

# 基于图片 AI 识别的维修质量管控方案

黄媛媛, 甘顺水

(中国移动通信集团广西有限公司, 广西 南宁 530000)

**摘要** 文章首先介绍了目前通信企业基站市电维修费用的管控现状;其次,从故障维修需求派单到维修费用结算稽核的过程,指出维修费用存在风险管控问题;最后,提出一种基于图片智能化识别自动稽核方案,实现对故障维修过程的设备类型准确性、数量真实性、维修质量等问题管控,以期能为有效提升维修费用的使用效能提供借鉴。

**关键词** 基站市电维修;费用稽核;图片识别技术

**中图分类号**: TP317.4; TN91

**文献标识码**: A

**文章编号**: 1007-0745(2023)06-0022-03

目前在通信动力设备系统日常故障维护框架合同费用使用管理过程中,对维修需求工单无统一的稽核标准,系统无法确认材料数据与维修现场实际情况是否一致,导致审核人员稽核效率低、稽核效果差;其次,在框架合同维护使用过程中时常发生“偷工减料”现象,框架维护费用支出存在审计风险。为增强维修施工质量和降低维护费用管控风险,本文提出了一种基于图片智能化识别技术的维修费用稽核方案。

当前维修费用稽核人员主要依靠系统提供的数据进行稽核,主要存在人工稽核工作量巨大、稽核效果比较差、费用审计风险大等问题,为提高稽核人员对费用稽核的管控能力,增加维修现场情况的把控能力,在稽核环节增加图片化识别和自动稽核功能,增强维修工程量和维修质量的辨识能力,从而降低虚报工程量、设备型号错漏等廉洁风险问题。

## 1 图片化 AI 识别稽核方案原理

维修需求工单中的设备图片,通过数据接口上传到图片识别智能平台系统<sup>[1]</sup>,实现对图片的设备类型、站点信息、站点经纬度、拍照时间等关键信息的获取。然后与设备图片库中的设备类型信息进行自动比对匹配,系统自动输出判定结果到应用模块,由应用模块返回比对结果到需求派单系统。

图片 AI 识别稽核实现流程如图 1 所示,主要分为以下五个阶段:

1. 触发阶段:用户通过派发维修需求工单,同时在需求工单里上传对应站点维修过程相关设备图片。
2. 工单数据整合:按既定的格式分类汇总各类型设备图片。
3. AI 智能分析:采用 GoogLeNet 特征算法、OCR

信息提取算法,对图片关键信息进行提取识别。

4. 数据评估匹配:对图片识别的信息进行算法模型匹配,确认图片的类型、站点信息、数量等识别结果。

5. 返回稽核结果:返回识别结果给需求工单系统,由工单系统确认结果是否通过。

## 2 图片数据预处理

1. 为提高图片数据处理效率,主要通过以下三种方式对图片进行集中预处理。

2. 图像缩放归一化处理<sup>[2]</sup>:对图片数据每一个维度的值进行重新调节,使数据向量落在 [0,1] 或 [-1,1] 区间内。例如电力电缆原始图片的像素值在 [0,255] 区间中,将像素值除以 255,归一到 [0,1] 区间内。

3. 样本均值消减处理:假设所有图片数据是平稳的,即图片数据每一个维度的统计都服从相同分布,在每个图片数据样本上减去图片数据的统计平均值,可以消除图像平均亮度值。

4. 特征数据归一化处理:图像数据特征标准化操作目的是使得图片数据在每一个维度上具有零均值以及单位方差。例如在电力电缆图片中,首先使用全体数据来计算每一个维度上数据的均值,然后在每一个维度上都减去该均值,最后在数据每一维度上除以该维度上数据的标准差,以此来实现图片的特征数据归一化处理。

## 3 图片识别算法实现

### 3.1 GoogLeNet 图片识别算法

图片识别技术主要采用 GoogLeNet 算法<sup>[3]</sup>,主要特点使用  $1 \times 1$  的卷积结构,在多个尺度上并行卷积然后再组合在一起,相比传统的图片识别算法,其主要优点如下:

