



图像识别技术在电信运维质检上的应用

范博, 邱芸, 沈雷

(中国电信股份有限公司上海研究院, 上海 200122)

摘 要: 随着电信行业的发展, 持续增长的宽带业务需要更高的人工运维成本。在电信运维的装维质检中, 需要人工识别施工现场图片以评估装维质量。传统的装维人工质检存在检测准确率低且人力成本高的问题。近年来基于深度学习的图像识别技术不断发展, 应用领域不断拓宽。若采用图像识别技术, 则可显著节约人力成本并提高准确率。将多种图像识别技术应用在电信运维业务中的装维场景, 并分析比较各种图像识别技术的性能与准确率。基于分析, 进一步提出了一种融合的图像识别模型, 应用于电信运维的质检流程, 显著提高了运维图片识别准确率和质检效率。

关键词: 人工智能; 电信运维质检; 图像识别; 融合模型

中图分类号: TN929.5

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2019086

Application of image detection techniques in telecom operation and maintenance system

FAN Bo, QIU Yun, SHEN Lei

Shanghai Research Institute of China Telecom Co., Ltd., Shanghai 200122, China

Abstract: With the development of telecom industry, growing web service operations need more people to maintain the quality of the service. In the branch of telecom operation service, installation and maintenance service, the service quality is evaluated by human through assessing picture of construction site. The traditional way of evaluating quality of operation service brings low accuracy and high labor cost. In recent years, image recognition technology developed well with the deep learning raised, the related application field were broaden. If image recognition techniques were applied in telecom operation and maintenance system, the labor cost would be saved and the recognition accuracy would be improved. Several image recognition techniques were applied in installation quality assessment of telecom operation and maintenance system, and the performance and accuracy were analyzed. Based on analysis, a combining image recognition model was further proposed. The image recognition accuracy and efficiency of quality assessment was improved.

Key words: AI, quality assessment of telecom operation and maintenance, image recognition, combining model

1 引言

随着电信宽带业务的快速发展,2017年中国电信有线宽带用户数已经达1.34亿户。对于庞大的用户基数,运维需要越来越多的人力维持服务质量。例如光纤安装,装维人员对施工现场所拍照片会进入数据库,作为后期装维质检评估的重要考据。而装维图片这类非结构化数据无法通过在数据库中制定一系列规则来检查,只能通过人工检测。但是日益增长的电信宽带业务给装维质量评估带来很大难度,人工评估的方式不再适用于爆发性增长的装维图片数据量。

图像识别技术可追溯至1960年,最早的图像识别技术基于模式识别,被应用在办公自动化的领域^[1]。随着深度学习的发展,图像识别技术依托于高质量的大规模数据集与不同的层次结构的神经网络,不再依赖于传统的机器学习技术,在学术界和工业界都取得了令人瞩目的成绩,刷新了人们对图像识别技术的认知。例如公开的千万级图像数据集ImageNet、SUN397^[2]等,极大地推动了图像识别技术的发展。2012年Krizhevsky等人^[3]提出了Alexnet,在ImageNet竞赛上取得了突破性的成绩,此后基于卷积神经网络(convolutional neural network, CNN)的更深层的网络不断被提出,在图像识别和目标分割等领域有着显著效果。

电信运维过程中会产生大量的非结构化数据,例如工单、图片数据等。传统的机器学习难以对非结构化数据中的特征进行提取分类,而深度学习由于其特殊的网络结构,被广泛应用于非结构化数据中的特征提取。因此本文针对电信装维业务中所产生的大量施工现场图片数据,为其后期评估工作提供解决方案,应用多种基于CNN的图像识别技术,并比较各类技术的优势劣势。此外根据装维评估业务的要求,基于多种图像识别技术,本文提出一种融合模型,综合了各类

