

铁路车辆监测图像识别模型训练及验证平台研究

杨 凯, 张 淼, 祁苗苗

Training and verification platform for railway vehicle monitoring image recognition model

YANG Kai, ZHANG Miao, and QI Miaomiao

引用本文:

杨凯, 张淼, 祁苗苗. 铁路车辆监测图像识别模型训练及验证平台研究[J]. 铁路计算机应用, 2023, 32(6): 26–30.

YANG Kai, ZHANG Miao, QI Miaomiao. Training and verification platform for railway vehicle monitoring image recognition model[J]. *Railway Computer Application*, 2023, 32(6): 26–30.

在线阅读 View online: <http://tljsjyy.xml-journal.net/2023/16/26>

您可能感兴趣的其他文章

Articles you may be interested in

图像识别技术在BIM运维管理中的应用研究

Image recognition technology applied to BIM operation and maintenance management

铁路计算机应用. 2021, 30(7): 76–79

代价敏感正则化有限记忆多隐层在线序列极限学习机及图像识别应用

Cost-sensitive regularized finite-memory multi-hidden-layer online sequential extreme learning machine and its application in image recognition

铁路计算机应用. 2018, 27(5): 18–23

基于LeNet-5的编组站内机车车号识别系统的研究

Locomotive number identification system in marshalling station based on LeNet-5

铁路计算机应用. 2019, 28(9): 16–20

基于Kubernetes的铁路客运营营销深度学习平台的设计与实现

Deep learning platform of railway passenger transport marketing based on Kubernetes

铁路计算机应用. 2021, 30(1): 57–61

高铁周边环境的空天光学遥感监测及特点分析

Aerospace optical remote sensing monitoring and characteristic analysis to high-speed railway surrounding environment

铁路计算机应用. 2021, 30(7): 66–71

基于卷积神经网络优化算法的列车智能测试系统技术研究

Train intelligent testing system based on convolution neural network optimization algorithm

铁路计算机应用. 2019, 28(5): 1–5,10



关注微信公众号, 获得更多资讯信息



铁路车辆监测图像识别模型训练及验证平台研究

杨 凯, 张 森, 祁苗苗

(中国铁道科学研究院集团有限公司 电子计算技术研究所, 北京 100081)

摘 要: 针对铁路车辆轨边图像检测系统现有图像自动识别模型训练及评价过程中训练数据不足、数据质量不高、评价标准不一致等问题, 研究铁路车辆监测图像识别模型训练及验证平台。设计故障图像数据统一接入, 专家标定数据形成, 自动识别模型接入、训练、对比评测等方法, 为故障图像自动识别模型提供标准训练数据、统一评测验证与管理服务的能力。实践表明, 该平台实现了车辆故障图像数据的集中汇总与统一管理, 为铁路车辆监测图像自动识别技术的发展提供了有力支持。

关键词: 轨边图像检测系统; 安全监测; 图像识别; 模型训练; 数据标定

中图分类号: U279 : U298.1 : TP39 **文献标识码:** A **DOI:** 10.3969/j.issn.1005-8451.2023.06.05

Training and verification platform for railway vehicle monitoring image recognition model

YANG Kai, ZHANG Miao, QI Miaomiao

(Institute of Computing Technologies, China Academy of Railway Sciences Corporation Limited, Beijing 100081, China)

Abstract: Aiming at the problems of insufficient training data, low data quality and inconsistent evaluation standards in the training and evaluation process of the existing image automatic recognition model of the railway vehicle trackside image detection system, this paper studied the training and verification platform for railway vehicle monitoring image recognition model, designed methods such as unified access to fault image data, expert calibration data formation, automatic recognition model access, training, comparative evaluation, etc., provided the ability of standard training data, unified evaluation, verification, and management services for fault image automatic recognition models. The practices show that this platform implements centralized collection and unified management of vehicle fault image data, provides strong support for the development of automatic recognition technology for railway vehicle monitoring images.

