

doi: 10.7690/bgzdh.2024.05.006

面向装备试验数据的管理平台

李胜昌, 胡明哲, 谢志豪

(中国兵器工业计算机应用技术研究所网络信息体系论证与研发部, 北京 100089)

摘要: 为高效管理装备试验数据资源, 基于 SpringCloud 微服务架构和 Kubernetes 容器调度平台, 设计装备试验数据管理平台。融合基于 DBNet+CRNN 的通用文字识别算法, 实现试验文书、图像、视频、音频和结构化数据的采集、整编以及任务管理等功能。结果表明: 该平台能有效提升对装备试验数据的全生命周期管理能力, 为进一步挖掘分析装备试验数据潜在价值提供全面有效的数据资源支撑。

关键词: 试验数据; 采集整编; 微服务; 文字识别

中图分类号: TP315 **文献标志码:** A

Management Platform for Equipment Test Data

Li Shengchang, Hu Mingzhe, Xie Zhihao

(Demonstration and Research Department of Network Information System, Computer Application Technology Institute of China North Industries Group Corporation Limited, Beijing 100089, China)

Abstract: In order to efficiently manage equipment test data resources, an equipment test data management platform is designed based on SpringCloud micro-service architecture and Kubernetes container scheduling platform. The general character recognition algorithm based on DBNet + CRNN is integrated to realize the collection, reorganization and task management of test documents, images, videos, audios and structured data. The results show that the platform effectively improves the whole life cycle management capability of equipment test data, and provides comprehensive and effective data resource support for further mining and analyzing the potential value of equipment test data.

Keywords: test data; collection and compilation; micro-service; character recognition

0 引言

武器装备在型号论证、研制生产、试验鉴定以及运用保障全生命周期过程中, 会产生并积累大量装备静态参数指标数据以及动态作战使用性能数据^[1]。随着装备数字化建设的不断加快, 装备试验数据积累总量也呈现爆发式的增长趋势^[2], 但由于早期装备试验历史数据大多以纸质试验文书的形式存储, 数字化程度低, 很难在当今数字化信息系统中有有效利用起来^[3]。而且, 装备试验数据存在类型异构、数据格式不一致等问题, 导致装备试验数据在不同的生命周期期间互相脱节, 无法有效关联使用^[4]。同时, 由于装备试验数据采集手段多样、采集设备分散、采集部门繁多, 海量的装备试验数据分散存储在试验基地、靶场^[5]以及工业部门实验室等不同的地点, 数据的汇聚、流转和共享障碍重重, 形成“试验数据孤岛”, 难以发挥数据融合分析的效能^[6]。

为满足装备试验鉴定相关方对试验数据有效管理利用的需求, 解决目前装备试验数据在采集整编方面存在的问题, 需要研制专门针对装备试验数据

的采集整编管理平台, 提供强耦合装备试验全流程的数据采集、整编以及管理服务, 为装备试验数据进一步挖掘分析潜在价值提供全面有效的数据资源支撑。

1 现状

在试验数据管理系统研究方面, 美国 Newtera 公司研发的 Newtera TDM (test data management) 试验数据管理系统是行业内的标杆产品, 其应用范围广泛, 包括通用、福特、波音等主要面向民用领域的制造业企业, 也包括洛克希德·马丁、霍尼韦尔等军工企业^[7]。国内相关企业也开发了类似的试验数据管理系统, 天健通泰科技研发的神鹰 TDM 系统针对试验数据管理和利用效率问题, 提供试验数据的一站式管理功能, 能有效缩短试验工作周期。此外, 神州普惠的 APP TDM 试验数据管理系统对分散的试验数据进行统一的存储管理, 提高了试验数据的一致性和安全性^[8]。总体来说, 国内试验数据管理系统相较于国外, 在系统稳定性以及数据精细化管理方面还有一定差距。

收稿日期: 2024-01-23; 修回日期: 2024-02-25

第一作者: 李胜昌(1987—), 男, 湖北人, 硕士。

虽然现有试验数据管理系统基本具备对试验数据的统一存储管理能力，但大多缺乏对结构化数据的规范整理，且无法进一步有效提取非结构化数据信息，难以满足海量装备试验数据的规范化、数据特征的智能化提取需求。

