doi:10.6041/j.issn.1000-1298.2024.S2.008

基于LOF的联合收获机制造质量检测与分级系统研究

黄胜操[1] 赵军杰[2] 李茂林[1] 倪昕东[1] 毛 旭[1] 陈 度[1,3]

（1. 中国农业大学工学院，北京 100083；2. 青海大学机械工程学院，西宁 810016；
3. 智能农业动力装备全国重点实验室，北京 100083）

摘要： 随着制造业对于产品质量的要求越来越高，机器学习技术在制造质量控制中的应用开始受到关注。针对联合收获机制造质量检测过程自动化和集成化程度较低、缺乏定量评价手段等问题，设计开发了一套联合收获机制造质量终检系统，在此基础上提出了“终检系统+二次分级”的制造质量混合检测方法，通过终检软件排查合格区间以外的异常数据，筛选劣质产品；通过分级模型对合格产品进行二次检测，标记质量隐患。在整合和分析联合收获机制造质量检测需求的基础上提出了检测流程并通过Visual Components数字车间仿真平台对总体方案进行仿真和测试。根据实际需求和检测功能开发了基于LabVIEW平台的联合收获机终检系统软件，并设计了人机交互界面。试验结果表明系统可以满足各项检测需求并实现产品质量检测功能，初步验证了系统可行性。结合使用场景选用局部异常因子（Local outlier factor，LOF）作为二次分级算法，根据异常检测原理将其集成到检测流程中，并建立了制造质量检测与分级算法架构，依据处理结果将初筛合格的产品二次分类并标记为“good”和“tracked”，进而完善制造过程质量检测-评价体系。训练结果表明LOF可以在差异性不显著的数据集中识别异常样本，性能验证过程中该方法可以准确识别并标记测试数据集中的“tracked”样本，且与四分位图的分布一致，进一步验证了该混合检测方法的有效性。本研究开发的联合收获机制造质量检测系统和提出的分级方法具有应用价值，将数字车间架构与机器学习方法应用于农机装备产品制造质量检测，为复杂农机装备制造质量控制提供了解决思路和方法。

关键词： 联合收获机；制造质量；数字车间；异常检测；局部异常因子

中图分类号： S225.3；S3 **文献标识码：** A **文章编号：** 1000-1298(2024)S2-0075-10

LOF-based Combine Harvester Manufacturing Quality Detection and Grading System

HUANG Shengcao[1] ZHAO Junjie[2] LI Maolin[1] NI Xindong[1] MAO Xu[1] CHEN Du[1,3]

(1. *College of Engineering*, *China Agricultural University*, *Beijing* 100083, *China*
2. *School of Mechanical Engineering*, *Qinghai University*, *Xining* 810016, *China*
3. *State Key Laboratory of Intelligent Agricultural Power Equipment*, *Beijing* 100083, *China*)

Abstract: With the increasing demand for product quality in the manufacturing industry, the application of machine learning (ML) technology in manufacturing quality control has been under attention. To address the low automation and integration, as well as the lack of quantitative evaluation methods in the manufacturing quality inspection for combine harvester, a combine harvester manufacturing quality end-of-line inspection system was designed and developed. Based on this system, an "end-of-line inspection + secondary grading" manufacturing quality hybrid inspection method was proposed, which used the inspection software to screen out abnormal products outside the qualified range and select superior and inferior products. The secondary grading model performed a secondary inspection on qualified products and marks hidden problems. Firstly, based on the integration and analysis of the combine harvester manufacturing quality inspection requirements, the detection flow was designed. The overall design of the system was tested and simulated by using the Visual Components digital workshop platform. The LabVIEW-based end-of-line inspection software was developed according to the actual requirements and detection functions, and corresponding user-friendly human-machine interfaces were designed. The results of the end-of-line workshop inspection

收稿日期：2024-07-30 修回日期：2024-09-25

基金项目： 农机研发制造推广应用一体化试点项目(69194014)

作者简介： 黄胜操(1989—)，男，博士生，主要从事农机装备智能测控技术研究，E-mail: huangshengcao@chiaic.com

通信作者： 陈度(1981—)，男，教授级高级工程师，主要从事智能农机装备和计算机测控技术研究，E-mail: tchendu@cau.edu.cn

tests showed that the system can meet various inspection requirements and achieve software functions, preliminarily verifying the feasibility of the system. Secondly, local outlier factor (LOF) was selected as the secondary grading algorithm according to the scenario, and it was integrated into the detection flow based on its anomaly detection principle. Then, a manufacturing quality inspection and grading framework was established, and the grading process classified the initially screened qualified products into "good" and "tracked" groups based on the processing results, thereby improving the manufacturing quality inspection and evaluation system. The training results indicated that LOF-based method can identify anomalous samples in the dataset with insignificant differences. In the performance validation process, this method accurately identified the four "tracked" samples in the testing dataset, which was consistent with the distribution of the quartile plots, further validating the effectiveness of this hybrid detection method. The developed end-of-line inspection system for the manufacturing quality of combine harvesters and the proposed grading method had important practical application value, promoting the application of digital workshop concept and ML on agricultural machinery, and providing solutions and methods for agricultural machinery manufacturing quality control.

