

基于卷积神经网络的煤炭运载车辆识别

马传香^{1,2} 汪炆杰¹ 王旭¹

1 湖北大学计算机与信息工程学院 武汉 430062

2 湖北省教育信息化工程研究中心 武汉 430062

(mcx838@hubu.edu.cn)

摘要 为了杜绝或避免矿产品资源如煤炭、砂石矿等行业因不开票而导致偷税漏税现象的发生,利用深度卷积神经网络自动识别空车重车是一种有效途径。本文在 AlexNet 模型基础上,针对空车重车图像的差异性,提出 5 种改进思路,最终得到一种基于 maxout+dropout 的 6 层卷积神经网络的结构。对 34220 张空车重车图片的测试结果表明,模型在准确度、敏感度、特异性、精度等方面都取得了良好的效果。此外,模型还具有高度的鲁棒性,可以成功识别大量不同角度和不同场景的空车重车图像。

关键词:空车重车识别;深度学习;卷积神经网络;AlexNet;maxout

中图法分类号 TP391

Identification of Coal Vehicles Based on Convolutional Neural Network

MA Chuan-xiang^{1,2}, WANG Yang-jie¹ and WANG Xu¹

1 School of Computer and Information Engineering, Hubei University, Wuhan 430062, China

2 Hubei Engineering Research Center for Educational Informationization, Wuhan 430062, China

Abstract In order to prevent or avoid the occurrence of tax evasion and taxation caused by non-invoicing of mineral resources such as coal, sand and gravel, it is an effective way to use the deep convolutional neural network to automatically identify empty vehicles. Based on the AlexNet model, this paper proposes 5 kinds of improvement ideas for the difference of empty car and heavy vehicle images, and finally obtains a structure of 6-layer convolutional neural network based on maxout+dropout. The test results of the picture of the 34220 empty cars and loaded cars show that the model has achieved good results in terms of accuracy, sensitivity, specificity and precision. In addition, the model is highly robust and can successfully identify a large number of empty car images with different angles and different scenes.

Keywords Empty car and loaded car identification, Deep learning, CNN, AlexNet, maxout

煤炭行业是国家限定开采的能源支柱产业,煤炭素有“工业粮食”的美誉。在煤炭企业出入口抓拍运载车辆的装载情况,确保一车一票开具矿产品销售确认单,杜绝不开票行为,这是税务部门对煤炭企业的重要监管手段。目前,税务部门在使用视频监控管控煤炭行业的纳税行为主要以人工排查为主,即安排人员对煤矿企业的出入口进行实时监管。然而,煤炭企业的监控图像具有数量巨大、时间跨度较大以及数据质量参差不齐的特点。以贵州省毕节市为例,全市煤炭企业一天的监控图像多达上万张,拍摄时间跨度一般为早上八点到零点,且需要人为判别的图像不到十分之一。如果单靠人工排查出入口的车辆装载情况,既费时又费力。对于空车重车识别问题,2018 年 Chen^[1]设计出一种基于翻车机仪器的空车重车识别方法,虽然可用于实际应用,但其原理是通过人为翻转翻车机来进行人工排查,效率低下且人工成本高昂,因此并不适用于税务部门利用视频图像监管煤炭行业的纳税行为这种情况。

随着大数据和深度学习技术的发展,尤其是卷积神经网络(Convolutional Neural Networks, CNN)在图像分类上体现出来的优越性,使得越来越多的人将 CNN 应用到图像识别

领域。AlexNet^[2]是 Alex 等于 2012 年设计的一个 8 层神经网络模型,一经公开即被广泛应用。Zhong 等将 AlexNet 应用于左右眼识别^[3],获得了 99% 的准确率;Lv 等对 AlexNet 模型进行了改进并将其应用于阿尔茨海默病的早期诊断^[4],得到了较好的诊断效果;Chen 等^[5]将 AlexNet 改进后用于乳腺疾病的诊断,获得了 98% 的识别率;Li 等^[6]采用梯度下降法对 AlexNet 进行调整,在小样本睡眠漂浮物的识别上也获得了较好的效果。

基于 AlexNet 的 CNN 在图像识别上具有准确性高且训练速度快的特点,本文在空车重车识别问题上引入了 AlexNet 模型,让模型在大量的训练样本中自动学习图像特征,并进行图像分类。通过改变图片的亮度、对比度,以及将图片进行旋转,扩大了图片的数据量,增强了模型的鲁棒性;针对图片数据在 CNN 中容易出现过拟合的问题,改进 AlexNet 原有的模型结构,将 maxout^[7]和 dropout^[8]结合使用,并采用五重交叉验证的训练方法,对比原模型分类准确率,减轻了模型的过拟合问题,使模型具有更好的鲁棒性;在分类器的选择上,对比了 softmax 和 sigmoid 在空车重车识别上不同的效果,通过验证模型的准确率、敏感度、特异性、错误率等指标,

基金项目:湖北省自然科学基金(2019CFB757)

This work was supported by the Natural Science Foundation of Hubei Province, China(2019CFB757).