2 装备试验数据管理平台

2.1 技术架构

装备试验数据管理平台采用 B/S 架构，基于

SpringCloud 微服务架构^[9]和 Kubernetes 容器调度平台^[10]进行开发设计，相比传统本地服务调用的模式，微服务模式将项目按功能需求拆分成独立的服务，并运行在独立的虚拟机或 Java 进程里，支持软件的敏捷、松耦合、模块式开发，满足平台的可扩展性开发需求。系统包括数据源、支撑层、数据采集整编、数据资源、采集整编工作管理、运维管理、数据资源库管理以及统一门户等。系统技术架构如图 1 所示。



图 1 系统技术架构

1) 数据源。

数据源包括试验文书经过扫描后形成的电子文档、音频、视频、图像、专用格式二进制文件等电子文件数据，以及多种关系型数据库表等，涵盖大范围的试验数据存储介质。

2) 支撑层。

平台的支撑层包括 SpringCloud 微服务框架、Kubernetes 容器管理平台、数据存储、数据传输和数据统一服务接口。微服务框架提供功能模块微服务化拆解耦、微服务注册与发现、调用、路由、鉴权、限流、降级和进程通信等功能；数据存储提供关系数据库、NoSQL 数据库和分布式文件系统，支持各类型装备试验数据的存储；数据传输提供在线数据传输、离线数据传输，满足软件多级分布式部署数据传输的需求。

3) 数据采集整编。

数据采集整编包括数据采集和数据整编 2 部分，数据采集主要提供试验文书采集、电子文件采

集及结构化数据抽取功能；数据整编主要提供纸质数据图像内容识别、结构化数据规范格式和非结构化数据特征提取等功能。

4) 数据资源。

数据资源包括原始数据库和整编数据库，原始数据库是经过数据采集模块采集后统一存储形成的原始库；整编数据库是原始库中存储的各类数据经过数据整编模块处理后形成的高可用的数据资源。

5) 采集整编工作管理。

采集整编工作管理是装备试验数据采集整编工作流程管理相关的内容，用于支撑各级试验鉴定部门开展装备试验数据采集整编工作，主要提供数据采集整编任务管理、采集活动信息管理、整编活动信息管理及数据采集整编态势展现功能。

6) 数据资源库管理。

数据资源库管理主要提供数据资源库配置、系统资源分配、备份恢复及迁移、数据检索、数据资源库目录管理和元数据管理等功能。

7) 运维管理。

运维管理提供用户管理、角色及权限管理、基础运行数据管理、系统配置和系统日志管理等功能。

8) 统一门户。

统一门户是装备试验数据采集整编平台的统一入口和展现层,为各级平台用户提供统一工作界面。

2.2 业务流程

平台业务流程如图 2 所示,数据采集整编任务管理模块向数据采集整编模块发送任务,由其执行具体数据采集整编任务。在数据采集模块中,经过试验文书采集、电子文件采集、结构化数据抽取后,生成原始数据入到原始数据库中。在数据整编模块中,接收原始数据库中的数据,进行图像内容识别、非结构化数据特征提取、结构化数据规范格式等,生成整编结果数据,并存入到整编数据库中。数据采集和数据整编模块分别向采集活动信息管理和整编活动信息管理模块发送采集/整编活动信息,并发送给数据采集整编态势展现模块。