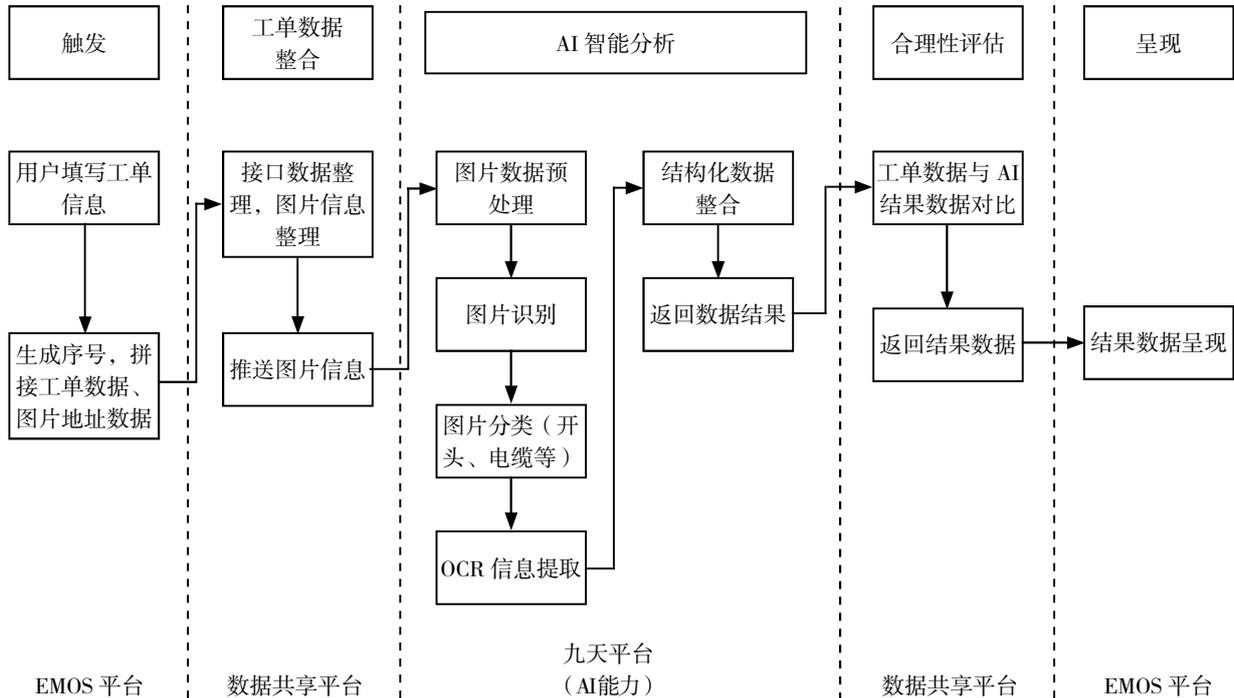


图 1 图片化 AI 识别稽核实现方案流程图

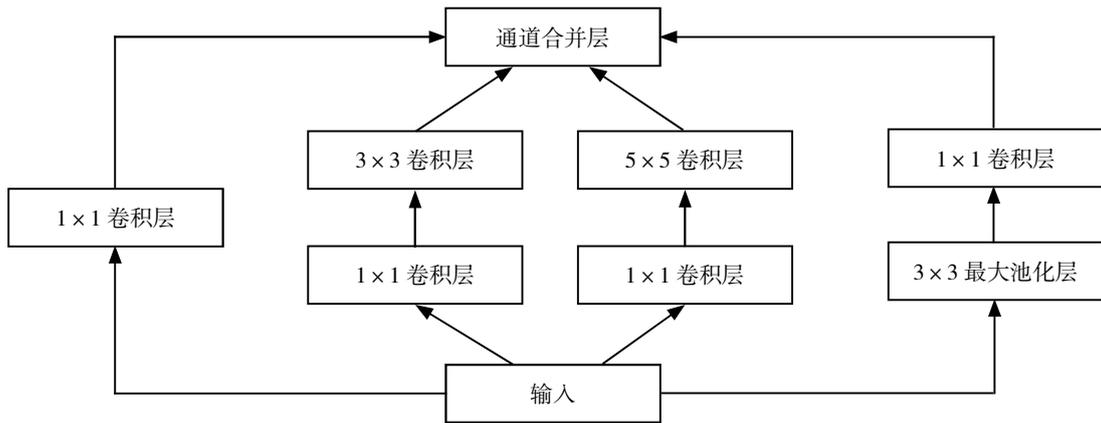


图 2 GoogleNet 算法流程结构

1. 采用 CNN 典型结构: 一堆卷积层 + 若干全连接层。
2. 通过增加层数和层的尺寸来解决大规模数据集问题、解决过拟合问题。
3. 采用深度网络结构增加网络的表现能力和数据降维, 解决图像数据处理计算瓶颈。
4. R-CNN 对象检测方法首先利用低级特征在分类不可知的情况下寻找潜在目标, 然后使用 CNN 分类器识别潜在目标所属类别, 而 GoogLeNet 算法在这两个阶段都进行了增强效果。

GoogLeNet 算法的核心思想是利用 Inception 模块来

叠加应用, 如图 2 所示。

### 3.2 OCR 信息提取算法

OCR 信息提取算法<sup>[4]</sup>主要是通过“端到端文字检测识别、快速定向文本定位、倾斜文字检测”三个算法的组合应用来实现对照片文字信息进行提取。

端到端文字检测识别算法涉及文字检测和文字识别两个方向, 其基本原理是设计一个同时具有检测单元和识别单元的模型, 共享其中两者的特征提取单元, 并且实施联合训练活动。在推断阶段, 该模型可以在一个前向传播中预测场景图像中的文本位置和内容信息。

