** 580.30 520.60

## Fast feature extraction algorithm for high-speed railway clearance intruding objects based on CNN

Wang Yang, Yu Zujun, Zhu Liqiang, Guo Baoqing

(School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing 100044, China)

**Abstract:** The high-speed railway clearance intrusion detection system is used to detect whether there is object intruding the safety clearance of the high-speed railway. To enhance the reliability of the system, a new CNN based fast feature extraction algorithm is proposed. Aiming at the problem of slow feature calculation speed, a simplified full connected network structure is proposed, and the structure of the neural network is simplified to two full connected convolutional layers. To avoid the accuracy decreasing caused by simplifying network structure, the convolutional kernels of the convolutional layers are pre-trained. Finally, in order to prevent the symmetric feature extraction caused by full connection, fast feature extraction algorithm with sparse parameters added is proposed, and the network is trained with sparse coding algorithm. The improved CNN accelerates the calculation speed while ensures the accuracy. At the same time, the new algorithm satisfies the requirements of real time capability and high accuracy. Experiment result shows that the speed of processing single image is 0.15 s and the accuracy is 99.5%.

**Keywords:** intruding object recognition; convolutional neural network (CNN); pre-trained convolutional kernel; feature map fast extraction; sparse coding

## 1 引 言

目前中国已经成为世界上高速铁路运营速度最高,运营里程最长以及在建规模最大的国家。高速铁路异物侵限检测系统作为铁路安全运营的重要保障,必须要在不同铁路的场景中出现的山体滑坡、落石、行人穿越等异物进行检测,判断是否有异物侵入铁路安全限界事件

的发生,特别是要保证系统在风、雨、雪、雾以及白天夜间等多种气候以及光照变化条件下仍能可靠运行。由于异物侵限事件具有突发性、无规律性和不可预测性等特点,且人工巡视等常规方法有着可靠性差、工作量大和漏检率高的缺点,因此高速铁路异物侵限检测系统作为高速铁路防灾安全监控系统的重要组成部分,对高速铁路的安全运营起到了至关重要的作用。高速铁路异物侵限检测系统主要设置在山体宜滑坡地段、公跨铁桥、隧道口、

公路铁路并行地段等,检测是否有异物侵入安全限界,为运营中心的调度、指挥及维护管理提供预警、报警信息,对高速列车的安全运行有着十分重要的意义。

根据检测原理的不同,高速铁路异物侵限检测系统<sup>[1]</sup>主要分为接触式和非接触式两种。其中接触式异物侵限检测系统的典型代表是以双电网传感器技术和光纤光栅传感器技术为主的防护网式检测系统,其主要原理是使用接触式传感器直接测量防护网自身的形变程度来判断是否有异物侵限事件的发生。非接触式异物侵限检测系统主要分两类:1)采用红外光幕<sup>[2]</sup>、激光扫描<sup>[3]</sup>等检测技术,其主要原理是通过测量红外或激光光幕被侵限异物遮挡的缺失量,进而计算侵限异物的位置、尺寸等参数;2)采用机器视觉技术<sup>[4]</sup>,其主要原理是通过对视频当前帧和背景帧进行差分计算来提取侵限异物<sup>[5]</sup>。机器视觉技术因为设备安装维护方便、可视化效果好等优点,在非接触式异物侵限检测系统中得到越来越多的应用。但是机器视觉技术在铁路场景中的应用受到环境因素影响较大,由于铁路场景的复杂性,包括每天 24 h 内不同的光照条件变化、相机剧烈抖动以及雨雪雾等恶劣的自然气候,严重影响帧间差分计算的结果,使得前景异物与背景难以准确分离,导致大量误报警情况的发生。为解决这些问题,文献[6]采用一维灰度投影快速去抖算法,进而通过背景差分、目标标记、合并与特性分析提取前景物体;文献[7]提出通过区域提取、背景建模、阈值分割和图像滤波来减少检测钢轨表面缺陷时所遇到的光照变化、反射不均、特征少等影响;文献[8]通过连续帧之间的背景融合算法来消除渐变光照强度的影响。但是现有算法并不能完全补偿列车驶过引起的相机剧烈抖动以及车体反光引起的光照突变,本文为解决这些问题,引入深度学习中的卷积神经网络(convolutional neural network, CNN)。CNN 有着良好的旋转、平移不变性,并且能够识别不同背景中的物体种类,能够减少相机抖动与光照变化的影响。文献[9]提出先通过四联通提取瑕疵区域,进而使用 CNN 来对火炮身管的损伤进行分类;文献[10]提出中央-周边差算法,提取铸件缺陷的微小特征,然后使用深度卷积神经网络对微小缺陷进行识别与分类。因此通过 CNN 来对报警图像中的侵限异物进行识别与分类,进而剔除正常行驶的列车及空场景等误报警情况,将大大减少高速铁路异物侵限检测系统的误报率,提高系统的可靠性。但是传统卷积网络的理论模型和在铁路场景中的实际应用之间差异性太大,并且现有应用 CNN 大多集中在分析处理事后汇总的图像数据而不是实时计算,无法直接套用于铁路场景的异物检测。传统的卷积神经网络是由 LeCun Y. L. 等人<sup>[11]</sup>提出,使用反向传播及梯度下降法来训练卷积神经网络模型,该模型在手写数字数据集 MNIST 上表现出当时最好的性能;机器学习领