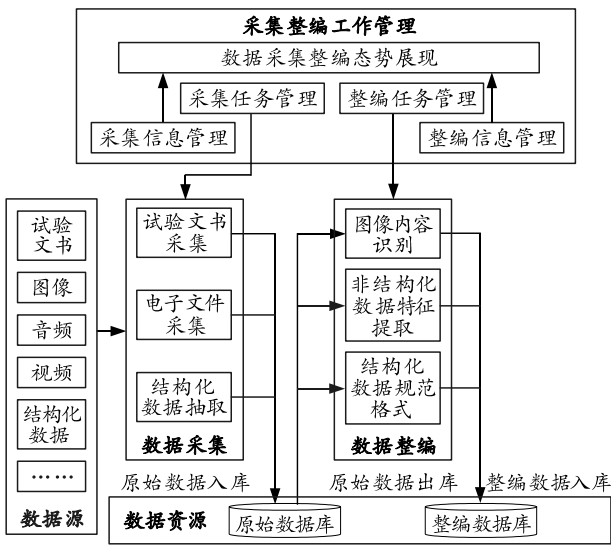


图 2 采集整编业务流程

3 关键技术实现

3.1 微服务架构

装备试验数据管理平台具备多个不同的应用,通过基于微服务架构的设计开发对不同的应用进行有效拆分,不同的业务服务彼此之间功能独立,每个微服务对外提供 Restful API 接口,微服务之间通过服务网关统一进行接口调用。基于容器技术封装各个微服务,能够单独运行部署,从而将应用和服务拆解为小型、松耦合的组件,便于平台进行功能升级和业务扩展^[11]。平台的微服务架构如图 3 所示。

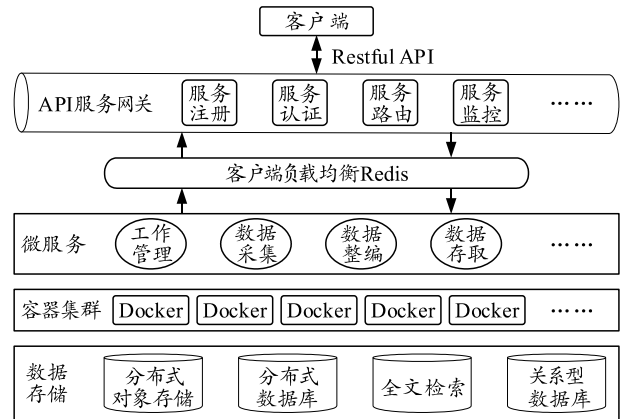


图 3 平台微服务架构

平台的 API 网关统一对外提供 Restful API, Redis 实现均衡负载,使用 Docker 承载服务端的微服务,并通过 Kubernetes 调度容器集群,从而使得任务下发、数据采集、数据整编以及态势展现服务都可以独立的部署、运行、升级。不同的微服务都有对应的数据存储空间,不同的微服务在结构上实现“松耦合”,但在功能上则体现统一,具体表现为界面风格统一、权限管理统一、安全策略统一、上线过程统一、调度方式统一、访问入口统一等^[12]。

3.2 文字智能识别

试验文书的数字化是装备试验数据管理的重要一环。由于试验文书形态多样,并存在诸如弯曲、倾斜、垂直等各种角度和形状,传统 OCR 算法无法很好地检测出文字区域以及识别文字。基于深度学习的文字识别相比于传统算法能够更好地识别复杂文本。平台通过基于 DBNet 的文本检测算法^[13]和基于 CRNN 的文字识别算法^[14]组合完成对文本数据的智能识别,实现纸质试验文书的电子化入库。

3.2.1 DBNet 文本检测算法

DBNet 文本检测算法是基于图像语义分割的文本检测方法,通过一个可微分二值化函数,期望将二值化操作融入分割网络中进行联合参数优化,形成一个自适应学习二值化阈值的过程,进而简化后处理过程,加快处理速度。DBNet 算法整体网络结构如图 4 所示。

DBNet 网络结构可分为 3 部分:1) 特征提取模块,由 FPN 结构构成,输入的图像通过卷积层得到 5 个特征图,分别为原图的 1/2, 1/4, 1/8, 1/16, 1/32 大小,将特征图进行上采样处理到原图的 1/4 大小,再合并拼接得到特征图 F,其大小为原图的 1/4;2) 预测模块,即特征图 F 生成概率图 P 和阈