Key words: combine harvester; manufacturing quality; digital workshop; anomaly detection; local outlier factor

0 引言

检测与测试是制造过程不可或缺的环节,可以帮助控制制造质量、降低制造成本、减少资源浪费、筛查产品缺陷,也是保证制造过程标准化、产品一致性和制造质量的常用方法[1-2]。工业4.0推动制造过程向更快更高效的方向发展[3],农机装备制造过程也正向着智能制造转型。然而,农机装备制造过程的在用检测与测试方法自动化、集成化程度较低,大部分流程主要依靠人力实现;同时缺少客观的量化评价方法,进而成为制约农机装备产品市场竞争力的重要因素。

联合收获机是结构复杂的农机装备,其制造过程属于典型的离散制造过程,具有物料品类多样、生产环节繁杂、过程控制多变等特点;而且其作业环境恶劣,制造质量将影响工作性能和无故障时间,加强制造质量检测尤为重要。传感器技术[5]、图像处理技术[6]和精密测量[7]技术手段的发展为高效准确的制造质量检测提供了解决方案。农机装备检测与测试的研究中,针对整机性能开展的传动性能测试[8]和结构耐久性测试[9-10]均取得一定成效,针对零部件开展的脱粒滚筒和螺旋输送器焊接质量在线检测[11-12]、发动机安装误差在线检测[13]等也为联合收获机制造质量检测提供技术支撑。但针对整机制造质量尚未形成较为成熟的检测和评价体系,缺少成套的自动化检测系统,亟需开展集成化联合收获机制造质量检测方法研究,完善制造-检测体系。

在用的联合收获机制造质量检测过程大多通过核对合格区间评价产品质量,但仅通过自动化检测设备和系统获得的数据可能并不可靠[14]:尽管合格区间对检测结果做出限制,但仍可能存在部分隐性离散数据,即质量隐患。作为一种用于开发算法和统计模型的先进技术,机器学习[15]可以在参数间相关性不明确或没有相关性的条件下实现异常检测,也可以根据历史数据提取数据特征,发现隐性问题。机器学习用于异常检测包括监督学习和无监督学习:监督学习是基于标签来训练预测模型;而无监督学习是通过对相似数据进行聚类来揭示数据集特征。对于制造质量这类多变量数据集而言,在不明确其分布的条件下,无监督学习是一种更可行的方法。异常检测的应用研究中,LOF在变压器多运行参数异常检测[16]、锂电池的储能效果诊断[17]和多指标评价[18]中都有优异的表现。所以本文将基于LOF开展联合收获机制造质量检测与评价方法研究。

本文针对联合收获机制造质量检测与评价过程存在技术空白和检测体系不健全等问题,在分析制造质量检测需求的基础上,开发一套集成化制造质量终检系统,依据本文提出的检测流程对整机性能进行系统化检测,筛选劣质产品;在此基础上开展基于LOF的制造质量评价方法研究,辨识质量隐患,健全制造-检测体系;最后通过实际车间测试对系统可行性和有效性进行验证。

1 联合收获机终检系统开发

1.1 检测项目及检测流程

联合收获机制造质量检测项目选定依据主要包括:①国家及相关行业对联合收获机制造质量提出的规定和要求[19-24]。②联合收获机制造企业依据产品性能和特点选定的技术指标和规格参数。结合以上两方面相关规范和企业具体需求,本文提出的联合收获机制造质量终检系统将从作业安全、环保指标和电液系统性能3方面进行质量检测和评价(表1)。

根据选定的检测项目,以“检测流程有序、合理高效利用空间”的原则对终检工位进行布局规划,设计了多

表1　整机检测项目

Tab. 1　Inspection items

检测类别	检测项目	检测参数	检测方法/仪器设备
作业安全	转向轮定位	前束值、外倾角	四轮定位仪
	制动性能	制动力、脚踏板力、阻滞力、轴重	反力式滚筒试验台
	灯光	光照强度	灯光检测仪
环保指标	排放	光吸收系数	不透光烟度计
	噪声	声强	声级计
电液系统	电气系统	用电器工况、起动性能、发动机参数及故障	电流传感器、CAN诊断仪等
	液压系统	工作压力、油箱温度	压力传感器、温度传感器

功能单元线式检测系统。多功能单元线式检测系统将各检测项目集成于单个车间工位中，方案设计源于单元线（Cell line）[25]生产模式，具有占用空间小、建线周期短、检测人员投入少等优点。

检测系统需要根据检测项目提供量化的检测结果，检测人员可以利用这些信息辨识故障位置，指导产品检修。在每台待测联合收获机的检测过程中，首先应在位置开关和检测人员的协同帮助下将联合收获机行驶并停放至检测位置；然后检测人员固定传感器或相应地调整仪器，为提高效率，驾驶员及检测人员的操作将显示在操作指示屏幕上；完成检测准备后，检测人员将操作检测软件进行数据采集和结果分析。联合收获机终检系统检测流程如图1所示。

1.2　车间建模及仿真

传统车间设计基于设计经验和平面图纸等描述车间布局和工艺流程，方案设计抽象、优化周期长且不便于数据管理，而工业4.0背景下的数字车间[26]概念使车间设计可视化、工艺流程数字化，为高效、灵活、准确的车间设计提供解决方案。本文选用Visual Components数字车间仿真平台搭建终检系统仿真环境，以组件形式将车间各物理元素的三维模型导入仿真环境，综合考虑现场空间、元素间干涉碰撞、布局合理性、操作流程顺畅性等问题，建立了单元线式终检系统的数字车间（图2）。基于该数字车间环境可以从元素间碰撞干涉、车间布局设计、检测流程等方面对整体方案进行评估和优化。