军人物 Hinton G. E. 等人<sup>[12]</sup>明确提出了深度学习神经网络的两个主要特点:使用多层人工神经网络提取特征向量和逐层训练深度网络;为了能够通过多层网络逐层提取原始图像的特征信息并对这些特征信息进行组合生成特征地图,Krizhevsky A 等人<sup>[13]</sup>定义了深度卷积神经网络的经典结构 AlexNet,5 个局部连接的卷积层、3 个池化层和 3 个全连接的特征地图层,通过非对称局部连接的方式,强迫网络逐层提取不同的特征地图;Zeiler M. D. 等人<sup>[14]</sup>通过使用反卷积计算从而获取 AlexNet 中间网络层的可视化效果,进而分析每层学习到的特征;Szegedy C 等人<sup>[15]</sup>GoogleNet 于 ILSVRC2014 比赛中在 AlexNet 的基础上,提出了模块化卷积神经网络结构和最优局部稀疏结构;Microsoft 亚洲研究院的 He K 等人<sup>[16]</sup>的 SPPNet<sup>[16]</sup>采用了空间金字塔式的池化方法提取不同尺度的卷积特征。可以看出现有的 CNN 都集中在增加网络宽度与深度,设置多个卷积层和采样层,逐层提取原始图像的特征信息,进而得到高层的特征地图,然后用于分类,因此需要大量的计算时间,满足不了高速铁路异物侵限检测系统高实时性的要求。而简单的通过减少网络的深度与宽度来加快计算将会使网络对原始图像低层特征的提取不够精细,进而导致高层特征地图组合程度不足,最后导致分类准确率偏低。

针对卷积神经网络在铁路场景应用时难以平衡准确率和实时性的问题,本文提出了基于 CNN 的特征快速提取算法,预先训练卷积网络的卷积核,使其能更好地提取铁路场景信息,提高分类准确率;同时提出将非对称局部连接的多个卷积层和特征地图全连接层合并为全连接的第二卷积层,并加入稀疏性参数,在保证准确率的基础上减少了计算时间。

实验表明,改进后的卷积神经网络能够应对铁路特定场景中复杂的场景细节以及光照变化,能够更好地提取不同光照条件下物体的特征信息,并且能够在 0.15 s 的时间内处理单幅图像,满足较高的实时性要求,且分类准确率也保持在 99.5%。

## 2 基于 CNN 的特征快速提取网络模型

传统 CNN 尽管有着较高的识别准确率,在铁路场景中却鲜有应用,主要原因是其自身理论模型与在铁路场景中的实际应用之间存在较大的差异,导致计算量大幅增加,难以在保证高准确率的基础上实现实时性检测。针对这一问题,提出了改进的 CNN 模型,通过预先训练卷积核来提高准确率,简化网络结构来加快计算速度,从而快速准确地识别高速铁路场景中的各类物体,有效地平衡了准确率与计算时间。

2.1 建立高速铁路场景样本库

将传统 CNN 的样本集合如 ILSVRC 数据集与铁路场景原始图像做对比,如图 1 所示。由图 1 可以发现,铁路场景原始图像有着尺寸较大、物体占据图像比例较小、包含多个物体以及物体种类总数较少等特点。

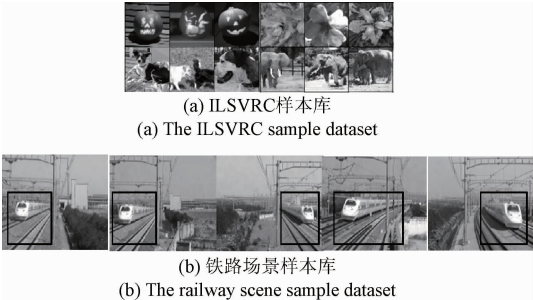


图 1 ILSVRC 与铁路场景样本库对比

Fig.1 Comparison between ILSVRC and railway scene sample datasets

表 1 所示为传统 CNN 样本图像与铁路场景原始图像之间的差异。铁路场景中,安全限界范围(见图 1 黑色方框标识范围)内主要包含了轨道板、接触网支架、行驶列车、行人及异物等多类物体,直接将原始图片作为传统 CNN 的输入层,将会大大增加网络参数数量,增加训练与计算时间。针对这一问题,提出建立高速铁路场景中各类物体的样本库,训练中尺寸的特征快速提取 CNN 模型来扫描安全限界,并对限界以内的物体进行分类。

表 1 传统 CNN 样本与铁路场景图像的差异

Table 1 The difference between traditional CNN sample dataset and railway scene sample dataset

	包含物体	分类总数	背景影响因素
传统 CNN 样本	1 个	1 000	物体占据原始图像中央 占据绝大部分比例 背景影响较小
铁路场景样本	多个		轨道板 物体占据比例较小 接触网 光照突变 列车 相机抖动 侵线异物 铁路沿线环境差异性大 行人 不同物体交叉重叠

从目前公开的资料库中并没有查找到有关于铁路场景中各类物体集合的数据库,因此在沪杭高速铁路沿线建立了基于机器视觉技术的异物侵限检测系统,并从具有代表性的地点采集铁路场景原始图像,如公跨铁桥、弯道、隧道进出口、同一位置的上下行方向等。然后提取铁路场景中的中尺寸大小的轨道板、列车、接触网、异物及行人等图像建立样本库。提取过程如图 2 所示。

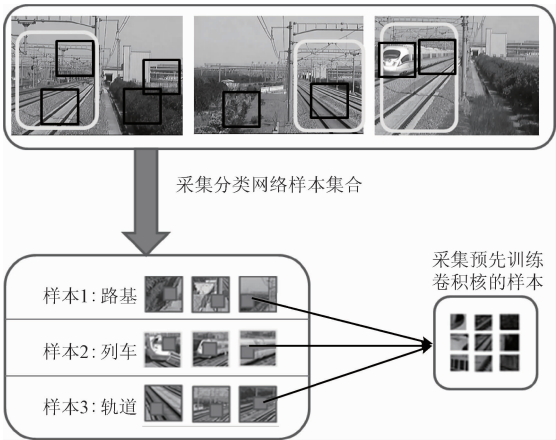


图 2 CNN 样本集合与卷积核预先训练样本集合的采集

Fig.2 Acquisition of CNN sample set and the sample set with pre-trained convolutional kernel

1) 使用的图像是整幅铁路场景的全尺寸图像(分辨率为  $384 \times 528$ , RGB 三色),图像中人工划分出安全限界区域(图见 2 白色圆角方框区域),在安全限界范围内使用中尺寸的图像采集框(图 2 中黑色方框,分辨率为  $65 \times 65$ ) 在安全限界范围以内随机位置采集轨道板和接触网等样本;在安全限界范围以外随机位置采集绿化带、建筑等并统归于背景一类。