值图 T 的过程；3) 可微二值化模块，将概率图 P 和阈值图 T 经过 Differentiable binarization，得到近似二值图。最后将目标区域利用二值化结果框选出来。将标准二值化中的阶跃函数进行近似：

$$\hat{B} = 1 / (1 + e^{-k(P_{i,j} - T_{i,j})}) \quad (1)$$

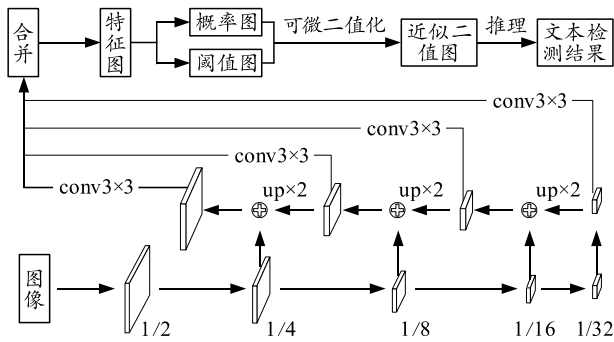


图 4 DBNet 网络结构

可微二值化函数其本质上是一个添加膨胀因子 k (经验型设置为 50) 的 sigmoid 函数，取值范围为 $(0,1)$ ， $P_{i,j}$ 为概率图像素点， $T_{i,j}$ 为阈值图像素点。

3.2.2 CRNN 文本识别算法

CRNN 是一种基于卷积神经网络 (convolutional neural network, CNN)、双向长短期记忆人工神经网络 (bi-directional long-short term memory, BiLSTM)、时间序列算法 (connectionist temporal classification, CTC) 的模型结构^[15]。CRNN 文本识别算法将文字的切割融入深度学习中，可在不需要手工提取特征以及一些预处理步骤的情况下，用于图像中序列文本的识别。CRNN 网络结构如图 5 所示。

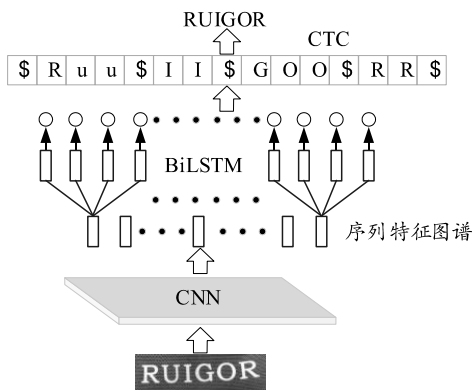


图 5 CRNN 网络结构

基于 CRNN 的文本识别主要由 3 层网络结构组成，包括 CNN、双向 LSTM 和 CTC，识别流程为：

- 1) 对经过 DBNet 网络后输出的图片进行读取，通过卷积神经网络提取图像特征；
- 2) 待识别的文本序列往往都是长短不一的，针

对不定长文本序列，通过双向 LSTM 网络来进行不定长文本序列预测；

- 3) 针对样本对齐困难问题，引入 CTC 模型，解决样本分割对齐问题。

4 试验数据管理平台原型效果展示

采集整编任务下发功能实现如图 6 所示。任务制定人员通过选择任务下发单位、指定任务负责人员、设置任务名称以及配置任务起止时间等，实现具体任务的下发。



图 6 任务下发界面

数据采集界面如图 7 所示，平台支持对非结构化数据、结构化数据以及专用二进制文件的采集。



图 7 数据采集界面

数据整编界面如图 8 所示，以试验文书的整编为例，采集的数据入原始库后，点击内容识别，可智能识别纸质试验文书中的文字内容。



图 8 数据整编界面

5 结束语

笔者针对目前装备试验数据管理存在的采集过程不透明、整编过程不智能、数据格式不规范、态势展现不全面等问题，以 SpringCloud 微服务和 Kubernetes 容器调度为基础，融合 DBNet+CRNN 的智能文字识别算法，设计可应用于装备型号论证、研制生产、试验鉴定以及运用保障全生命周期过程的装备试验数据管理平台。通过试验数据采集整编任务的分发、活动信息的上传、试验数据的采集与整编以及采集整编过程的态势展现，实现试验数据的集中可控管理，构建了装备试验数据管理流程的完整闭环，形成了标准的原始库和规范高可用的整编库，有利于精细化管理装备试验数据，对试验数据挖掘、关联分析、深度学习等数据工程应用也能发挥高效的数据支撑作用。

参考文献：

[1] 朱学锋. 装备试验数据资源目录体系构建[J]. 信息系统工程, 2016(5): 111-113.
 [2] 李赫才. 装甲装备在役考核数据采集研究[J]. 兵工自动化, 2022, 41(2): 14-17.
 [3] 姚鹏飞. 装备试验大数据应用架构研究[J]. 舰船电子工程, 2019, 39(1): 10-13.
 [4] 虞业沛. 卫星装备试验鉴定数据质量评价技术及实现[J]. 计算机测量与控制, 2021, 29(8): 233-237.