技术的优势,整体提高了准确率。最终展现模型的测试结果及现网运行情况。

2 运维质检图片识别关键技术

对于运维质检图片识别任务来说,首先改变传统运维质检流程,加入深度学习检测图片的步骤替代人工质检。其次在图片检测模型建立中,主要分为图片预处理、多模型训练和融合模型优化3个阶段。对于图片预处理,应该按不同类型的图片应用不同的处理方式,例如对于室内家庭网关图片,本文根据数据集的特点采用直方图均衡化的方法,对光照影响进行矫正。模型训练阶段应用图像分类与目标分割的技术,并结合两种技术优势提出一种融合模型。迁移学习阶段主要是通过后续数据集对模型进行优化。

2.1 基于深度学习的装维质检流程

以深度学习模型替代传统人工运维质检流程,架构变化与测试效果如图1所示。

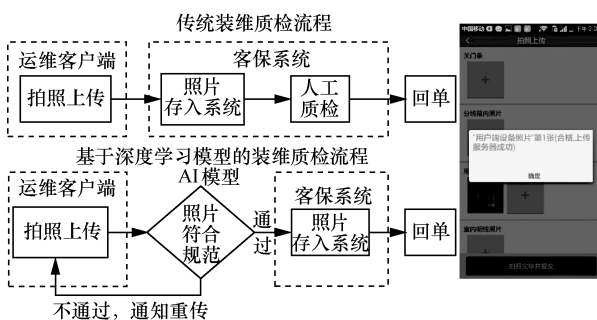


图1 装维质检流程

由图1所示,深度学习模型能够直接实时对运维人员上传的施工照片进行质检,若不合格则会通知重传。这样可以保证进入客保系统图片的质量,同时实时处理质检问题,提高运维效率,节约人工质检所带来的人力成本。

2.2 运维质检图片来源

本文所使用的实验数据来源于电信装维业务中施工人员在设备安装或维修结束后通过客户端上传的施工拍照凭证。具体可以分为家庭网关、



电视盒子与室外分线箱 3 类。在实际情况中,待质检图片会出现设备缺失、设备不清晰、设备错误等多种情况。图像数据如图 2 所示。

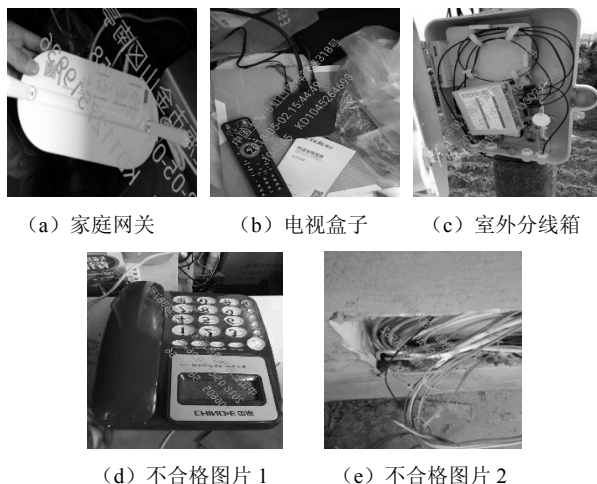


图 2 运维图片数据集

电信宽带业务量巨大,每天约产生 1 万张上述图片集,其中存在大量如图 2 (d) 和图 2 (e) 所示的图片。运维质检要求施工项目必须与图片中待测物一致,例如家庭网关安装项目中,图片中就必须出现家庭网关,若出现图 2 (d) 或图 2 (e) 就视为不合格。

2.3 基于直方图均衡化的图片预处理

首先应该根据图片的不同种类,采用不同的图片预处理方式。例如室内家庭网关,某些型号的网关颜色与背景相似,神经网络难以分辨待测网关与背景;家庭网关表面光滑,容易反射强光造成部分区域亮度过高。直方图均衡化可以对图片进行非线性的变化,重新分配像素值使输出图像直方图分布均匀。对于 $M \times N$ 大小的图像,其变换函数如式 (1)。