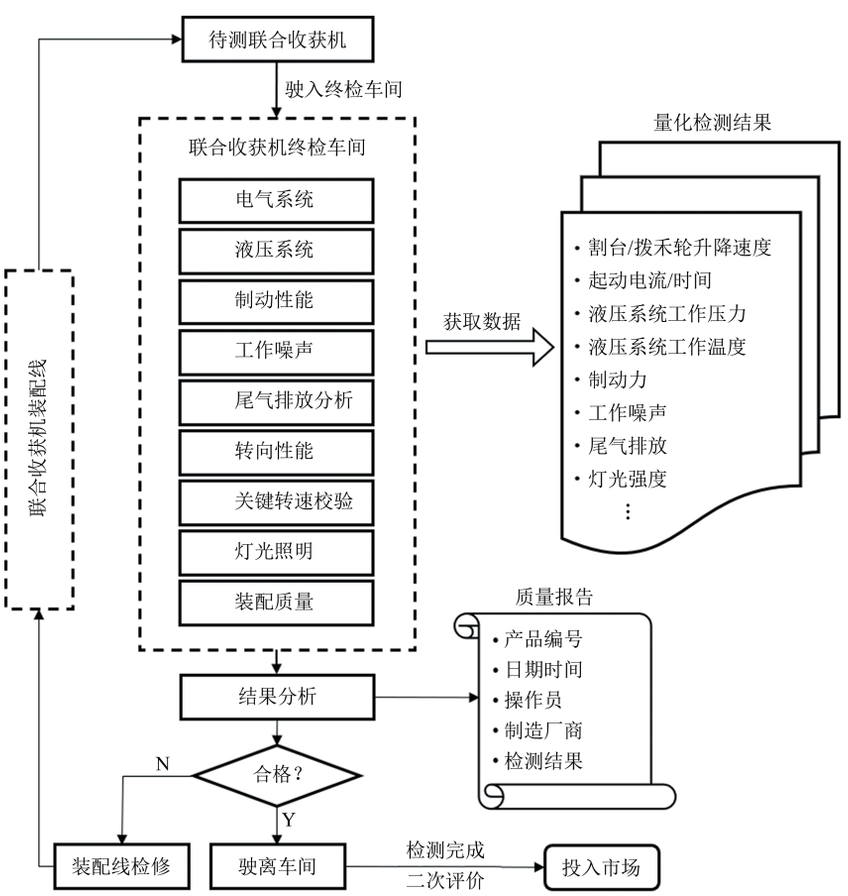


图1　终检系统检测流程图

Fig. 1　Test flowchart



图2 终检系统数字车间

Fig. 2 Digital workshop of end-of-line test system

为验证终检系统方案的可行性并进一步优化设计，基于搭建的数字车间以及Visual Components平台的测试功能开展动态系统仿真和工艺流程测试，从而评估终检系统的工作性能。仪器设备利用率、工作强度和检测效率是评估终检系统性能的重要指标：仪器设备利用率是检测设备在单位时间内的利用率，评估设备的使用效率和车间资源利用情况；工作强度是检测人员在检测过程中所承受的工作负荷，包括心率、移动步数等，评估工作条件和工作效率；检测效率是终检系统在单位时间内的检测能力，包括检测数量、效率等。终检系统仿真结果如表2所示。

表2 检测系统性能仿真结果

Tab. 2 System simulation results

参数	单台检测耗时/min	单位时间检测数量/台	仪器设备使用效率/%	工人平均步数/步	最高心率/(次·min^{-1})	平均工作效率/%
取值	8.38	7	15.9	1 343.7	110	6.46

其中，系统检测效率主要评估单位时间内(1 h)完成检测数量和单台检测时间，以第一台联合收获机驶入工位作为开始并以其驶离工位作为完成检测的标志。仿真结果显示单台联合收获机完成检测共耗时8.38 min，系统单位时间内共有8台联合收获机驶入工位，其中7台完成检测任务。仪器设备利用率以系统中自动化程度较高的灯光检测仪为例，当灯光检测仪的感光探头产生位移即判断已进入工作状态。单位时间内仿真结果显示灯光检测仪总工作时间为9.53 min，工作效率为15.9%。仿真过程中为车间配备3名检测人员，根据检测需求和车间布局为他们布置检测任务，其中：1号负责转向轮定位、灯光、排放和噪声检测；2号负责制动、液压及电气系统检测；3号负责协调驾驶员和系统操作。3人工作步数分别为1 652、2 093、286步，平均1 343.7步；最高心率分别为110、110、109次/min，最高心率110次/min；单位时间内工作效率分别为8%、7.2%和4.2%，平均工作效率为6.46%。

1.3 车间及现场总线

前文对终检系统检测工艺的仿真表明，各项检测功能均能根据实际需求和相关标准顺利实现。进而，基于以上工作创制了联合收获机制造质量终检车间，如图3所示。

在开发的终检车间基础上，部署了车间现场总线，如图4所示。现场总线是一种现代工业通信网络，用于连接车间中的设备和仪器，并在它们之间实现通讯和信息传递。车间中的设备仪器包括制动测试台、四轮定位仪、烟度计、灯光检测仪和各类传感器，它们在实现采集和控制功能的同时需要与控制终端交换数据。为了满足这些要求，选择了Modbus、CAN总线或TCP/IP协议实现现场总线通信，并将它们集成到终端中，实现车间内实时数据传输和控制。在部署现场总线时，根据不同的检测项目和功能需求设计了独立的人机交互功能，整合