通信作者:汪炆杰(wyj_zizy0310@163.com)

表明 sigmoid 更适合对空车和重车进行分类。

1 卷积神经网络

1.1 深度学习

深度学习是构建一个多层级的网络,每一层的特征由上一级的特征进行提取组合,以此来发现数据的分布式特征表示。深度学习的特征由机器提取,避免了人工提取特征的繁琐以及时间的损耗,但是每一层的特征数并没有相应的公式,只能由经验来设定。特征数过多,模型容易出现过拟合;特征数太少,则模型对数据的拟合度不高,从而达不到预期的效果。深度学习多层结构如图 1 所示。

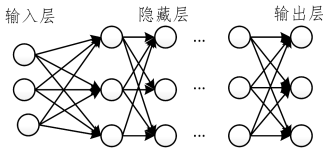


图 1 深度学习多层结构

Fig. 1 Multi-layered structure for deep learning

深度学习主要是模仿人脑进行分析和学习的神经网络,模仿人脑学习数据的模式来解释图像、数据、语音等。深度学习是基于大量数据来实现的。一般来说,数据量越大越好。但是在基于深度学习的图像识别领域,数据集的获取会因场景、隐私性等条件的限制导致深度学习在这一领域研究的很少,但是随着深度学习在计算机视觉上体现出来的优势,深度学习在图像识别领域的应用已成为不可阻挡的趋势。

1.2 卷积神经网络

卷积神经网络^[9]是深度学习的神经网络模型,其原理是利用设定好的卷积核以及池化运算对图片交替操作,从而得出更加复杂的特征^[10],卷积神经网络采用前向传播输出特征值,再通过后向传播调整权重和偏置,层与层之间采用局部连接(即感受野^[11])。典型的卷积神经网络模型由卷积层、池化层和全连接层组成。卷积神经网络的学习过程如下:

(1)将原始图片调整为一定尺寸并作为模型的输入,令 $x = \{x_0, x_1, \dots, x_{N-1}, x_N\}$,其中 N 为输入图像的数量。

(2)卷积运算

卷积是使用两个大小不同的矩阵进行的数学运算,卷积层主要是实现对图像的过滤以及强化效果,设在卷积网络的第 L 层,将前一层的特征图与卷积核进行卷积操作,并通过激活函数得出该层的特征图 y_i^L ,如式(1)所示:

$$y_i^L = f\left(\sum_{i \in x^{L-1}} x_i^{L-1} \otimes k_{i,i}^L + b_i^L\right) \quad (1)$$

其中, f 代表激活函数, \otimes 代表卷积操作, i 代表层数, b_i^L 代表偏置, $k_{i,i}^L$ 代表卷积核, x_i^{L-1} 表示第 i 层输入特征的集合。

通过卷积获得特征之后,针对特征图像具有的一种“静态性”属性,需要通过下采样来减少卷积层的特征维数,对不同位置的特征进行聚合统计,如式(2)所示:

$$y_i^L = f(\text{down}(y_i^{L-1}) * \omega_i^L + b_i^L) \quad (2)$$

其中, $\text{down}(\dots)$ 为函数,常用的有最大池化和平均池化, ω_i^L 为权重。

全连接层一般连接在池化层后,会有一个或多个全连接层,全连接层将池化层输出的所有特征串联,变成矢量,输出图像特征,如式(3)所示:

$$F = f(\omega_i * x^{L-1} + b_i) \quad (3)$$

卷积运算^[12]主要有两种方式:“稀疏矩阵”和“参数共

享”。稀疏矩阵是指卷积核的大小远远小于输入图像的大小;参数共享是指在卷积神经网络中,卷积核的每一个元素都被用在输入图像的每一个像素点上,而在过程中只需学习一个参数集合就能把这个参数应用到所有的图片元素中。这样做的好处是:例如输入一个 $227 * 227$ 的重车图片,如图 2 所示,如果使用传统的神经网络,将产生 227^4 个连接;如果使用卷积神经网络,且卷积核大小为 $5 * 5$,那么参数量只有 $227 * 227 * 5 * 5$ 个连接,而且如果卷积神经网络利用同一个卷积核对图像进行操作,不论隐层神经元的尺寸如何,这之间的共享参数只有 25 个,这是其他神经网络所不能比拟的。