$$S_k = \sum_{j=0}^k \frac{n_j}{M \times N}, k=0,1,2,\dots,L-1 \quad (1)$$

其中, n_k 是当前灰度级像素的个数, L 是图像中灰度级的总数。

通过直方图均衡化,原始图像的灰度动态范围得到增加,提高对比度,从而利于神经网络的

特征提取。如图 3 中网关与背景均为黑色的装维图片,应用直方图均衡化处理。

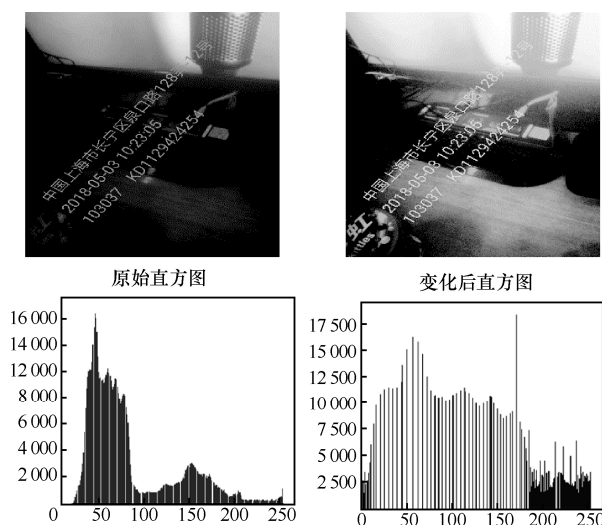


图 3 基于直方图均衡化的装维图片预处理

原始图片灰度分布不均匀可能导致 CNN 在提取特征时难以分辨出待测网关与背景。通过直方图均衡化方法,图片整体对比度明显提升,待测网关边缘特征更加明显。由于除家庭网关外的运维质检图片主要为室外分线箱,室外照明条件较好,光线均匀,故室外图片不采用直方图均衡化方式,也可直接进入模型训练。

2.4 运维质检图片识别建模

图像识别技术可以分为 4 个等级:物体检测、物体定位、场景检测和场景解析^[1]。物体检测是最低级别的需求,即检测图像中是否包含待检测物,例如针对手写数字识别问题的 5 层神经网络 LeNet-5,本质上是一个多分类问题;物体定位是指不仅要识别图像中是否包含待检测物,还要具体定位出待检测物的位置;场景检测在物体定位的基础上继续扩展,即定位出图像中包括待检测物在内的所有物体的位置,例如 COCO 2015 竞赛中的 Faster R-CNN 和 RPN 框架;场景解析是指理解图片的语义。这 4 种级别的图像技术是以基于 CNN 的不同深度和不同层次结构的深度学习网络实现的。

本文根据质检数据集特点,主要应用基于

CNN 的图像分类与目标分割的图像识别技术。

CNN 浅层通过对提取的图片普遍性特征,例如图片的边缘、纹理等,进行组合进入高层,成为有针对性的特征。例如在人脸识别中,低层提取人脸的点、线和边缘等特征,由于 CNN 局部感知和共享权值的特点,仿射变化不会影响基础特征的提取,在高层这些基础特征就可以组合成鼻子、眼睛等有针对性的特征,形成多个特征图。

2.4.1 基于 CNN 的图像分类技术

图像分类的 CNN 网络本质上是一个多分类器。从 2012 年被提出的 AlexNet 发展到 2014 年在 ImageNet 大规模视觉识别挑战(ILSVRC 2014)排名前 7 的 VGGNet 和 GooleNet^[4],除了网络在不断加深以外,新的优化技术如 GooleNet 中 Inception 模块将深层网络的训练变得可行,并且更加细致化了特征提取的过程。

AlexNet 与 VGGNet 结构比较相似,两个网络对于特征提取的过程都是顺序进行,即浅层与高层是一一对应的关系。但是相较于 AlexNet, VGGNet 添加了更多的卷积层与池化层,连续卷积层更多,计算量更大。以 VGGNet 中的 VGG-16 为例,图 4 展示了 AlexNet 与 VGG-16 的异同。