2) 采集大小不一,行驶方向及姿态各不相同的高速列车、穿越行人及异物等样本图像,统一调整为标准大小(分辨率为  $65 \times 65$ )。

3) 最终得到的每类样本各有 8 000 张图像,将来自 1~6 号相机的样本集合用于训练和测试网络,而 7、8 号相机只用做测试样本集合,以验证系统的通用性。

4) 从训练样本图像中随机位置用小尺寸采样框(分辨率为  $8 \times 8$ ) 采集 10 000 张图像补丁用来预先训练第一卷积层的卷积核。

引入交叉验证,将高速铁路场景各类物体样本图像分为 3 组:训练组占 60%、测试组占 20%、交叉验证组占 20%。使用训练组与交叉验证组作为训练数据集训练网络参数,将测试组与交叉验证组作为测试数据集,分别计算训练集合和测试集合的代价函数。如图 3 所示,当卷积核的数量取值小于最优值时,两组集合的代价函数值都偏高属于欠拟合;当卷积核的数量大于最优值时,测试集合的代价函数值将反弹意味着出现了过拟合现象。

$$J_{\text{train}}(\theta) = \frac{1}{2m_{\text{train}}} \sum_{i=1}^{m_{\text{train}}} (h_{\theta}(x^{(i)}) - y^{(i)})^2 \tag{1}$$

$$J_{\text{test}}(\theta) = \frac{1}{2m_{\text{test}}} \sum_{i=1}^{m_{\text{test}}} (h_{\theta}(x^{(i)}) - y^{(i)})^2 \tag{2}$$

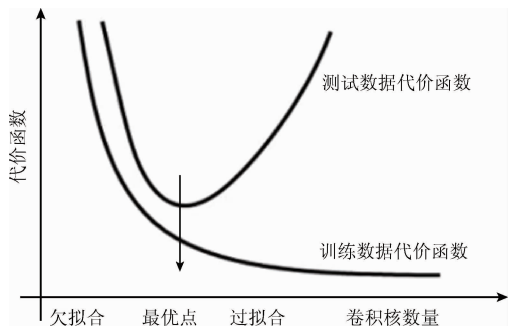


图 3 卷积核个数的选取

Fig. 3 Selection of the number of the convolutional kernels

由于后期将采用梯度下降法来更新卷积核的值,为快速准确的实现梯度下降,通过式(3)对样本集合进行均值化处理。其中  $x_j^{(i)}$  表示第  $i$  个样本的第  $j$  维特征值,总计  $m$  个样本,  $x_{j,\max/\min}^{(i)}$  表示样本在该特征维度上出现的最大值或最小值。

$$x_j^{(i)} = \frac{x_j^{(i)} - \frac{1}{m} \sum_{n=1}^m x_j^{(n)}}{x_{j,\max} - x_{j,\min}} \quad (3)$$

2.2 特征快速提取网络模型的结构

由于 CNN 要实现只包含一个物体的单幅图像的 1 000 分类,因此必须采用较深的网络结构来增加特征地图的可分性,而铁路场景的特点有两个:包含物体总种类较少但同时包含多个物体;场景组成元素单一但变化很大。高速铁路场景一般只包含轨道板、接触网、路基设施等,但整条线路上的同类物体在不同环境的光照条件下变化较大。本文针对同类物体变化较大的特点,提出预先训练卷积核来提取更好的原始图像特征信息以便提高特征地图可分性;针对分类总数较少的特点,提出简化网络结构以加快计算速度。最终实现在减少计算时间的同时保证准确率。

传统的深度卷积神经网络(deep CNN, DCNN)通常包含多个卷积层与采样层,如 AlexNet 采用的就是一个包含 5 个局部连接卷积层、3 个采样层和 3 个特征地图全连接层的深度卷积神经网络,其中前 5 层局部连接卷积层的卷积核是随机初始化的,采用非对称式局部连接,目的在于强制前 5 层卷积网络层提取不同的特征地图,并通过 GPU 显卡并行计算提高计算速度。考虑到铁路场景中的实际应用及成本问题,预报警图象将在铁路沿线监控中心的计算机上进行处理,无法大规模使用 GPU 显卡,因此只能通过简化神经网络结构来减少计算时间。因此将网络结构精简为 2 层卷积层以及两层采样层和一层分类层,如图 4 所示,计算时间维持在 0.15 s,但是准确率骤减到了 72%。在不增加运算时间的情况下为了提高准确率,就必须对卷积核进行预训练,使其

对原始图像中的色彩以及特征点、特征边缘更敏感,从而提取出可分性更好的特征地图,才能提高系统的准确率。

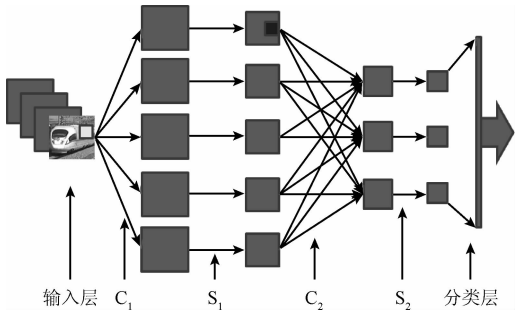


图 4 网络结构示意图

Fig. 4 Schematic diagram of the network architecture

改进后的卷积神经网络结构,第一卷积层  $C_1$  的输入层为 RGB 三色图像,分辨率  $65 \times 65 \times 3$ ,如图 6(a)、(b)所示,每个卷积核大小为  $8 \times 8 \times 3$ ,如图 6(a)、(b)所示共 100 个卷积核,扫描步长为 1 个像素,与原始图像进行卷积运算后,第一卷积层将输出 100 张分辨率为  $58 \times 58$  大小的特征地图。

第一采样层  $S_1$  采用 Meanpooling 方式的均值采样,采样框大小为  $2 \times 2$ ,步长为 2 个像素,因此 100 张特征地图将一一对应生成 100 张分辨率为  $29 \times 29$  大小的特征地图,如图 5 所示。