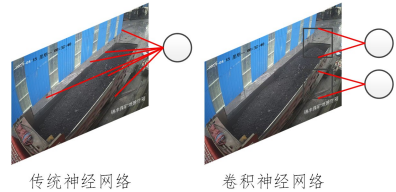


图 2 卷积神经网络与传统神经网络的对比

Fig. 2 Comparison between CNN and traditional neural network

2 数据预处理及实验环境

2.1 数据获取

本文收集了贵州省毕节市大方县不同煤矿企业的监控图像,所有图像资料均由国家税务总局大方县税务局提供。为应对不同出入口的识别问题,从大方县小屯煤矿和大方县瑞丰煤矿两家煤矿共获取图片 3 260 张,图像均由海康 DS-2CD3T36DWD-i5 网络摄像机拍摄,图片尺寸为 $2048 * 1536$,所有照片均来自海康网络摄像头自带的触发事件抓拍,保证了数据集的质量。所有的空车重车类别的图片都经过了人工识别,保证了数据标签的正确性。由于特征差异,在对图片进行卷积操作时,不同的特征对模型有着不同的影响。因此,对图像进行标准化处理会使得从图片中提取出来的不同的特征具有相同的尺寸。这样,在使用反向传播算法时,不同特征对参数的影响是一致的。同时,进行标准化处理之后可以加快模型的收敛速度,在一定程度上也可以提高模型的精度。

2.2 数据增强

本文的数据集拍摄时间均为上午八点半到下午五点半。然而,根据实际应用情况以及数据库的信息比对发现,有部分煤炭企业的开票时间会持续到凌晨。所以最初的图片并不能完全满足实际需要,应该对数据集的亮度和对比度进行调整,以满足实际需要。为此,本文按原图的 50%, 90%, 130% 的亮度和对比度,通过自动算法,得到不同亮度和对比度的图片,效果如图 3 所示。通过该方法将原始数据集扩大了 9 倍,并将 90% 的图片作为训练集,其余 10% 作为测试集。



(a)调整前

(b)调整后

图 3 调整了亮度和对比度的图像对比

Fig. 3 Image comparison after adjusting brightness and contrast

2.3 实验环境

本文采用的深度学习框架为 TensorFlow,训练过程采用 GPU 的训练方式,实验配置为 Xeon(R) Silver 4114 CPU,两块 NIDIA Telsa M60 显卡,使用 AdamOptimization 优化函数进行权值优化。

3 卷积神经网络的设计与实现

3.1 卷积神经网络 AlexNet

研究表明^[3],当模型为二分类模型时,AlexNet 较其他深层模型表现更好,因此选用 AlexNet 模型。模型的结构如图 4 所示。

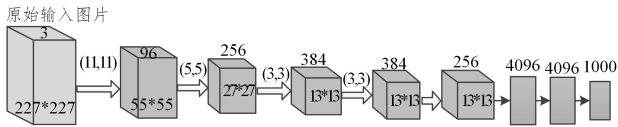


图 4 AlexNet 网络结构图

Fig. 4 Diagram of AlexNet network structure

模型输入为 227 * 227 * 3,227 * 227 为输入图片的尺寸,3 代表彩色三通道 RGB 图片。从模型分解来看,AlexNet 前两个完整的卷积层包括一层卷积、一层 Rectified Linear Units、一层 MaxPooling,一层 normalization。就第一个完整的卷积层来说,首先对 227 * 227 的图片进行卷积操作,其中卷积核(kernel size)大小为 11 * 11,步长(stride)为 4,卷积后的图片尺寸变换为 55 * 55 * 96(55 = ((227 - 11) / 4) + 1),其中 96 为自行设置的卷积核个数,经过 Relu 激活函数,图片尺寸大小不变。再做最大池化操作,其中 Kerner size 为(3 * 3),步长(stride)为 2,最终得到的图片大小为 27 * 27 * 96(27 = ((55 - 3) / 2) + 1),随后即为对数据的归一化处理,公式如式(4)所示,local_size(即 N)设置为 5,Alpha(即缩放因子 α)设置为 0.0001,beta(指数项 β)为 0.75,得到下一层的输入为 27 * 27 * 96。