AlexNet	VGG-16
Input (227×227 RGB image)	Input (224×224 RGB image)
conv-96 (11×11)	conv-64 (3×3)
	conv-64
	maxpool
conv-256 (5×5)	conv-128 (3×3)
	conv-128
	conv-128
	maxpool
conv-384 (3×3)	conv-256 (3×3)
conv-384	conv-256
conv-256	conv-256
	maxpool
	conv-512 (3×3)
	conv-512
	conv-512
	maxpool
	conv-512 (3×3)
	conv-512
	conv-512
	maxpool
	FC
	FC
	FC
	softmax

图 4 AlexNet 与 VGG-16 网络结构对比

VGG-16 通过效仿 AlexNet 的顺序结构,加入额外的连续卷积层,将 8 层网络扩展到 16 层。

VGGNet 与 AlexNet 主要不同在于 VGG-16 运用更多更小的卷积核提取特征,例如在图 1 中, VGG-16 采用 3×3 的卷积核代替了 AlexNet 中 11×11 和 5×5 的卷积核。通过这种方法在特征提取的过程中能保存更多的基础特征信息,但是也会带来更大的计算量。

GooleNet 在 VGGNet 基础上新加入了 Inception 模块,从而改变了整个网络结构,使得浅层到高层不再是一一对应的关系。在 Inception 模块中,卷积核本身变成了一个单独的子网络,如图 5 所示。

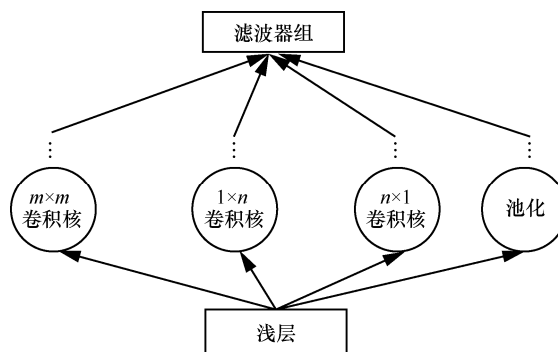


图 5 Inception 模块结构

在特征提取过程中, Inception 模块对浅层输入的特征图进行不同的卷积或池化运算,在滤波器组组合所得的多个结果,形成更深的特征图。不同于 VGGNet 在每层使用固定大小的卷积核提取图信息, Inception 模块通过组合不同卷积核所能提取的不同图信息,更好地保留了原始图像信息。但是使用 Inception 模块会使得网络变得更深,容易造成过拟合与梯度消失等问题。

运维质检图片识别本质上是一个单类识别问题,因此本文在实验中采用 3 种经典的图像分类网络,即 AlexNet、VGG-16 与 InceptionV3。一般来说,图像分类网络卷积层的值采用正态分布初始化和 Imagenet 权值初始化两种方式,本文针对两种初始化方式进行对比试验。

(1)采用随机初始化权值的训练过程如图 6 所示。

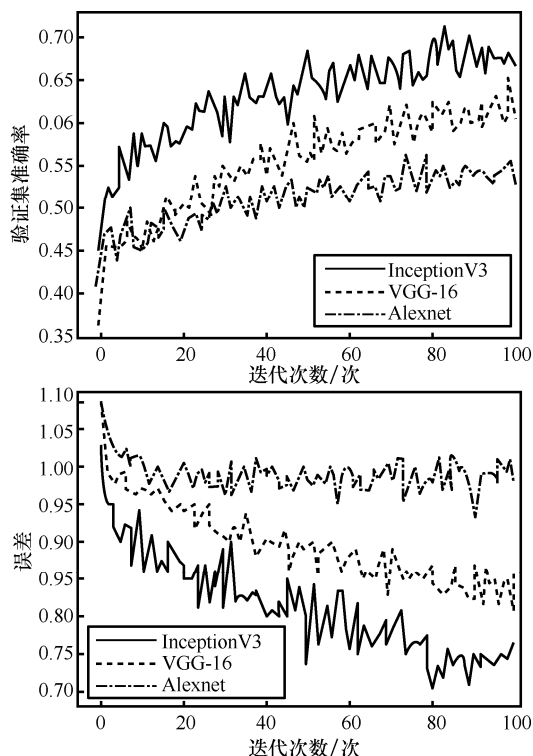


图6 随机初始化训练过程对比

图6中均使用随机权值初始化,展示了训练过程中验证集上准确度与损失函数的收敛情况。在100个迭代期内,InceptionV3的准确率,VGG-16其次,AlexNet最差。根据损失函数收敛情况,相较于InceptionV3与VGG-16,8层的AlexNet效果最差,在20个迭代期后就只在小范围内波动。然而在随机初始化权重中,效果最好的InceptionV3达到收敛时,验证集准确率也不超过70%。