图 5 原始图像及第一采样层  $S_1$  得到 100 张中的 4 张

Fig. 5 Original image and the 4 in 100 feature maps obtained in first sampling layer  $S_1$

第二卷积层  $C_2$  将采用全连接,生成 10 张分辨率为  $22 \times 22$  的特征地图,因此需要预先训练 10 个大小为  $8 \times 8 \times 100$  的卷积核。 $C_2$  生成的可视化效果较好的一幅特征地图如图 6(c)、(d)所示。

第二采样层  $S_2$  依旧采用 Meanpooling 方式,10 张特征地图将一一对应生成 10 张分辨率为  $11 \times 11$  大小的特征地图。采样框大小为  $2 \times 2$ ,步长为 2 个像素。

最后将这 10 张特征地图,像素尺寸  $11 \times 11$  展开成列向量,输入到最后一层的 Softmax 分类层。

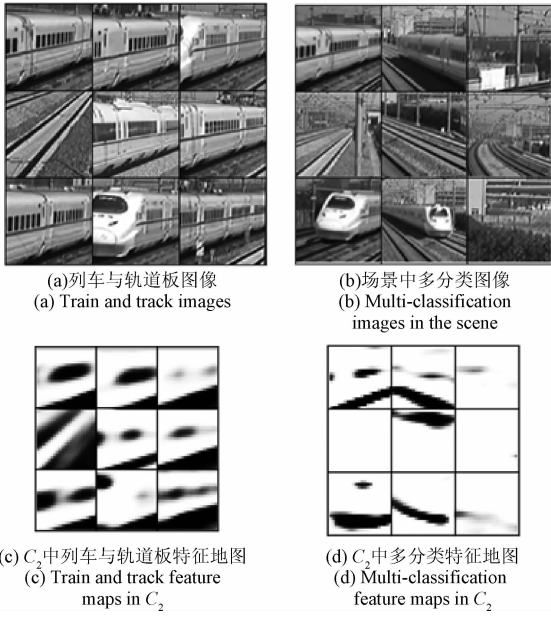


图6 多分类的原始图像及第二卷积层的特征地图

Fig. 6 Multi-classification original images and the feature maps in second convolutional layer  $C_2$ 

与传统卷积网络相比,改进的 CNN 的卷积层采用了预先训练的卷积核以便更好地提取铁路场景信息;同时整体网络结构仅采用两层卷积网络,可以快速提取原始图像特征,在保证准确率的基础上加快了特征提取速度,满足实时计算的要求。

### 3 基于稀疏编码的快速特征提取算法

如上文所述,CNN 在铁路场景中的应用最大的问题在于平衡计算时间和准确率。新卷积神经网络最终将确定卷积核的大小为  $8 \times 8$ ,并对卷积核进行预先训练来弥补失去的准确率;同时采用两个卷积层、两个采样层以及一个分类层的结构,减少计算时间。每个卷积层的卷积核将采用无监督自学习算法来进行预训练,最后使用预先训练的卷积核构建特征快速提取与分类网络结构。

#### 3.1 基于 Autoencoder 网络的卷积核训练算法

卷积核的预先训练如图 7 所示,将从铁路场景中采集到的彩色小图像补丁(像素大小为  $8 \times 8 \times 3$ )使用用来预训练第一层卷积网络的卷积核。

采用第 2 节中提出的随机提取样本方法,采集到了 10 000 张的小图像块(像素尺寸  $8 \times 8$ ,RGB 三色),这些图像块随机采取自铁路场景以及高速列车车身。将每个图像补丁展开成列向量(大小为  $8 \times 8 \times 3$ )输入到 Autoencoder 神经网络中,网络设计采用 3 层网络(见图 7),输出输出层都设定为  $8 \times 8 \times 3$  个神经元,隐含层神经元节点

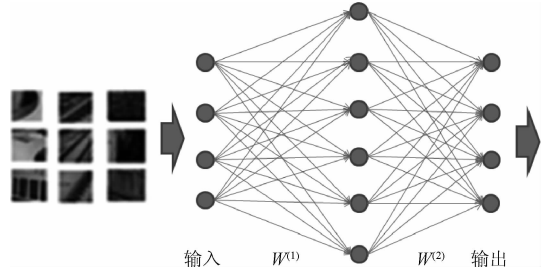


图7 卷积核预训练过程

Fig. 7 Convolutional kernel pre-trained process

为 100 个,设定输出层输出等于输入层的输入,采用梯度下降法来无监督自学习网络传递参数  $W_{j,i}^{(1)}$  和  $W_{m,j}^{(2)}$ ,上标  $l = 1, 2$  表示网络层数,下标  $i, j$  表示相应位置的传递参数。

卷积核的预先训练中每一次迭代按照以下过程更新网络参数  $W, b$ 。

$$\begin{cases} W_{i,j}^{(l)} = W_{i,j}^{(l)} - \alpha \frac{\partial J(W, b)}{\partial W_{i,j}^{(l)}} \\ b_i^{(l)} = b_i^{(l)} - \alpha \frac{\partial J(W, b)}{\partial b_i^{(l)}} \end{cases} \quad (4)$$

式中:  $W_{i,j}^{(l)}$  表示从第  $l$  层第  $i$  个神经元到第  $l+1$  层第  $j$  个神经元的传递参数,  $\alpha$  是学习率,  $b$  是偏置量,  $J(W, b)$  代价函数。无监督自学习过程如下。

1) 每个样本做前向传导

隐含层输入  $z_i^{(2)}$ :

$$z_i^{(2)} = \sum_{j=1}^{n_m} W_{i,j}^{(1)} a_j^{(1)} + b_i^{(1)} = \sum_{j=1}^{n_m} W_{i,j}^{(1)} x_j + b_i^{(1)} \quad (5)$$

隐含层输出(激活值)  $a_i^{(2)}$ :

$$a_i^{(2)} = f(z_i^{(2)}) \quad (6)$$

输出层输入  $z_i^{(3)}$ :

$$z_i^{(3)} = \sum_{j=1}^{n_{hidden}} W_{i,j}^{(2)} a_j^{(2)} + b_i^{(2)} \quad (7)$$