$$b = a / ((l + \beta / N) \sum (a)^2)^\beta \quad (4)$$

第二层的过程与第一层一样,不同的是卷积核大小、步长以及特征数等参数的设定,第三层和第四层只做了卷积和 Relu。第五层与第一层类似,只是没有做 norm。在得到第五层的结果后,将其输入到一个传统的神经网络中,经过 3 个全连接层,应用 softmax 得到分类值^[13],softmax 的公式如式(5)所示:

$$p(y = l | x) = \frac{e^{x^T w_l}}{\sum_{k=1}^K e^{x^T w_k}} \quad (5)$$

本文将 AlexNet 原始模型在本文的应用算法暂称为算法 1,其算法流程图如图 5 所示。

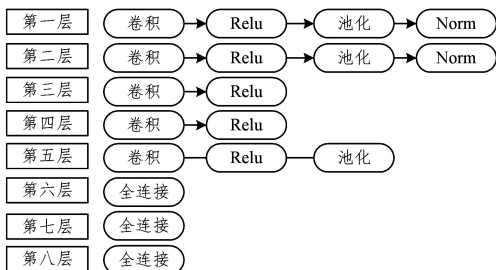


图 5 算法 1 模型结构图

Fig. 5 Model structure of algorithm 1

其中算法 1 的损失函数趋势图如图 6 所示。

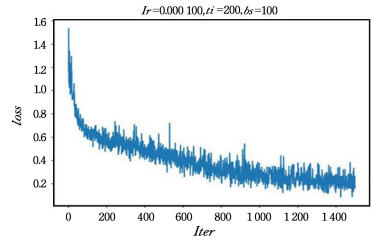


图 6 算法 1 的损失函数趋势图

Fig. 6 Trend graph of loss function of algorithm

本文在研究算法准确率上采用测量模型的准确度、灵敏度和特异性等几个关键评判标准作为本文评估模型的依据。算法 1 的模型评估数据如表 1 所列。

表 1 算法 1 的模型评估数据

Table 1 Model evaluation data of algorithm 1 (单位: %)

Accuracy	Sensitive	Specificity	Precision
78.81	79.88	77.75	78.21

3.2 AlexNet 结构优化

根据图 5 以及表 1 的结果显示,算法 1 的准确度并不高,可以得出算法 1 出现了过拟合现象的结论。对于这种情况,文献[3]中提出将一种简化的 AlexNet 模型用于二分类的问题,并取得了较好的效果。以此为启发,本文设计了一个包含 4 个卷积层和 2 个池化层的模型,每一个卷积层后面都接了一个最大池化层和一个 dropout 层,前两个卷积层中,卷积核大小为 5 * 5,后两个卷积核大小为 3 * 3,最大池化层的核为 3 * 3,第一个卷积层中步长为 3,第二个卷积层中步长为 2。其他卷积层中步长均为 1,池化层中步长均为 2。考虑到空车重车识别为一个二分类问题,以及算法 1 在模型中的表现,本文将全连接层的特征大大减小。为了探究最适合空车重车识别模型的特征数,本文从探究算法 1 特征数的角度出发,从原始特征数开始做公差为 2 的批次处理,最终得到了图 7 所示的模型结构,并将其暂命名为算法 2。

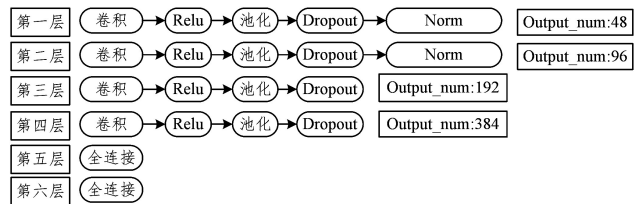


图 7 算法 2 的最终模型结构

Fig. 7 Model structure of algorithm 2

在得出如上结构后,本文对模型的输入做了测试,分别将图片的输入尺寸在算法 1 的基础上做公差为 128 的批次处理,得出在空车重车识别的模型中,640 * 640 的输入尺寸较好的结论,其模型评估数据如表 2 所列。