图3 联合收获机制造质量终检车间

Fig. 3 Combine harvester end-of-line test workshop

控制、采集和通讯功能。同时提供直观的数据可视化，帮助用户更好地理解和管理车间设备和数据，提高管理效率和测试质量。

1.4 软件开发及系统测试

虚拟仪器是一种基于计算机技术和硬件支持的检测和测试工具。它具有灵活性、可重复性、自动化程度高、数据处理和分析能力强等特点。

虚拟仪器的检测功能可以灵活地调整和扩展，能够同时满足不同检测需求，具有较强的适应性。此外，虚拟仪器的检测结果具有较高的可重复性，可以准确地复现检测条件和数据结果。检测流程可以通过程序自动完成，减少人工干预，提高检测效率和准确性。同时，虚拟仪器提供的显示和操作界面直观且易于操作。所以，本文选用LabVIEW虚拟仪器平台作为终检系统软件开发平台。

基于LabVIEW的开发架构由3部分组成(图5)：最底层为硬件驱动层，主要通过提供硬件设备通讯接口，实现数据采集、控制设备等功能；中间层为虚拟仪器层，通过构建虚拟仪器模拟传统硬件仪器的功能，包括数据采集、信号处理、控制和可视化等，同时虚拟仪器层也作为应用程序接口层提供功能函数和测试方法等，用于实现虚拟仪器和硬件设备之间的信息交互；最上层为用户界面层，用于实现人机交互、系统控制和数据可视化等。

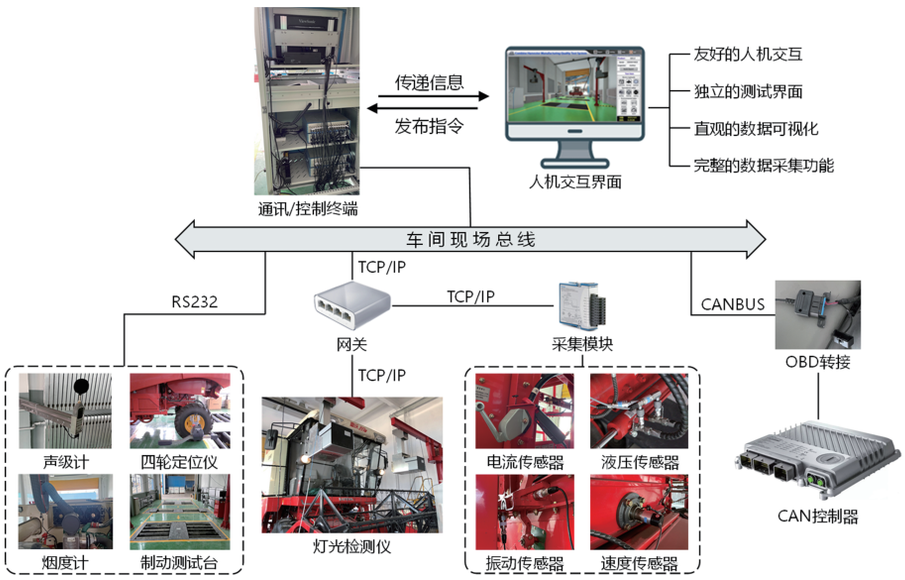


图4　车间现场总线

Fig. 4　Workshop fieldbus

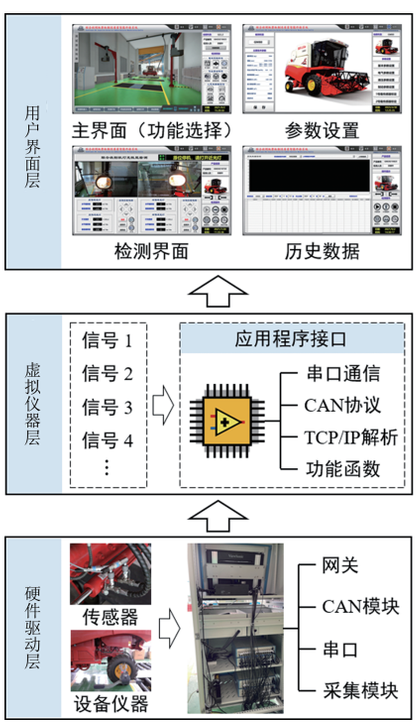


图5　软件架构

Fig. 5　Software structure

此外，为便于产品数据管理，为车间配置了本地数据库及云端数据库以存储检测数据和回调历史数据。为测试联合收获机制造质量终检车间及系统软件的可行性和有效性，对完成装配的10台联合收获机进行了电气系统检测[27]，结果表明系统检测功能运行稳定，实现了联合收获机电气系统快速检测，符合实际生产需求。

2　制造质量检测与分级架构

2.1　数据集预处理

尽管终检系统已经可以实现整机制造质量检测功能，但在系统测试过程中发现尽管合格区间对检测参数作出限制，但少量检测结果存在问题：它们靠近合格区间边界或与大多数结果偏离。所以本文提出猜想：仅靠合格区间评价制造质量可能并不可靠。