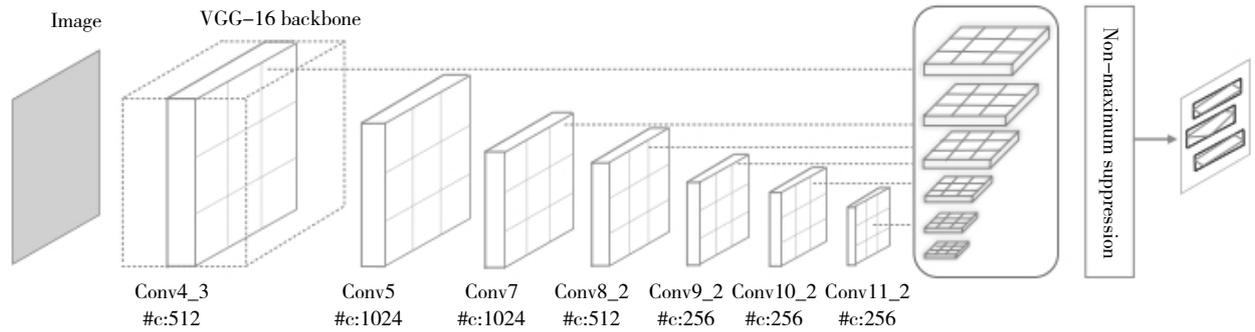


图3 TextBoxes++ 模块功能流程示意图

模型由主干网络的特征提取单元、检测单元、识别单元三个子单元，以及检测单元与识别单元的一种交互模式组成。

在主干网络模块中引入空洞卷积使其有效的输出序列为8个，维持较大的接收域，便于捕获自然场景中尺寸上有巨大变化的文本信息。识别单元由一个 LSTM-Decoder 来依次预测各个字符，直到预测类别为终止字符串才停止。文本识别任务粒度较细，需要更密集的特征，为此引入了多尺度特征融合的方法来加强输入识别单元的特征。通过 RoI Masking 新的交互模式来过滤相邻文本和背景信息，免去对文本特征的矫正操作。

通过骨干卷积网络 Resnet50 实现文本快速定位检测和识别，使用骨干网络最后一层卷积特征，结合了浅层结构特征和深层语义特征。在骨干网络输出层上实现密集文本框预测，在骨干网络之后增加一个 6\*6 卷积层，分别为该像素属于文本区域的概率、该像素到文本框上下左右的距离和文本框的旋转角度。

由于图像中的文本大小差异大，而识别网络通常只能输入固定大小的矩形窗口<sup>[5]</sup>，因此需要对目标区域做缩放和填充使其与识别网络的输入窗口相匹配。具体实现就是 ROI Align 在保持目标纵横比不变的情况下将目标缩放至窗口内，使目标长边与窗口对齐，再沿目标短边方向补零使其与窗口对齐。此外，由于窗口内的像素格点位置在缩放前很可能不在原图的像素格点位置上，故使用插值法计算该点处的特征值<sup>[6]</sup>。

通常文本目标具有较大的纵横比，为了更好地适应歪斜的文本，快速定向文本定位中的目标框提案具有旋转角度。基于双向 LSTM+CTC，从 ROI Rotate 的输出开始，经过一系列卷积池化层，输入到双向 LSTM，最后模型通过 CTC 层进行路径概率求和输出最终文本结果。TextBoxes++ 只包括卷积层和池化层的全卷积结

构，可以检测多角度文本，能够处理任意形状图像。其用卷积层取代了平均池化，有利于多尺度的训练和测试，功能网络结构如图3所示。

TextBoxes++ 继承了 vgg-16 的结构并且保留了 conv 1\_1 到 conv5\_3 层，把最后两个全连接层变成卷积层。另外八个卷积层被分为四个阶段（conv8 到 conv11），通过最大池化形成不同的分辨率。

#### 4 系统

当前通信基站市电故障维修具有工作量大、站点比较分散、维修内容多变、施工管控难度大等特点，维修框架使用过程中，对现场维修工作质量的把控，缺乏强有力的支撑手段，导致维修施工单位“质量弱化”风险大。对施工人员的管理、操作、服务等过程实施电子化手段进行管控，促进维护施工人员提高施工维修的规范性及合理性，减少审核工作量，提高了费用稽核效果，节约了管理成本，同时大大降低了框架使用管理廉洁风险。

#### 参考文献：

- [1] 张可. 物联网及其数据处理 [M]. 北京: 国防工业出版社, 2018.
- [2] 高隆昌, 杨元. 数据建模基础理论 [M]. 北京: 科学出版社, 2007.
- [3] 安凤平. 自适应图像处理算法及应用研究 (2010 版) [M]. 北京: 科学出版社, 2019.
- [4] 同 [3].
- [5] 言有三. 深度学习之图像识别: 核心技术与案例实战 [M]. 北京: 机械出版社, 2019.
- [6] 郑列, 李刚. 图像灰色模型理论与算法 [M]. 北京: 科学出版社, 2017.