表 2 算法 2 的模型评估数据

Table 2 Model evaluation data of algorithm 2 (单位: %)

Accuracy	Sensitive	Specificity	Precision
81.19	89.9	72.6	76.57

3.3 maxout + dropout

由表 2 的结果可以看出,算法 2 与算法 1 相比,模型在准确度、灵敏度、特异性等方面均有了很大的提高,对过拟合的效果也有所改善。由文献[6]可以发现,dropout 作为一种抑

制过拟合的方法,通过忽略隐藏层的神经元可达到稀疏化的效果,从而抑制过拟合。为了使 dropout 抑制过拟合的效果最大化,由文献[14]可以发现,relu, sigmoid 等只能拟合二维函数,但是 maxout 作为一种分段函数可以拟合任意维度的函数,在深度学习中有较好的拟合效果^[15]。maxout 作为一种前向传播结构,从 k 个隐层的延伸节点选取一个最大值作为输出,由于 K 个节点是线性的,故 maxout 函数在每一处都是线性的。唯一的不足之处是,由于给隐层节点扩大了 K 个延伸节点,因此 maxout 函数不是稀疏的,但是 dropout 可以使输出稀疏化,从而代替 relu 函数。依照上面的思路,我们对模型的激活函数作了调整,调整后的模型暂称为算法 3,算法 3 的结构如图 8 所示,Maxout 的公式如式(6)所示。

$$h_i(x) = \max_{j \in [1, k]} z_{ij} \quad (6)$$

其中, x 为输入, $z_{ij} = x^T w_{...ij} + b_{ij}$, W 为权重, b 为偏置。

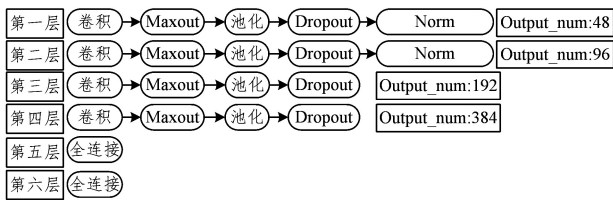


图 8 算法 3 的模型结构

Fig. 8 Model structure of algorithm 3

模型的评估数据如表 3 所列。

表 3 算法 3 的模型评估数据

Table 3 Model evaluation data of algorithm 3 (单位: %)

Accuracy	Sensitive	Specificity	Precision
83.25	86.15	80.4	81.44

3.4 maxout + dropout 对不同层级的卷积神经网络的影响

对比表 2 的实验结果以及文献[16]的结果得出, AlexNet 在第三层和第四层的提取效果较好。为了更好地提升模型的精度,本文以算法 3 为基础,以文献[4]作为优化思路,更改了层级结构,去掉了三、四层的池化层,增加了一个卷积层和池化层,以及一个全连接层,将其命名为算法 4,其结构图如图 9 所示。

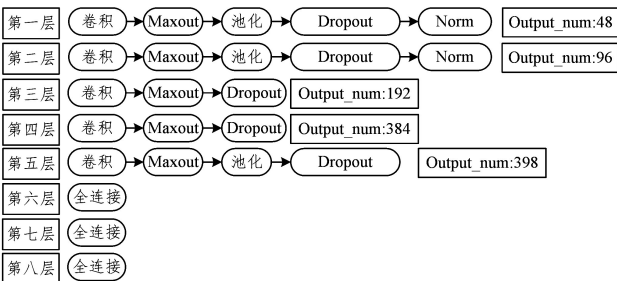


图 9 算法 4 的模型结构

Fig. 9 Model structure of algorithm 4

模型的评估数据如表 4 所列。

表 4 算法 4 的模型评估数据

Table 4 Model evaluation data of algorithm 4 (单位: %)

Accuracy	Sensitive	Specificity	Precision
77.15	91.1	63.2	59.0

3.5 模型分类层的函数比较

Sigmoid 是一种在 logisti 回归中经常使用的函数,常用

于二分类问题上,其公式如式(7)所示:

$$s(x) = \frac{1}{1 + e^{-x}} \quad (7)$$

通过前面 4 种算法的优化路线,本文对模型的结构、激活函数的搭配,以及参数的调整做出了优化。为了最大程度地优化模型,本文以前面 4 种改进算法和优化思路为基础,比较了 softmax 和 sigmoid 两种分类器,并将其在模型中的效果做了对比,其中模型的结构与前述算法中效果最好的算法 3 一致,并将其命名为算法 5,算法 5 的评估数据如表 5 所列。