基于开发的终检车间，获取70台检测软件判断为“合格”的联合收获机及它们的检测数据，组成70个制造质量数据集。验证上文猜想的本质是聚类[28]，即从70个数据集找到离群样本，根据数据集聚类结果得到离群点特征。聚类是一种典型的无监督学习，是一个把数据对象划分成多个组或多个簇的过程，使簇内的对象高度相似的同时与其他簇中对象极不相似。本文的聚类过程采用常见的高斯混合模型（Gaussian mixture model，GMM）对数据集进行预处理，数据集被分为“good”和“tracked”两类，其中“good”表示产品无异常，“tracked”表示产品包含离群项，可能存在质量隐患。

GMM是使用高斯概率密度函数（正态分布曲线）精确地量化对象，将一个对象分解为若干模型。它假设所有的数据样本都是由一个给定参数的多元高斯分布生成的。其中，给定聚类个数K，对于给定样本x，一个高斯混合模型的概率密度函数可以由K个多元高斯分布组合而成的混合分布表示，即

$$p(x)=\sum_{i=1}^{K}\omega_i p\left(x|\boldsymbol{\mu}_i,\boldsymbol{\Sigma}_i\right) \tag{1}$$

式中，$p\left(x|\boldsymbol{\mu},\boldsymbol{\Sigma}\right)$是以$\boldsymbol{\mu}$为均值向量，以$\boldsymbol{\Sigma}$为协方差矩阵的多元高斯分布的概率密度函数，高斯混合模型则是由K个不同的多元高斯分布共同组成，其中每一个分布被称为高斯混合模型中的一个成分，而ω_i是第i个多元高斯分布在混合模型中的权重，且$\sum_{i=1}^{K}\omega_i=1$。假设已有一个高斯混合模型，样本空间中样本的生成过程为：以ω_1，ω_2，…，ω_K为概率选择出一个混合成分，并根据该混合成分的概率密度函数产生出相应的样本。

利用GMM进行聚类的过程是利用GMM生成数据样本的“逆过程”：给定待聚类数据集的聚类簇数K，以期望-最大化（Expectation-maximization，EM）方法搜索出混合成分参数（均值向量$\boldsymbol{\mu}$、协方差矩阵$\boldsymbol{\Sigma}$和权重$\boldsymbol{\omega}$），每一个多元高斯分布成分即对应聚类后的一个簇。最后采用极大似然估计法将样本点分配到概率最大的簇中：

$$L=\lg\prod_{j=1}^{m}p(x)=\sum_{j=1}^{m}\lg\left(\sum_{i=1}^{K}\omega_i p(x|\mu_i,\Sigma_i)\right) \tag{2}$$

式中 L——极大似然估计

m——样本数量

在预处理过程中，有70个数据集样本（样本数量）；由于数据集分布尚不明确，所以在K为20的条件下进行尝试。图6为一个制动性能检测数据集预处理结果。

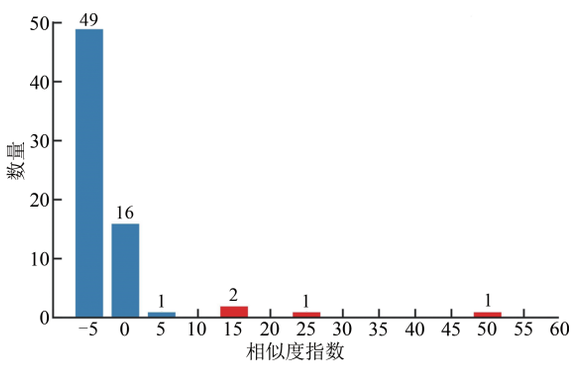


图6 制动数据集的GMM预处理结果

Fig. 6 Preprocess on braking performance dataset using GMM

在图6中，横坐标表示每个样本的相似度指数，即属于K个聚类之一的可能性。横坐标越大表示样本与K个聚类之间的差异性越大。该参数是通过公式（2）计算得出的。纵坐标表示具有相似聚类结果的样本数量。如图6所示，蓝色部分表示有66个样本聚类较紧密，而一些红色样本则相对分散，表明确实存在一些样本与大多数偏离，即它们对应的联合收获机可能存在潜在的制动问题，同时也初步验证了本文提出的猜想。

2.2 LOF算法及原理

LOF是基于密度的经典算法[29]，传统异常检测算法大多基于统计学理论实现，或是借用聚类算法识别离群点：基于统计的异常检测算法需要以数据样本服从特定概率分布为前提，但这种假设通常具有局限性；而聚类方法在输出结果时大多只能给出是或不是的判断，无法量化异常程度。与它们相比，基于密度的LOF算法计算更直接，结果更直观，且对数据集分布没有特殊要求，以量化形式呈现异常程度，所以在制造质量检测场景中适用性更好。

LOF算法在高维空间中可以自适应地检测异常，其基本思想是，LOF认为离群样本周围的样本密度环境与正常样本周围的样本密度环境不同：在稠密的区域中，离群样本周围的局部密度较低；而在稀疏的区域中，离群样本周围的局部密度较高。因此，LOF算法主要通过比较每个样本的局部密度以及它周围样本的局部密度，来计算每个样本的异常程度。其中，异常程度即局部异常因子，它本质上是反映异常程度的评分，也是样本所处位置局部相对密度。

对于样本p，异常检测步骤为：①计算样本p与其他k个样本的距离，并排序。②找到对应的k近邻并计算LOF分数

$$\mathrm{LOF}_k(p)=\frac{\sum_{o\in N_k(p)}\mathrm{lrd}_k(o)}{|N_k(p)|\mathrm{lrd}_k(p)} \tag{3}$$