(上接第 13 页)

[18] PANG K, XIONG Q. Semantic Modeling Framework for Mission-Oriented Military Systems and Combat Cloud Control[C]//2021 IEEE 7th International Conference on Control Science and Systems Engineering (ICCSSE). IEEE, 2021: 200-204.
 [19] BENZEKKI K, FERGOUGUI A E, Elalaoui A E. Software-defined networking (SDN): a survey[J]. Security & Communication Networks, 2016, 9(18): 5803-5833.

(上接第 16 页)

从智能化的设计角度来说，近年来各个领域广泛地应用了智能设计，这也是枪弹未来发展的大趋势。笔者主要结合企业实例对普通弹头自动装配一体化技术在企业中的应用方法进行探究，希望能对相关人士提供帮助。

参考文献：

[1] 刘正. 基于 UG 的枪弹弹头一体化设计技术研究[D].

[5] 李永哲. 靶场开展武器系统作战试验鉴定的思考[J]. 兵工自动化, 2020, 39(1): 28-31.
 [6] 杨成伟. 基于区块链的装备质量信息管理[J]. 兵工自动化, 2022, 41(4): 23-27.
 [7] 龚昕. 基于 Newtera TDM 平台的装备管理系统设计[J]. 信息技术, 2011, 35(8): 186-188.
 [8] 冯兴. 试验设备调度及试验数据管理系统设计与实现[D]. 成都: 西南交通大学, 2019.
 [9] 翟永超. Spring Cloud 微服务实战[M]. 北京: 电子工业出版社, 2017: 51-67.
 [10] 张磊. 深入剖析 Kubernetes[M]. 北京: 人民邮电出版社, 2021: 253-304.
 [11] 李文俊, 杨学强, 杜家兴. 基于云计算的装备保障信息系统集成[J]. 计算机集成制造系统, 2021, 27(7): 1941-1950.
 [12] 王铭. 一种基于微服务的分布式论坛系统: 中国, 201911291742.3[P]. 2020-05-15.
 [13] LIAO M, WAN Y, YAO C, et al. Real-Time Scene Text Detection with Differentiable Binarization[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 11474-11481.
 [14] SHI B, BAI X, YAO C. An End-to-End Trainable Neural Network for Image-Based Sequence Recognition and Its Application to Scene Text Recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39(11): 2298-2304.
 [15] 程博. 基于深度学习的供应商资质证书智能识别技术研究[J]. 数字通信世界, 2021(4): 79-82, 133.
 [20] LI Q, XIONG Q. Application Description Based Task Composition and Reasoning in WSN[C]//2012 8th International Conference on Wireless Communications, Networking and Mobile Computing. IEEE, 2012.
 [21] HORROCKS, IAN, PATEL-SCHNEIDER, et al. SWRL: A Semantic Web rule language combining OWL and RuleML[R/OL]. <http://www.daml.org/2003/11/swrl/rules-all.html>, 2004.
 [22] ALEXANDER C. 解放军东沙岛夺岛作战兵棋推演及其分析[EB/OL]. <https://zhuanlan.zhihu.com/p/163521290>. 太原: 中北大学, 2018.
 [2] 余海勇, 曾云翔, 史慧芳, 等. 机枪弹弹头连续成型技术[J]. 四川兵工学报, 2011, 32(3): 7-12.
 [3] 段红, 虞波, 陶光成. WJ20426—2016《枪弹自动装配生产线设计要求》标准解析[J]. 中国管理信息化, 2017, 20(21): 96-100.
 [4] 范庆辉, 杨志清, 李作武, 等. 弹头自动装配机控制系统的设计与应用[J]. 兵工自动化, 2019, 38(8): 78-81.
 [5] 金翰林, 李锦, 谷岩波. 某枪弹装配系统的控制系统设计[J]. 兵工自动化, 2020, 39(6): 27-29.