Keywords: trackside image detection system; safety monitoring; image recognition; model training; data calibration

随着铁路运输的快速发展, 铁路车辆对运行安全的需求日益增强。为此, 中国国家铁路集团有限公司(简称: 国铁集团)在全国铁路推广应用了货车故障轨边图像检测系统(TFDS, Trouble of moving Freight car Detection System)^[1]、客车故障轨边图像检测系统(TVDS, Train of passenger Vehicle failures Detection System)^[2]、动车组运行故障轨边图像检测

系统(TEDS, Train of EMU failures Detection System)等车辆运行安全监测检测系统^[3], 保障了铁路车辆运行安全。

上述系统中的图像自动识别模型存在训练素材分散、故障数据缺乏标定、无统一的模型评测验证指标等问题, 故障图像自动识别报警的准确度无法满足实际需求, 目前仍主要依靠人工来分析过车图像。随着过车频次的日益增加, 人工监测作业强度增大, 人工检视的效率降低。面对铁路车辆逐渐增长的安全监测需求, 为保证监测效率, 亟需深入研究图像识别技术在车辆故障自动报警方面的应用。

收稿日期: 2022-11-30

基金项目: 中国国家铁路集团有限公司科技研究开发计划项目(N2021J017)

作者简介: 杨 凯, 副研究员; 张 森, 研究实习员。

当前，铁路车辆运行安全监控系统所使用的图像识别技术依赖于大量的图像型数据，需要从中提取学习特征，训练模型并对图像内容进行自动识别。有较多学者利用机器学习技术如支持向量机、神经网络（RNN、CNN、R-CNN 等）^[4-5]，以及深度学习方法如 YOLOv4 等^[6]，对铁路车辆的受电弓、车体、走行部等关键部件（如轮对、螺栓、折角塞门等）进行故障图像识别模型研究，在其自身的数据集中，上述研究模型的故障识别率可超过 75%^[7]。对于故障图像自动识别模型的训练而言，输入的数据至关重要，需要大量相同格式的标准数据及高质量数据以实现图像识别模型的最佳表征。为此，需要人工对数据进行注释、标记并将准确的结论纳入模型训练过程，以提高数据质量。

基于以上研究，本文开展故障图像自动识别模型训练与统一验证平台研究，实现铁路图像类故障数据统一接入管理、专家分类标定，提供模型训练、评测等服务能力，从而提高故障图像自动识别报警准确率。

1 方案设计

1.1 总体架构

本文设计了故障图像数据获取、模型接入、图像数据标准化、模型训练及其验证等方法，构建铁路车辆监测图像识别模型训练及验证平台（简称：训练验证平台），主要包括数据接入、训练模块、评测模块、管理模块和服务模块，实现故障图像数据采集存储、专家标定故障库管理、模型训练与评测验证等功能。平台总体架构如图 1 所示。

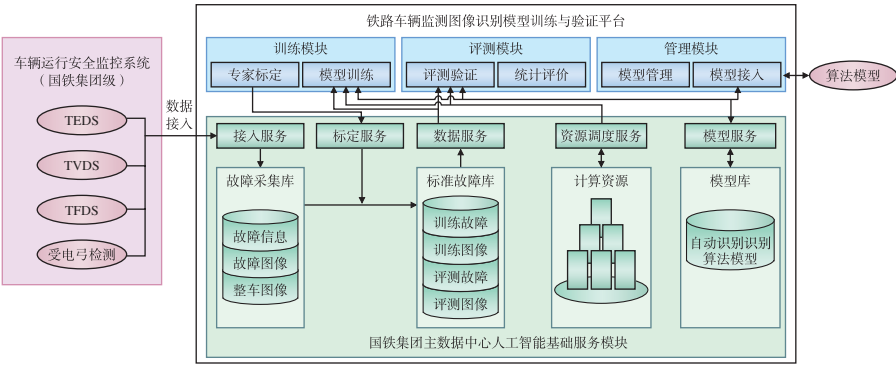


图1 平台总体架构

（1）在数据接入过程中，利用国铁集团级车辆运行安全监测相关系统（TEDS、TVDS、TFDS、受电弓检测）的内部数据共享接口，获取故障图像文件、故障详细信息等数据，接入服务模块故障采集库中。