(2) 采用基于Imagenet数据集的权值网络初始化,在100个迭代期内研究不同的权重初始化方法对训练过程的影响,如图7所示。

图7展示了预加载Imagenet权值初始化的两种网络训练过程,可以看到两种模型的验证集准确率大大超过了采用随机初始化权重的方法。这是因为对权重采用合适的初始值避免了深层网络中梯度消失的问题,使得验证集准确率有显著提升,其中收敛时InceptionV3准确率更是超过了90%,相较于随机初始化方法提升了30%。从损失函数收敛情况来看,InceptionV3作为更深层的

网络,其收敛效果仍然好于VGG-16。

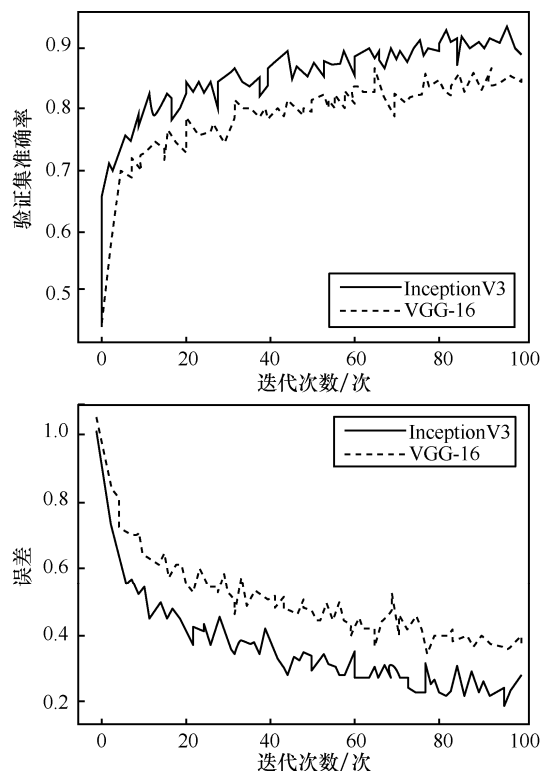


图7 预加载权重训练

本文采用两种训练集标注的方法研究背景复杂图片对模型准确率的影响:第一种为将训练集图片分为3类,即背景简单的合格图片、背景复杂的合格图片 and 不合格图片,分别用0、1和2代替;第二种为将训练集图片分为2类,即背景简单的合格图片为0类,背景复杂的合格图片 and 不合格图片为2类;分类模型在测试集上的准确率如图8所示。

图8(a)是第一种分3类的标注方式,虽然InceptionV3识别率总体高于VGG-16,但是仍有很多合格图片(0类)被错分为不合格图片(2类)。根据混淆矩阵精度判别指标,InceptionV3存在false negative率过高的问题。由于运维质检业务需要假阴性率相对较低的模型,VGGNet具有更低的假阴性率,所以仍然选择VGG-16作为待优化的基础模型。图8(b)是采用第二种标注方式的测试结果,模型准确率远高于第一种标注方式,由此可见,图片背景复杂是影响模型识别率的主要因素。

样本分类	识别结果	VGG-16	识别率	Inception V3	识别率
0 (337)	0	284	87.8%	296	84.2%
	1	33		39	
	2	20		2	
1 (384)	0	52		50	
	1	298	85.6%	329	77.6%
	2	34		5	
2 (478)	0	34		49	
	1	15		90	
	2	429	70.9%	339	89.7%

(a)

样本分类	识别结果	VGG-16	识别率
0 (300)	0	289	96.3%
	2	11	
2 (170)	0	8	95.3%
	2	162	

(b)