输出层输出  $a_i^{(3)}$ :

$$a_i^{(3)} = f(z_i^{(3)}) \quad (8)$$

最终输出记作:

$$a^{(3)} = h_{W,b}(x) \quad (9)$$

其中:

$$f(z) = \frac{1}{1 + \exp(-z)} \quad (10)$$

2) 计算输出层的代价函数

单个样本  $i$  的代价函数:

$$J(W, b; x^{(i)}, y^{(i)}) = \frac{1}{2} \|y^{(i)} - h_{W,b}(x^{(i)})\|^2 \quad (11)$$

对  $m$  个样本  $(x^{(i)}, y^{(i)})$  样本集来说其代价函数为:

$$J(W, b) = \left[ \frac{1}{m} \sum_{i=1}^m J(W, b; x^{(i)}, y^{(i)}) \right] + \frac{1}{2} \lambda \left( \sum_{j=1}^{n_m} \sum_{i=1}^{n_{hidden}} (W_{i,j}^{(1)})^2 + \sum_{j=1}^{n_{hidden}} \sum_{i=1}^{n_m} (W_{i,j}^{(2)})^2 \right) =$$



$$\left[ \frac{1}{m} \sum_{i=1}^m \frac{1}{2} h_{w,b}(x^{(i)}) - y^{(i)2} \right] + \frac{1}{2} \lambda \left( \sum_{j=1}^{n_{in}} \sum_{i=1}^{n_{hidden}} (W_{i,j}^{(1)})^2 + \sum_{j=1}^{n_{hidden}} \sum_{i=1}^{n_{out}} (W_{i,j}^{(2)})^2 \right) \quad (12)$$

其中,令  $y=x$ ,则该网络将完成无监督自学习,  $\lambda$  为权重衰减项,目的是减小权重的幅度,防止过度拟合。

### 3) 反向求导计算各层残差

$$\delta_i^{(i)} = \sum_{j=1}^{n_{i+1}} \delta_j^{(i+1)} W_{j,i}^{(i)} f'(Z_i^{(i)}) \quad (13)$$

$$d_i^{(3)} = -[y_i - f(Z_i^{(3)})] f'(Z_i^{(3)}) = -(x_i - a_i^{(3)}) f'(Z_i^{(3)}) \quad (14)$$

$$\delta_i^{(2)} = \frac{\partial J(W, b; x, y)}{\partial Z_i^{(2)}} = - \sum_{j=1}^{n_{out}} (y_j - f(Z_j^{(3)})) f'(Z_j^{(3)}) \frac{\partial (Z_j^{(3)})}{\partial Z_i^{(2)}} = \sum_{j=1}^{n_{out}} \delta_j^{(3)} W_{j,i}^{(2)} f'(Z_i^{(2)}) \quad (15)$$

$$\delta_i^{(1)} = \sum_{j=1}^{n_{hidden}} \delta_j^{(2)} W_{j,i}^{(1)} f'(Z_i^{(1)}) = \sum_{j=1}^{n_{hidden}} \delta_j^{(2)} W_{j,i}^{(1)} f'(x_i) \quad (16)$$

该残差表明了该节点对最终输出值的残差所产生的影响。

### 4) 更新 $W, b$

$$\begin{cases} W_{i,j}^{(l)} = W_{i,j}^{(l)} - \alpha \frac{\partial J(W, b)}{\partial W_{i,j}^{(l)}} \\ b_i^{(l)} = b_i^{(l)} - \alpha \frac{\partial J(W, b)}{\partial b_i^{(l)}} \end{cases} \quad (17)$$

$$\begin{cases} \frac{\partial J(W, b)}{\partial W_{i,j}^{(l)}} = \frac{\partial J(W, b)}{\partial Z_i^{(l+1)}} \frac{\partial Z_i^{(l+1)}}{\partial W_{i,j}^{(l)}} = \delta_i^{(l+1)} a_i^{(l)} + \lambda W_{i,j}^{(l)} \\ \frac{\partial J(W, b)}{\partial b_i^{(l)}} = \frac{\partial J(W, b)}{\partial Z_i^{(l+1)}} \frac{\partial Z_i^{(l+1)}}{\partial b_i^{(l)}} = \delta_i^{(l+1)} \end{cases} \quad (18)$$

5) 重复步骤1)~4)直到残差小于阈值即认为参数收敛

将训练好的第一层传递参数  $W_{j,8 \times 8 \times 3}^{(1)}$  组合成第  $j$  个  $8 \times 8 \times 3$  的卷积核(见图8)作为第一卷积层的卷积核。

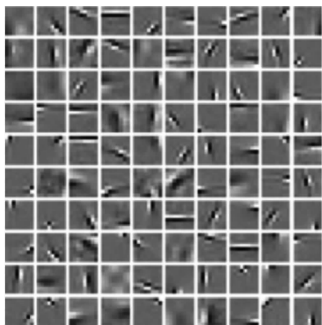


图8 预先训练的卷积核

Fig. 8 The pre-trained convolutional kernels

## 3.2 加入稀疏项的卷积核训练算法

改进后的特征快速提取 CNN,将不再采用传统卷积网络整体训练的方式,而是采用堆栈式网络训练模式,每层单独训练。第一卷积层的卷积核是通过采用 Autoencoder 的方式无监督自学习而来;第二卷积层卷积核也采用无监督自学习完成,设计一个新的3层卷积神经网络,将输入层  $X_i^{(1)}$ 、输出层  $X_m^{(3)}$  都定为第一采样层 S1 输出的 100 张特征地图,隐含层卷积核个数与第二卷积层的个数相同,为  $10 \times 100$  个  $8 \times 8$  大小的卷积核。训练过程采用 3.1 节的无监督自学习流程,其中前向传输变为式(19)、(20),残差计算及反向传输为式(21)、(22),旋转卷积核完成反卷积运算  $\text{rot180}(K_{j,i}^l)$ 。其中第一层的第  $j$  个输入  $u_j^{(l)}$ 、第  $m$  个输出  $x_j^{(l)}$  为:

$$\begin{cases} u_j^{(2)} = \sum_{i=1}^{100} x_i^{(1)} \otimes K_{j,i}^{(1)} + b_j^{(1)} \\ u_m^{(3)} = \sum_{j=1}^{10} x_j^{(2)} \otimes K_{m,j}^{(2)} + b_m^{(2)} \end{cases} \quad (19)$$

$$\begin{cases} x_j^{(2)} = f(\sum_{i=1}^{100} (u_i^{(2)} \otimes K_{j,i}^{(2)}) + b_{j,i}^{(2)}) \\ x_m^{(3)} = f(\sum_{j=1}^{10} (u_j^{(3)} \otimes K_{m,j}^{(3)}) + b_{m,j}^{(3)}) \end{cases} \quad (20)$$

$$\delta_i^{l-1} = \delta_j^l \otimes \text{rot180}(K_{j,i}^l) f'(u^{(l)}) \quad (21)$$

$$\frac{\partial J}{\partial K_{i,j}^l} = x_i^{l-1} \otimes \text{rot180}(\delta_j^l) \quad (22)$$

将代价函数式(9)改为式(23)并增加了稀疏项,即式(24)。

$$J(K) = \left[ \frac{1}{s} \sum_{i=1}^s \frac{1}{2} (h_{k,b}(x_i^{(1)}) - x_i^{(3)})^2 \right] + \beta \sum_{j=1}^{10} \left( \rho \log \frac{\rho}{\eta_j} + (1 - \rho) \log \frac{1 - \rho}{1 - \eta_j} \right) \quad (23)$$

$$\eta_j = \frac{1}{s} \sum_{i=1}^s \sum_{u=1}^{29} \sum_{v=1}^{29} x_j^{(2), (i)} \quad (24)$$

式中:  $\beta$  为控制稀疏性惩罚因子的权值,  $\rho$  为稀疏编码的稀疏性参数,取接近 0 的较小数值 0.05,  $\eta_j$  为隐含层生成的第  $j$  张特征地图在  $S$  个样本的平均激活度。代价函数收敛以后将  $K_{j,i}^{(1)}$  作为第二卷积层的卷积核。

改进后的快速特征提取神经网络的第一卷积层将采用 3.1 节训练的卷积核,第二卷积层将采用 3.2 节训练的卷积核,加上采样层后完成特征快速提取,使用带有标签的训练图像集合来训练 Softmax 分类层。

## 4 实验结果

高速铁路异物侵限检测系统会给出大小不一的预报警图像,最小的为仅包含异物的局部安全限界区域(分辨率为  $64 \times 64$  和  $96 \times 96$ ),最大的则是整个安全限界区域(分辨率为  $96 \times 96$  和  $128 \times 128$ )或整个站台。因此将

对每张预报警图像通过分割与缩放为1~5张图片,依次进行分类与识别。

为验证改进后的CNN对铁路场景中物体的识别与分类的准确率,与传CNN做对比实验。如表2所示,在相同网络结构的条件下,使用相同的铁路场景图像集合作为测试集,结果表明改进后的CNN具有较高的准确率。

表2 不同网络类型的准确率

Table 2 The accuracies of different types of networks	第一层卷积层	第二层卷积层	准确率/%
	卷积核个数	卷积核个数	
传统CNN网络	30	20×30	75
传统CNN网络	100	10×100	72.5
改进后	100	10×100	99.5
改进后	100	0	97.2

进一步的对比实验,如表3所示。在采用预先训练卷积核的基础上,对比不同网络结构的处理时间及准确率。其中结构1采用的是两层卷积层、两层采样层及1层分类层,结构2采用的是3层卷积层、3层采样层及1层分类层。结果表明采用两个卷积层以及8×8大小的卷积核,使用分辨率为65×65的图像尺寸,有着较好的表现。且根据预报警图像大小的不同,单幅报警图像背缩放分割扫描成1~5张图像,对应的准确率在0.995~0.995<sup>5</sup>之间,处理时间在0.15~0.75 s,为最优的选择。

表3 不同网络结构对比实验结果

Table 3 Comparison experiment result of different network structures					
	卷积核大小	图像尺寸	训练	测试	
			平均迭代时间/s	准确率/%	单幅处理时间/s
结构1	3×3	62×62	1715	96	0.04
	5×5	64×64	3276	98.5	0.08
	8×8	65×65	6026	99.5	0.15
	8×8	97×97	10938	98	0.27
结构2	8×8	97×97	10670	92	0.26

视频监控系统根据可疑物体的大小将生成预报警图像(分辨率为64×64、96×96和128×128),下面以误报警的列车图像(见图9)和空场景(见图10)以及行人图像(见图11)为例,将系统的工作流程分为以下5个步骤。

1) 在图9(a)中,以黑色方框表示的安全限界内出现的列车为例,根据列车由远及近将生成3种预报警图像:远距离列车图像分辨率为64×64,中距离分辨率为96×96(见图9(b)上图)、近距离分辨率为128×128(见图9(b)下图)。分辨率为64×64的预报警图像将直接进行分类识别,剩余两种预报警图像将通过进一步缩放与分割扫描来生成一组5张图像进行分类,如图9(c)、(d)流程所示。