（2）训练模块通过专家手动操作，对故障采集库中的图像数据进行重新标定，生成标准化数据存入服务模块标准故障库之中。按照需要将数据分为训练故障数据、训练图像数据与评测故障数据、评测图像数据，对外提供相应数据服务。在模型训练过程中从服务模块的模型库、标准故障库分别调用模型与训练故障图像数据，配置计算资源进行模型训练，训练结果返回模型库。

（3）评测模块使用标准故障库中评测故障图像数据，从服务模块的模型库、标准故障库分别调用模型与评测故障图像数据，将运算的故障识别结果与真实的故障进行比对，计算模型准确率与识别率，评价模型的性能。

（4）管理模块负责故障图像自动识别报警模型的上传，并与服务模块模型库交互，将符合规范的模型存入服务模块的模型库供模型训练与评测验证调用，并提供模型训练结果反馈功能，以供模型更新迭代。

（5）利用国铁集团主数据中心人工智能基础服务模块作为后台服务模块，用于提供数据存储服务，存储接入的故障图像数据与图像识别模型文件。通

过接口实现故障图像数据与模型文件调用；提供计算服务器资源，实现模型训练与评测验证。

1.2 数据采集及存储设计

故障图像数据采集与存储主要方式如图2所示。

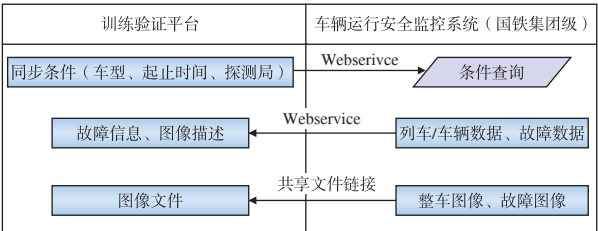


图2 数据采集与存储示意

（1）故障图像数据的采集通过接入国铁集团级车辆运行安全监测相关系统共享接口同步数据的方式实现。训练验证平台以车型、起始时间、结束时间、探测局作为查询条件，获取上述系统数据库中的整车原始图像数据、人工填报的故障数据及对应故障图像。

（2）故障图像数据以共享文件链接的方式获取，根据检车作业的不同流程，车辆故障图像存在原始图像、预览缩略图像与人工分析故障截图3种模式，在接入过程中按照原始图像、缩略图、截图的优先顺序获取故障图像数据，包含图像文件、列车及车辆信息、故障详情、图像描述等信息。

（3）图像文件需要转换为统一格式，按照接入日期分别存储于服务器对应磁盘目录，车辆信息、故障详情、图像描述等数据存入故障采集库。图像文件与故障采集库通过图像名称、图像文件路径建立链接。

1.3 功能设计

1.3.1 数据接入

与国铁集团级车辆运行安全监控系统数据同步，设置车型、起止时间、探测局查询条件，获取国铁集团级服务器中相应车型故障数据与图像文件，并以列表形式展示包括接入时间、探测局、监控中心、探测站、故障通道、故障位置、故障模式、图像模式等信息，可手动删除不需要的数据。

1.3.2 专家标定

专家对接入的故障数据与图像数据进行标定操

作，对列表中同步的数据进行重新标定，专家在车辆图像上用红框标出故障范围（X轴、Y轴、宽、高），并备注故障位置、故障模式、检测难易度，实现故障图像数据的标准化。

1.3.3 故障库管理

完成专家标定的数据标记为评测故障数据，在故障库管理界面可按照故障来源、车型、部位、模式与难度筛选数据，由管理员选择数据将其更改标记为训练故障数据。训练故障数据也可重新标记为评测故障数据，对不需要的故障数据可永久删除。