图8 测试结果对比

2.4.2 目标分割

对于背景复杂图片,可采用目标分割的方法具体定位到图中待测物位置以判断图片是否合格。Mask RCNN^[5]是近年来基于深度学习的最成功的目标分割检测算法之一。基于 Fast RCNN,添加一个 Mask 分支基于现有检测对待测目标进行并行的像素分割。以运维质检图片为例的 Mask RCNN 为例,架构如图9所示。

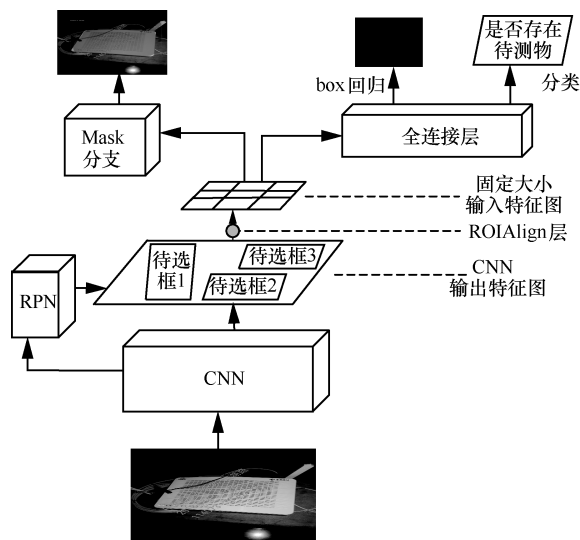


图9 Mask RCNN 结构

原始图像先通过 CNN 提取特征,同时进入

RPN 网络学习待测物所在区域,RPN 网络使用滑动窗口的方法,选出数个待测物可能所在候选框,并与 CNN 输出特征组合成特征图。ROIAlign 层将特征图缩放为全连接层输入格式。Mask 分支分割出具体待测物像素,全连接层判别待测物是否存在,同时根据特征图对每个候选框打分,并使用一个回归器预测新的候选框。

Mask RCNN 在运维质检中有良好的效果,但是需要在数据标注过程中将图片中待测物像素分割,这会带来巨大的标注工作量。因此将目标分割与图像分类的模型相融合,既改善图像分类模型中对复杂背景图片识别率低的问题,又减轻了目标分割模型所需人工标注工作量过大的问题。

2.5 融合模型

基于已训练的 VGG-16 模型,只需要对原始图片数据量约 1/3 的背景复杂图片进行训练,极大地减少了标注量。根据第 2.4.2 节中的分析,VGG-16 中背景复杂的不合格图片是影响模型准确率的主要因素。因此,可以采用加权投票的方式主要针对背景复杂图片进行投票,对各模型判别结果合并处理,变换如式(2):

$$P_{\text{vote}} = m \times P_{\text{vgg16}} + (1 - m) \times P_{\text{mask-rcnn}}, 0 < m < 1 \quad (2)$$

其中, P_{vgg16} 和 $P_{\text{mask-rcnn}}$ 分别为分类和定位模型中存在待测物的概率,根据训练结果可以使用浅层学习模型,如决策树、SVM 等,调整 m 大小调整融合效果。

3 现网试运行与迁移学习优化

3.1 现网试运行

在客保系统应用融合模型,对半天产生的现网数据进行实时检测,模型准确率见表1。

由表1可知,融合模型在 VGG-16 模型准确率上整体提高了约 7.5 倍,主要提高了背景复杂图片的识别率,改善了运维质检图片中施工场地复杂所导致的图片质量低影响准确率的问题。



表 1 融合模型准确率

数据量	分类	识别率
合计现网图片总量: 4 292 张	合格	91.10%
合格图片: 3 192 张	不合格	82.50%
不合格图片: 1 100 张		

3.2 迁移学习优化

当数据集改变时,深度学习模型需要新的数据集对其重新训练,以适应数据集的变化。但是面对电信运维质检业务所产生的海量数据,对模型重新训练可能会带来较大的时间成本。针对这个问题,本文使用迁移学习使融合模型能够在短时间内根据新的数据情况调整自身以适应运维图片数据的变化。