其中，lrd是样本p的局部可达密度（Local reachability density），即样本p与其他k个样本的平均可达距离的倒数，$N_k(p)$是$k-\mathrm{distance}(p)$点集，称为第k邻域。

通过以上步骤即可计算得到数据集中各样本的局部相对密度。其中，离群样本的局部相对密度将大大低于正常样本的局部相对密度，将所有样本的局部相对密度排序后即可通过其分布情况找到对应的离群样本。

2.3 检测与分级架构

基于1.1节提出的联合收获机制造质量检测流程，并以LOF算法为核心，提出了联合收获机制造质量检测与分级架构，如图7所示。通过计算各数据集中每个样本的局部异常因子，实现对联合收获机的制造质量分级。该架构主要包含以下环节：首先，利用开发的终检车间和检测软件，根据检测流程完成下线检测环节，以此获取联合收获机制造质量数据集并初步筛查质量异常，若合格则进行二次分级，若不合格则退回装配线检修；其次，根据LOF算法的异常检测原理得到各样本的LOF分数，并将样本标记为“good”和“tracked”两类，以此对合格产品进行二次分级，“good”类产品将直接投入市场，而“tracked”类产品在运维管理中将重点监测标记参数。

3 系统测试

为验证本文提出的“终检系统+二次分级”混合检测方法的可行性和有效性，首先基于已获取的70个电气系统、液压系统和制动性能数据集对LOF分级方法进行训练，借助聚类算法的簇内致密度和簇间离散度评价LOF算法的性能，再通过离散样本击中数量验证LOF分级方法的准确性；最后将分级方法与终检系统结合并部署在终检车间，通过车间测试验证系统性能和混合检测方法。

3.1 数据集重构

为了初步提取各检测项目中的离散样本并定位对应的产品编号，将70个联合收获机电气系统、液压系统和制动性能数据集导入并绘制四分位图（箱型图），如图8所示。四分位图根据样本分布将数据集分为3部分，包括上触须、上四分位线、下四分位线和下触须。

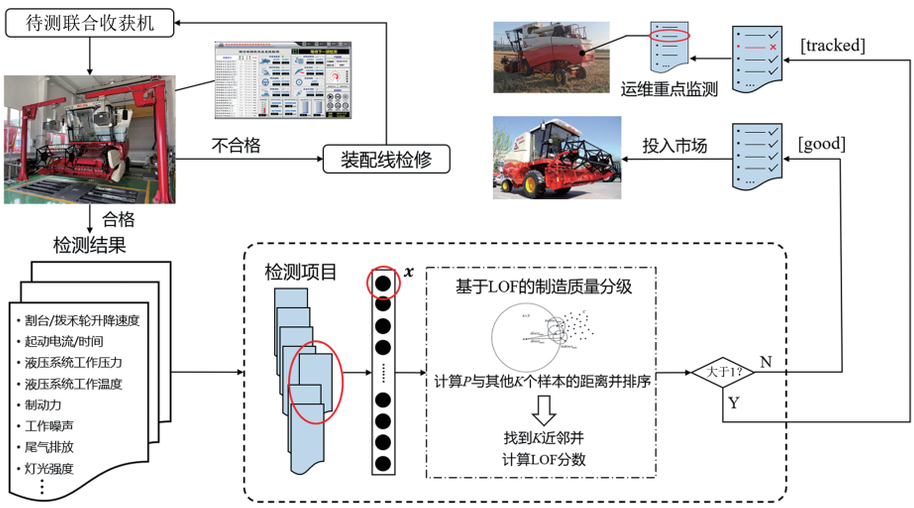


图7　基于LOF的检测与分级架构

Fig. 7　LOF-based detection and grading structure

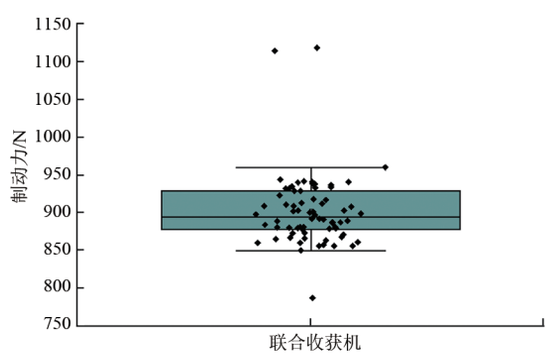


图8　右前轮制动力四分位图

Fig. 8　Quartile plot of front right wheel's braking force

这些四分位图将可视化离散样本，进而帮助标记和提取它们。统计标记结果后发现，电气系统数据集中包含8个离散样本；液压系统数据集中包含4个离散样本；制动性能数据集中包含11个离散样本。将这些离散样本从数据集中提取出来，并从中随机抽取若干个重新放入数据集并置于末尾，进而重新构建数据集并作为训练数据集。其中，电气系统数据集包含67个样本，其中有5个离散样本（63号~67号）；液压系统数据集包含68个样本，其中有2个离散样本（67号、68号）；制动性能数据集包含66个样本，其中有7个离散样本（60号~66号）。

3.2　训练及性能验证

将重构数据集作为LOF算法的输入，进行5次300轮训练和计算，本文训练及计算环境是：Intel（R）Core（TM）i7-10700KF CPU @ 3.80GHz，10 GB NVIDIA GeForce RTX 3080 GPU和32 GB运行内存。