1.3.4 模型训练

平台将上传的图像识别模型文件存储至本地，一个模型文件对应一种或多种车辆故障类型，按照训练需求，选择训练车型、故障部位、故障形式，从训练故障数据中匹配相应数据，利用服务模块中国铁集团数据中心的服务计算资源进行模型训练。

1.3.5 评测验证

可对已完成训练的模型进行图像识别能力评测验证，按照模型对应的车辆故障类型从评测故障数据中选择多条用于验证的数据，统计显示识别耗时、总体识别率与总体准确率。

1.3.6 统计分析

统计并展示包括每日（月）的接入故障图像数量；专家未标定、已标定数量；故障库中训练故障与评测故障数量；接入模型数；模型训练、评测验证总次数等数据。

2 图像识别模型训练及验证

2.1 资源标准设计

目前，接入故障采集库的故障图像数据为全国铁路各铁路局检车分析员的人工分析数据，而由于不同分析员的作业习惯存在差异，且图像文件存在原图、缩略图、截图等不同类型，导致图像数据存在故障标定误差、图像质量良莠不齐等问题，无法直接用于模型训练。

为提高模型训练的数据质量，并保证数据格式统一，本文设计通过专家标定实现图像数据标准化的方法，包括故障标定和标定审核2个步骤，其基本流程如图3所示。

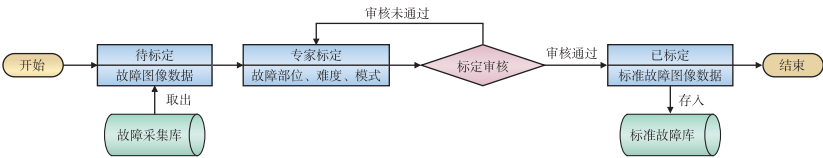


图3 故障图像数据专家标定流程

(1) 故障标定：由多名具有一定检车分析经验的专家对故障采集库中的图像数据统一标定，包括在故障图像上以故障框形式重新标定故障在图像中所处的位置、故障在车辆所处的部位、故障模式及故障识别难度。

(2) 标定审核：对于同一故障图像，在一位专家进行故障判断后，由另一专家对其标定位置、信息进行审核，判断是否准确，审核通过则该故障图像完成标定，存入标准故障库，否则需要由前一专家重新标定。

完成专家标定、审核过程的图像数据存放于标准故障库中，数据由人工分成评测数据与训练数据。评测数据仅供模型能力评测验证使用，训练数据为各模型故障识别训练使用，设计采用 Spring Cloud 微服务内部接口实现图像数据调用服务。

2.2 模型训练验证方法设计

为实现模型的训练，并满足多种模型评价及检验的需求，本文设计了故障图像自动识别报警模型的训练及评测验证方法，利用标准故障库中的训练、评测故障图像数据与服务模块的计算资源训练模型检验模型的性能。方法步骤如下：

(1) 模型以 .py、.jar、.dll 等文件形式上传至训练验证平台，能够读取图像文件后按照标准故障库的数据规范，输出故障图像部位、故障部件及故障模式。

(2) 模型训练时，按照车型、故障部位、故障模式为模型配置所需要的故障图像数据，服务模块提供计算资源，进行模型训练。平台记录训练结果，根据训练结果调整模型参数或重新上传文件实现模型迭代。

(3) 评测验证时，通过人工或随机选择故障图像数据的方式，开展评测验证的模型运算，将模型

的输出结果与检验图像数据的真实结果进行对比，得出模型的总体识别率与总体准确率，形成模型能力评估报告^[8-9]。

3 平台应用

结合上述方法，本文研发了铁路车辆监测图像识别模型训练及验证平台，实现数据接入、专家标定、故障库管理、模型接入训练、评测验证与统计分析等功能。平台现已应用于 TEDS、TFDS、TVDS 等系统的故障自动识别模型训练，模型训练界面如图 4 所示。