迁移学习只训练特定层,一般为最后的全连接层。由于全连接层参数占据了神经网络大部分待训练参数,所以迁移学习的方法是可行的。迁移学习能在保障模型准确度基础上减小训练时间和不需要大量数据集,训练效果如图 10 所示。

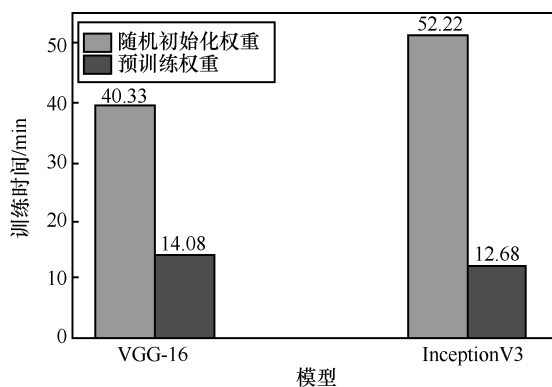


图 10 迁移学习训练时间对比

由图 10 可见,相较于全层训练,使用迁移学习的训练时间显著地减少,VGG-16 训练时间减少了 65%,InceptionV3 减少了 75.9%。在本文提出的融合模型中,只需要对基础的 VGG-16 模型的最后 3 层全连接层进行迁移学习。

4 结束语

文本针对电信运维的装维质检中,需要人工识别施工现场图片以评估装维质量的问题在深入

研究和对比了多种图像分类技术后,提出了一种适应于电信运维质检系统的融合模型,并开发了相应的系统,通过测试数据及现网数据的验证,取得良好的效果。

参考文献:

- [1] ANDREOPOULOS A, TSOTSOS J K. 50 years of object recognition: directions forward[J]. Computer Vision and Image Understanding, 2013, 117(8): 82789.
- [2] RUSSAKOVSKY O, DENG J, SU H, et al. ImageNet: large scale visual recognition challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211-252.
- [3] KRIZHEVSKY A, SUTSKEVER I, HINTONG E. ImageNet classification with deep convolutional neural networks[C]//The 25th International Conference on Neural Information Processing Systems, Dec 3-6, 2012, Lake Tahoe, Nevada, USA. New York: ACM Press, 2012: 1097-1105.
- [4] HE K, ZHANG X Y, RENS Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1914.
- [5] HE K, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[C]//2017 IEEE International Conference on Computer Vision, Oct 22-29, 2017, Venice, Italy. Piscataway: IEEE Press, 2017: 2980-2988.

[作者简介]



范博(1993—),男,中国电信股份有限公司上海研究院数据挖掘工程师,主要研究方向为图像算法研究和数据挖掘。



邱芸(1980—),女,中国电信股份有限公司上海研究院网络系统架构师,主要研究方向为基于网络数据的应用研究开发、物联网网络数据分析。

沈雷(1970—),男,中国电信股份有限公司上海研究院算法专家,主要研究方向为移动网络应用研究开发、人工智能算法。



国之重器出版工程

网络强国建设



多接入边缘计算 (MEC) 及关键技术

Multi-access Edge
Computing (MEC) and
Key Technologies

张建敏 杨峰义 武洲云 张郑锟 王煜炜 编著

人民邮电出版社

书号: ISBN 987-7-115-50235-3

定价: 128.00元

出版时间: 2019年1月

内容简介

本书结合未来5G网络演进趋势全面讨论了5G多接入边缘计算(MEC)及关键技术,内容涵盖5G网络需求及架构、MEC概念应用场景及需求分析、MEC系统架构及部署组网策略以及MEC系列关键技术,包括本地分流技术、缓存加速、网络能力开放、移动性管理、固移融合、计算卸载等不同层面。

本书可供具有一定移动通信技术基础的专业技术人员或管理人员阅读,也可作为高等院校相关专业师生的参考读物。

ISSN 1000-0801



邮发代号: 2-397 国外代号: M841 定价: 68.00元