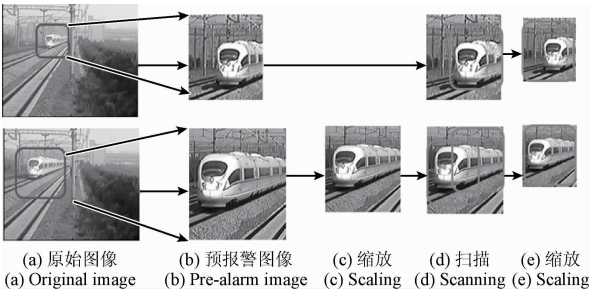


图9 识别列车流程

Fig. 9 Flow chart of recognition of the railway train

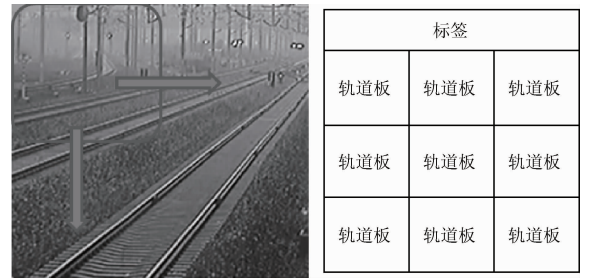


图10 剔除误报警图像流程

Fig. 10 Eliminating the false alarm images of the empty clearance using labels

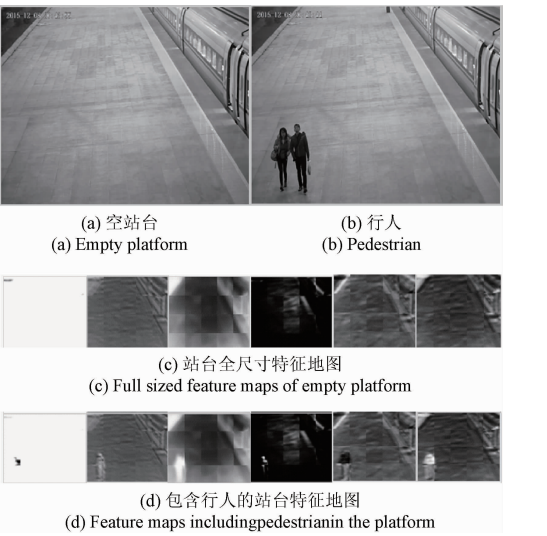


图11 行人与站台原始图像及特征地图

Fig. 11 Original images and feature maps of pedestrian and platform

2) 在图9(d)中白色方框表示扫描框,按照步长32个像素来扫描预报警图像,这一步将产生4张图像。

3) 将图像进一步缩小为像素尺寸64×64大小,如图9(e)所示,这样一幅列车图像就变为包含5张图像的图像集。

4) 将图像集输入到改进的CNN中获取每张图像的分类标签。

5) 综合 5 张图像的分类标签以及坐标位置,最终识别和确认出图像中的列车的存在与位置。

空场景误报警的情况是指预报警图像中并不存在列车和侵限异物,这种情况是由于光照强度变化或者相机抖动的引起的场景变化而导致的,因此需要对预报警图像进行处理并剔除误报警图像。过程将重复上述的 5 个步骤,按照步长 32 个像素来逐行逐列的扫描误报警图像,当分类标签都被归类为属于安全限界场景的“轨道板”标签时剔除误报警情况以降低误报率,如图 10 所示。

针对行人检测主要是防止旅客行人出现在禁止通行的区域,如站台两侧延伸部。同样采用扫描整体场景的方法确定前景物体位置并识别与分类。如图 11 所示,与空站台的全尺寸的特征地图做对比,可以识别并精确定位行人。图 11(c)、(d)所示为站台上无人和有人的多张特征地图,可以发现其中第一张特征地图中有人无人的对比最为明显。

高速铁路异物侵限检测系统在夜间运行时,由于铁路线路大部分位于野外,没有任何人工照明设备,普通相机在夜间拍摄到的将是几乎全黑的图像。考虑到工程量、工程造价和对信号灯影响等原因无法沿线安装人工照明设备,因此将普通相机更换为星光相机。星光相机是一种不需要任何辅助光源就能在微光条件下(如月光照射)采集清晰彩色图像的相机。在测试系统中采用了中星电子生产的 VS-8200 系列相机,其视频分辨率为 $1\,280\times 720$ ,25 帧/s,在光照强度为 0.001 Lux 下仍能采集清晰彩色画面的视频,使得高速铁路异物侵限检测系统在无人工照明条件的野外夜间仍然能够正常工作,相应场景拍摄效果如图 12 所示。由于夜间列车车体特征会发生巨大变化(见图 12(c)),列车的车头车尾以及车身的灯光会成为列车的主要特征,因此白天和夜间的列车样本差异性太大,导致系统在夜间工作时准确率下降。针对这一问题,在原有一组列车模板的基础上,增加了一组夜间车体模板。通过换用星光相机和增加夜间车体模型提高了系统在夜间的工作能力。不同工作环境下系统的准确率如表 4 所示。

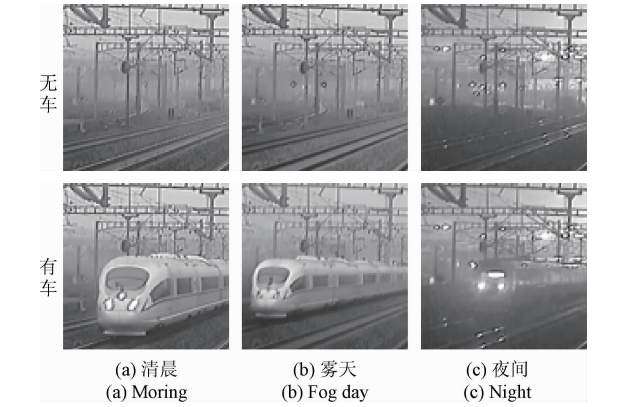


图 12 清晨、雾天以及夜间安全限界场景有车无车对比  
Fig. 12 The safety clearance areas with/without the train in the morning, on foggy day and at night

表 4 夜间工作环境下的准确率  
Table 4 The accuracy under night working environment

相机	光照条件及列车模型数量	准确率/%
星光相机	白天,单个列车模型	99.5
星光相机	白天+夜间,单个列车模型	88.2
星光相机	白天+夜间,两个列车模型	97

## 5 结 论

本文针对传统 CNN 在铁路场景中应用时所遇到的实时性和高准确率难以同时满足的问题,提出了适合铁路场景的卷积神经网络结构和算法,实验结果表明处理速度为 0.15 s,白天准确率为 99.5%,全天准确率为 97%,能够快速准确完成高速铁路场景中侵限异物的识别与分类任务。同时在所有测试实验过程中,是将现场 8 台相机中的 6 台相机数据作为纯训练样本,剩余 2 台的数据用作为纯测试样本。在试验中,在网络参数被训练好之后,将其导入测试相机系统中,对测试相机在其他铁路地段拍摄的图像进行分类与识别。实验结果表明高速铁路异物侵限检测系统具有良好的通用性。

由于高速铁路异物侵限检测系统监控的重点区域是安全限界以内,因此安全限界以外的区域信息是无用冗余的,在实验过程中,目前是通过手工划定检测范围的方法来减少系统的计算量,在之后的研究中,将进一步采用增加网络训练样本种类,对整个场景进行识别与分割,自动划分安全限界的范围。

## 参考文献

[1] 于革,贾利民,秦勇,等. 视频监控铁路限界内逗留物体的检测方法[J]. 中国铁道科学, 2013, 34(4): 105-109.  
YU G, JIA L M, QIN Y, et al. Method for detecting loitering (abandoned) objects within railway clearance by video surveillance [J]. Journal of the China Railway Society, 2013, 34(4):105-109.