3.2.1　评价指标

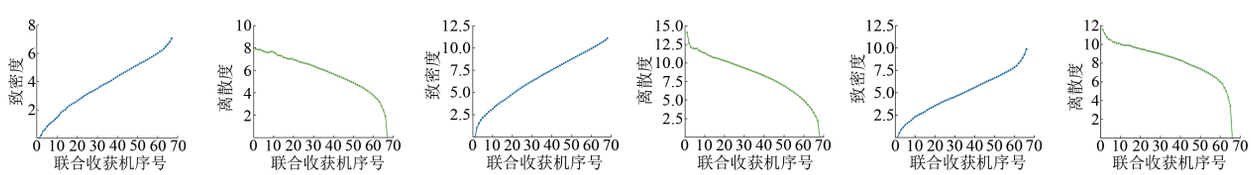
对于本文开展的联合收获机制造质量分级方法研究而言，其本质是在终检系统初步筛选存在明显异常及故障的基础上，再次对初筛合格的联合收获机进行分级，找到在合格范围内但与大多数合格产品偏离的离群项，属于无监督聚类问题：其目标是保证在同一个“簇”内的样本距离尽量小，而不同“簇”之间的样本距离尽量大，因此簇内致密度和簇间离散度是评价检测及聚类效果的重要指标。

（1）致密度（Compactness，CP）的评价通常采用均方误差和（Sum of squared errors，SSE），对于联合收获机制造质量分级而言数据集将被聚类为两个簇X和Y，聚类中心分别为X_0和Y_0，两个簇中所有的样本X_i和Y_i与各自簇中心的距离之和即为致密度的衡量指标。

（2）离散度（Separation，SP）是用来评价聚类效果的参数，可以量化聚类显著性。联合收获机制造质量分级中的离散度即“good”和“tracked”两个聚类簇间总距离，其计算方式是将“good”中的每一个样本与“tracked”中的每一个样本的间距求和。

3.2.2　样本异常度分析

基于上述评价指标，使用LOF分别对重构的电气系统、液压系统和制动性能数据集加以处理，根据数据集中各样本到样本中心的距离，将其从高到低排序后计算对应的CP和SP值，结果如图9所示。



(a) 电气系统数据集　(b) 液压系统数据集　(c) 制动性能数据集

图9　CP和SP曲线

Fig. 9　CP and SP curves

由图9可得样本与样本中心距离越大,对应的CP值越小,SP值越大。CP曲线反映聚类后簇内紧密程度:CP值越小,样本属于当前类的可能性越大;SP曲线反映聚类后簇间分离程度:SP值越大,样本属于其他类的可能性越小。随着样本与聚类中心的距离减小,簇内致密度越来越大,说明聚类效果越好;距离越大,簇间离散度越来越大,说明算法对两个簇的区分效果越好,以此初步验证了LOF算法对制造质量数据集的聚类性能。

3.2.3 击中数量

基于LOF方法的聚类和分离表现,为进一步验证LOF对离群样本的识别准确性,将它在电气系统、液压系统和制动性能数据集中识别到的离群样本序号和数量与3.1节中的四分位图结果进行比对,以反映对离群样本的辨识准确性和发现异常的能力,即击中数量。选取样本与样本中心距离排序的后20%样本,统计其中对插入的离群样本的击中数量,如表3所示。

结果表明LOF可以准确地识别和标记插入的离群样本,结合上文的数据集聚类表现,说明LOF可以对联合收获机的制造质量进行分级,验证了本文提出的混合检测方法的有效性,为联合收获机制造质量控制提供了新的解决方案。

表3 离群样本击中数量

Tab. 3 Hit amount of outliers

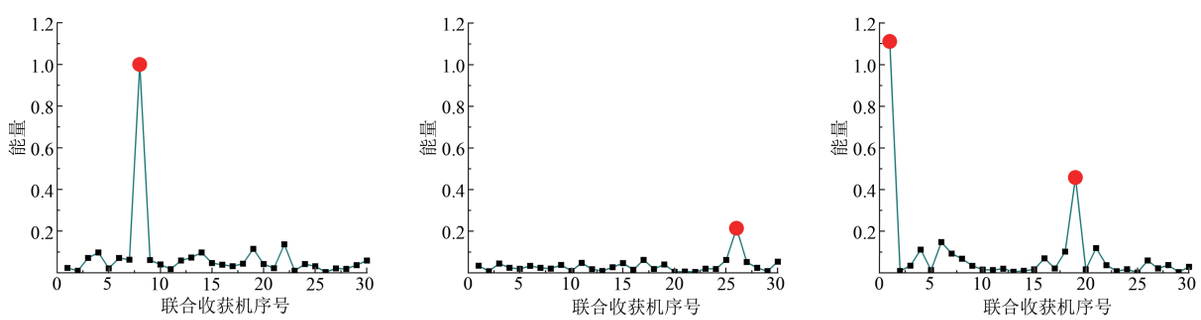
电气系统		液压系统		制动性能	
编号	击中结果	编号	击中结果	编号	击中结果
63	√	67	√	60	√
64	√	68	√	61	√
65	√			62	√
66	√			63	√
67	√			64	√
				65	√
				66	√