图4 模型训练界面

截至目前，该训练验证平台已实现 TEDS 故障图像数据的接入，存储图像数据 211 855 例，构建标准故障库 105 567 例（其中，训练数据 85 421 例，评测数据 20 146 例），已实验开展了 5 种非转向架部位图像自动识别模型的训练，其识别率变化如图 5 所示。

根据统计数据可见，通过多个周期的模型训练迭代，各模型识别率均有明显提升，平均增长率达到 29.8%，提高最多的可达 35.3%。由此可见，本文设计的训练验证方案、数据标准化等方法，显著提升了各模型的识别率。

4 结束语

本文研究了铁路车辆监测图像识别模型训练及

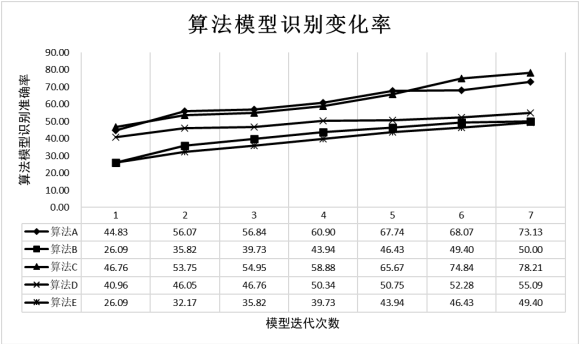


图5 各模型识别率变化

验证技术，研发了铁路统一的车辆图像识别模型训练及验证平台，利用国铁集团主数据中心服务器资源，实现铁路车辆运行安全监控故障图像数据的集中汇总、统一管理，为各类图像识别模型提供有效的训练资源；实现了统一、标准的评测验证功能，为推进铁路图像自动识别技术发展提供了有力的数据基础和技术支持。后续研究将集中于模型自动报警技术能力验证、模型运用等方向。

参考文献

[1] 赵长波, 陈 雷. 铁路货车安全监测与应用概论[M]. 北京: 中国铁道出版社, 2010.

[2] 张 雪. 客车故障轨边图像检测系统 (TVDS) 运用研究

[J]. 现代制造技术与装备, 2019 (3): 112-113.

[3] 刘 彬. 动车组运行故障图像检测系统 (TEDS) 运用研究与思考 [J]. 中国铁路, 2017 (12): 61-65.

[4] Zhou F Q, Zou R, Qiu Y F, et al. Automated visual inspection of angle cocks during train operation [J]. Proceedings of the Institution of Mechanical Engineers, Part F:Journal of Rail and Rapid Transit, 2014, 228(7): 794-806.

[5] Yao Z K, He D Q, Chen Y J, et al. Inspection of exterior substance on high-speed train bottom based on improved deep learning method [J]. Measurement, 2020, 163(5): 108013-108024.

[6] Hsieh C C, Lin Y W, Huang W H, et al. A YOLOv4 based cloud server for railway fault fastener detection YOLOv4 [J]. Journal of the Chinese Institute of Civil and Hydraulic Engineering, 2021, 33(4): 263-272.

[7] Tang R F, De Donato L, Bešinović N, et al. A literature review of Artificial Intelligence applications in railway systems [J]. Transportation Research Part C:Emerging Technologies, 2022, 140(2): 103679.

[8] 杨 凯. TEDS动车组故障图像数据库建立技术研究 [J]. 铁道机车车辆, 2020, 4 (6): 38-42, 81.

[9] 杨 凯, 刘 彬, 崔中伟, 等. 基于多源数据分析的TEDS故障识别技术研究 [J]. 铁路计算机应用, 2019, 28 (4): 14-17.

责任编辑 朱 一