[2] OH S, KIM G, LEE H. A monitoring system with ubiquitous sensors for passenger safety in railway platform[C]. Power Electronics, 2007:289-294.

[3] 郭保青,朱力强,史红梅. 基于快速 DBSCAN 聚类的铁路异物侵限检测算法[J]. 仪器仪表学报, 2012, 33(2):241-247.  
GUO B Q, ZHU L Q, SHI H M. Intrusion detection algorithm for railway clearance with rapid DBSCAN clustering[J]. Journal of Scientific Instrument, 2012, 33(2):241-247.

[4] 史红梅,柴华,王尧,等. 基于目标识别与跟踪的嵌入式铁路异物侵限检测算法研究[J]. 铁道学报, 2015(7):58-65.  
SHI H M, CHAI H, WANG Y, et al. Study on railway



- embedded detection algorithm for railway intrusion based on object recognition and tracking[J]. Journal of the China Railway Society, 2015(7):58-65.
- [5] 武治国, 王明佳, 丁南南. 基于视觉分析的机场跑道异物检测技术研究[J]. 仪器仪表学报, 2015, 36(增刊1):62-67.  
WU ZH G, WANG J M, DING N N. FOD detection and identification technology research based on the visual analysis[J]. Journal of Scientific Instrument, 2015, 36 (Suppl. 1):62-67.
- [6] 郭保青, 杨柳旭, 史红梅, 等. 基于快速背景差分的高速铁路异物侵入检测算法[J]. 仪器仪表学报, 2016, 37(6):1371-1378.  
GUO B Q, YANG L X, SHI H M, et al. High-speed railway clearance intrusion detection algorithm with fastback ground subtraction [J]. Journal of Scientific Instrument, 2016, 37(6):1371-1378.
- [7] 贺振东, 王耀南, 刘洁, 等. 基于背景差分的高铁钢轨表面缺陷图像分割[J]. 仪器仪表学报, 2016, 37(3):640-649.  
HE ZH D, WANG Y N, LIU J, et al. Background differencing-based high-speed rail surface defect image segmentation[J]. Journal of Scientific Instrument, 2016, 37(3):640-649.
- [8] 董宏辉, 葛大伟, 秦勇, 等. 基于智能视频分析的铁路入侵检测技术研究[J]. 中国铁道科学, 2010, 31(2):121-125.  
DONG G H, GE D W, QIN Y, et al. Research on railway invasion detection technology based on intelligent video analysis [J]. China Railway Science, 2010, 31(2):121-125.
- [9] 汤一平, 韩国栋, 鲁少辉, 等. 基于 CNN 的火炮身管全景图像疵病识别方法[J]. 仪器仪表学报, 2016, 37(4):871-878.  
TANG Y P, HAN G D, LU SH H, et al. Flaw recognition method for gun barrel panoramic images based on convolutional neural network[J]. Journal of Scientific Instrument, 2016, 37(4):871-878.
- [10] 余永维, 杜柳青, 曾翠兰, 等. 基于深度学习特征匹配的铸件微小缺陷自动定位方法[J]. 仪器仪表学报, 2016, 37(6):1364-1370.  
YU Y W, DU L Q, ZENG C L, et al. Automatic localization method of small casting defect based on deep learning feature [J]. Journal of Scientific Instrument, 2016, 37(6):1364-1370.
- [11] LECUN Y L, BOSER B, DENKER J S, et al. Handwritten digit recognition with a back-propagation network[C]. Advances in Neural Information Processing Systems. Morgan Kaufmann Publishers Inc., 1990:465.
- [12] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786):504-7.
- [13] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2012, 25(2):2012.
- [14] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks [J]. Computer Vision, 2013, 8689:818-833, doi:10.1007/978-3-319-10590-1\_53.
- [15] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions [J]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015:7-12.
- [16] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 37(9):1904-16.

## 作者简介



王洋, 分别在 2008 年和 2010 年于北京交通大学获得学士学位和硕士学位, 现为北京交通大学博士研究生, 主要研究方向为机器视觉检测及神经网络。

E-mail: 12116331@bjtu.edu.cn

**Wang Yang** received his B. Sc. and M. Sc. degrees both from Beijing Jiaotong University in 2008 and 2010, respectively. Now he is pursuing Ph. D. degree in Beijing Jiaotong University. His main research interest includes machine vision detection technology and neuron network.



余祖俊, 2008 年于北京交通大学获得博士学位, 现为北京交通大学教授、博士生导师, 主要研究方向为测控系统设计。

Email: zjyu@bjtu.edu.cn

**Yu Zujun** received his Ph. D. degree from Beijing Jiaotong University in 2008. Now he is a professor and Ph. D. supervisor in Beijing Jiaotong University. His main research interest is measurement and control system.



郭保青(通讯作者), 2009 年于北京交通大学获得博士学位, 现为北京交通大学副教授, 研究方向为铁路基础设施检测及机器视觉检测。

E-mail: bqguo@bjtu.edu.cn

**Guo Baoqing** (Corresponding author) received his Ph. D. degree from Beijing Jiaotong University in 2009. Now he is an associate professor in Beijing Jiaotong University. His main research interest includes detection and diagnosis of railway infrastructure, and machine vision detection technology.