表 5 算法 5 的模型评估数据

Table 5 Model evaluation data of algorithm 5 (单位: %)

Accuracy	Sensitive	Specificity	Precision
88.38	95.75	81	83.4

3.6 模型横向对比

VGG16^[17] 是牛津大学于 2014 年提出的模型,由于其简洁性和实用性,一举获得了 2014 年 ILSVRC 比赛的冠军,网络的输入采用 224×224 的 RGB 图像,卷积核大小为 3×3 和 1×1 两种,卷积层后跟有一个池化层,采用最大池化,池化大小统一为 3×3 ,步长均为 2,全连接层结构一致。并且激活函数采用 relu,每层参数与 AlexNet 一致。本文按照原始 VGG16 模型结构对空车重车识别做了测试,并将 VGG16 在本文的应用算法暂称为算法 6,其模型结构图如图 10 所示,模型评估数据如表 6 所列。

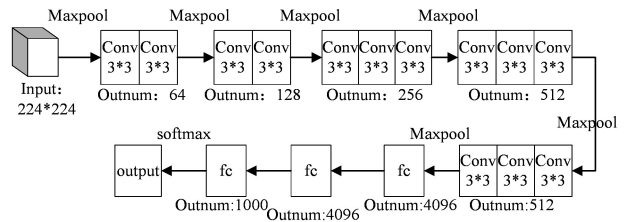


图 10 算法 6 模型结构图

Fig. 10 Model structure of algorithm 6

表 6 算法 6 的模型评估数据

Table 6 Model evaluation data of algorithm 6 (单位: %)

Accuracy	Sensitive	Specificity	Precision
51.19	100	2.38	50.60

4 实验结论

根据上述评估,本文研究出 6 个算法模型,其中为了评估模型的性能^[18],本文从模型的准确度、灵敏度、特效度和精度 4 个方面综合考虑模型的性能,其中判断模型的准确度的公式如式(8)所示:

$$accuracy = \frac{TP + TN}{P + N} \quad (8)$$

判断模型的敏感度,即模型的正例正确分类的准确率,如式(9)所示:

$$sensitive = \frac{TP}{P} \quad (9)$$

判断模型的特异性,即模型的负例正确分类的准确率,如式(10)所示:

$$specificity = \frac{TN}{N} \quad (10)$$

判断模型的精度如式(11)所示:

$$precision = \frac{TP}{TP + FP} \quad (11)$$

其中, TP 为真阳例, 指被分类器正确分类的正源组; TN 为真阴性, 指被分类器正确分类的负元组; FP 为假阳性, 指被错误标记为正元组的负元组; FN 为假阴性, 指被错误标记为负元组的正元组。考虑到煤炭运载车辆的识别是一个二分类问题, 本文选择空车作为正元组。

根据以上权衡模型分类性能的方法, 将 6 种算法的模型评估情况列入图 11 中。

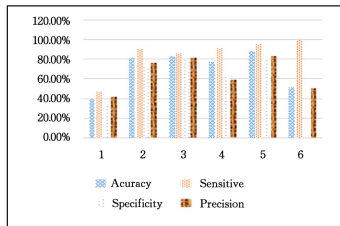


图 11 实验结果对比

Fig. 11 Comparison of experimental results

从图 10 可以看出, 算法 5 除了敏感度不如算法 6 外, 其他各项指标均优于其他模型, 达到了 95.75% 的敏感度以及 88.38% 准确率。对于算法 6 来说, 虽然其敏感度达到了 100%, 但是从实验结果看, 它是一个过拟合的模型, 将所有的模型均拟合成了空车, 从这一点来看, VGG16 作为一个比前几种算法更深的模型, 并不适用于空车重车识别场景。

结束语 本文在实地收集了大量空车重车图像的情况下, 利用卷积神经网络提出并验证了一个可以自动识别空车重车的新方法。该方法首次将深度学习应用到空车重车识别领域, 并获得了具有 95.75% 敏感度和 88.38% 准确率的模型。与以往的空车重车识别情况相比, 该方法具有更好的操作性, 只需将图片传入模型就可以得到结果, 并且本文在参考众多文献的基础上, 将模型做了最大的简化, 针对模型的过拟合问题也得到了有效的处理。在模型的测试方面, 除了一些常用的参数以外, 还测试了 AlexNet 和 VGG16, VGG16 层数更深, 但是在本文的表现却不如新简化的算法 5, 其原因可能是针对特征明显的二分类问题, 简化的模型反而能起到意想不到的效果。其次, 本文采用 maxout+dropout 结构, 更好地抑制了模型的过拟合, 且使用 maxout 这种激活函数也有更好的拟合效果, 使模型不会过拟合也不会欠拟合。空车重车识别是一个二分类问题, 本文在 TensorFlow 框架的基础上, 提出了一种 6 层的神经网络模型 (即算法 5)。与经典的 AlexNet 相比, 本文提出的算法 5 模型, 在减少计算量同时还大大提高了精度。除以之外, 本文还发现 maxout+dropout 这种层级结构适合用于空车重车识别, 与 softmax 分类器相比, sigmoid 也更适合作为空车重车模型的分类器。本文的研究也表明, 简单的卷积神经网络在二分类问题上也有不错的表现。接下来我们准备将算法 5 模型封装成 API 接口, 用在税务稽查的系统上, 同时从更多的角度考虑, 得出鲁棒性、准确率更高的空车重车识别模型。

参考文献

- [1] CHEN C. Design of Empty Car Recognition System for Undetected Tipper[J]. Lifting and transporting machinery, 2018, 11: 107-109.
- [2] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet Classification with Deep Convolutional Neural Networks[C]// International Conference on Neural Information Processing Sys-

tems. Curran Associates Inc, 2012; 1097-1105.

- [3] ZHONG Z Q, YUAN J, TANG X Y. Left-vs-Right Eye Discrimination Based on Convolution Neural Network[J]. Journal of Computer Research and Development, 2018, 55(8): 1667-1673.
- [4] LV H M, ZHAO D, CHI X B. Deep Learning for Early Diagnosis of Alzheimer's Disease Based on Intensive AlexNet[J]. Computer Science, 2017, 44(S1): 50-60.
- [5] CHEN S W, LIU Y J, LIU D, et al. AlexNet Model and Adaptive Contrast Enhancement Based Ultrasound Imaging Classification[J]. Computer Science, 2019, 46(S1): 146-152.
- [6] LI N, WANG Y X, XU S, et al. Recognition of Floating Objects on Water Surface With Small Sample Based on AlexNet[J]. Computer Applications and Software, 2019, 36(2): 245-251.
- [7] GOODFELLOW I J, WARDE-FARLEY D, MIRZA M, et al. Maxout networks[EB/OL]. [2017-0912]. <http://www-etud.iro.umontreal.ca/~goodfeli/maxout.pdf>.
- [8] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. Journal of Machine Learning Research, 2014, 15(1): 1929-1958.
- [9] 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.
- [10] SHUIWANG J, MING Y, KAI Y. 3D Convolutional Neural Networks for Human Action Recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2013, 35(1): 221-231.
- [11] HUBEL D H, WIESEL T N. Receptive field, binocular interaction and functional architecture in the cat's visual cortex[J]. The Journal of Physiology, 1962, 160(1): 106-154.
- [12] WANG X H. Tensorlow deep learning application practice[M]. Beijing: Tsinghua University Press, 2018.
- [13] HINTON G, DENG L, YU D, et al. Deep Neural Networks for Acoustic Modeling in Speech Recognition[J]. IEEE Signal Processing Magazine, 2012, 29(6): 82-97.
- [14] ZHAO H Z, LIU F X, LI L Y, et al. Improving deep convolutional neural networks with mixed maxout units[J]. Journal of Communications, 2017, 38(7): 107-114.
- [15] BAI C, HUANG L, CHEN J N, et al. Optimization of Deep Convolutional Neural Network for Large Scale Image Classification. Journal of Software[J]. 2018, 29(4): 1029-1038.
- [16] LIU F, SHEN C. Learning Deep Convolutional Features for MRI Based Alzheimer's Disease Classification[OL]. <http://arxiv-web.arxiv.org/pdf/1404.3366v1>.
- [17] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large Scale Image Recognition. [OL]. <http://arxiv.org/abs/1409.1556>. 2014.
- [18] FAN M, MENG X F. Data mining concepts and techniques [M]. Beijing: China Machine Press. 2012.



MA Chuan-xiang, born in 1971, professor, postgraduate supervisor. Her research interests include data mining and machine learning.



WANG Yang-jie, born in 1994, master. His research interests include deep learning and image recognition.