3.3 系统验证与讨论

3.3.1 车间部署及验证

将基于LOF的制造质量分级方法与联合收获机制造质量终检系统结合,建立联合收获机制造质量检测与分级系统。为验证系统性能,将其部署于联合收获机制造质量终检车间并开展下线检测测试。测试以潍柴雷沃重工股份有限公司当季完成装配并待测的30台联合收获机为对象,将它们依次驶入联合收获机制造质量终检车间并按照检测流程与分级框架完成检测。

首先经过联合收获机制造质量终检系统初步检测,30台联合收获机的各项检测参数均在合格区间范围内,即它们都满足企业及相关行业标准的基本要求,可以对它们进行制造质量分级。将30台联合收获机的制造质量检测数据集作为测试数据集,作为LOF分级模型的输入。基于70个联合收获机数据集的训练结果对30台联合收获机进行分级,辨识标记离群样本并分级,如图10所示。



(a) 电气系统 (b) 液压系统 (c) 制动性能

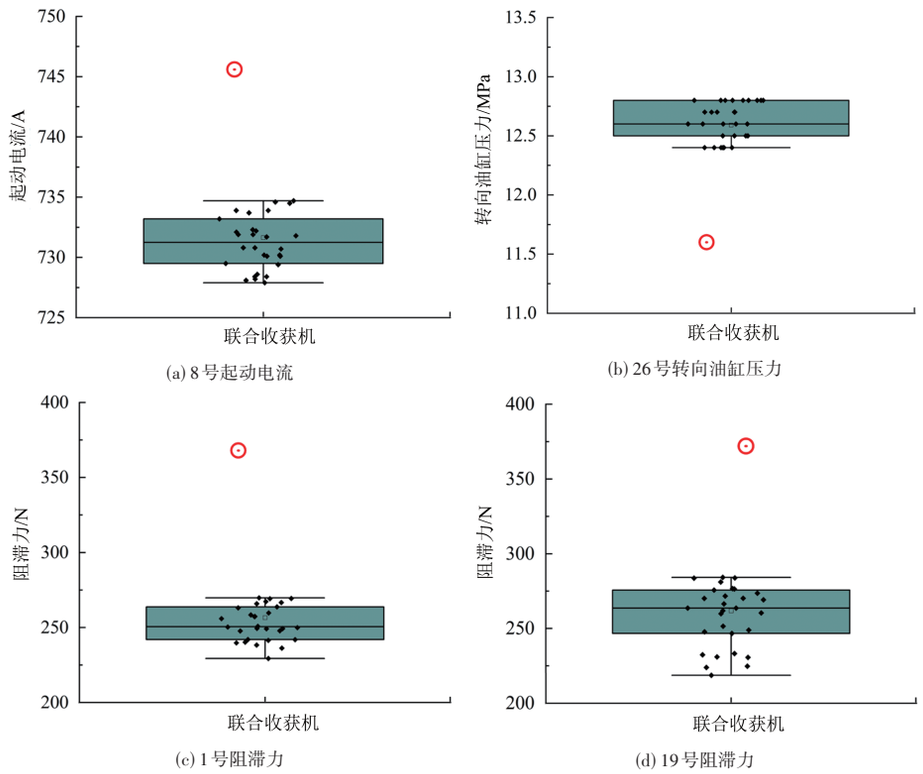
图10 测试集中的离群样本

Fig. 10 Outliers in test dataset

标记结果表明,电气系统测试集中,8号产品为离群样本,定位于起动电流;液压系统测试集中,26号产品为离群样本,定位于转向油缸压力;制动性能测试集中,1号和19号产品为离群样本,定位于1号产品和19产品的车轮阻滞力。为验证对离群样本的辨识准确性,采用四分线法对测试集进行复核(图11),二者结果一致。综上所述,30台联合收获机均可正常售卖,其中:1号、8号、19号和26号产品定级并标记为“tracked”;其余产品定级并标记为“good”。

3.3.2 讨论

机器学习在制造质量控制方面的应用日益普及,而LOF的探索只是一个开始。尽管本文结果表明LOF可以在识别制造质量隐患方面发挥作用,但在其部署过程中也发现了一些改进方向。首先,LOF是基于固定数据集训练的,当新样本加入数据集时,需要重新计算和训练,而随着数据集包含的样本体量越来越大,系统需要具备一定计算能力。将LOF与在线学习方法相结合或建立更稳定的算法模型可能会提高计算和处理效率。此外,另一个改进方向是部署集成方案。目前,该混合检测方法能够识别和标记异常样本,但如果LOF用作预处理步骤来识别异常样本,再使用决策树一类的决策支持工具构建决策模型,将建立一个更完整的“检测-识别-决策”体系,从而进一步完善制造质量控制过程。



(a) 8号起动电流　(b) 26号转向油缸压力　(c) 1号阻滞力　(d) 19号阻滞力

图11　离群样本分布图

Fig. 11　Quartile plots of test dataset with outlier

4　结束语

在整合相关标准和企业实际生产需求的基础上，设计了联合收获机制造质量检测方案，采用数字车间技术对检测流程进行动态仿真和设计优化，基于以上工作开发了联合收获机制造质量终检系统，包括终检车间和检测软件。在此基础上进一步提出了“终检系统+二次分级”的混合检测方法，搭建了检测与分级架构。终检系统初筛劣质产品后对合格产品进行二次分级，辨识标记故障隐患。其中，选用LOF算法作为二次分级的核心，训练及性能验证结果表明LOF具有较好的数据集聚类表现，可以准确地在数据集中找到离群样本，验证了该分级方法的可行性和有效性，同时说明LOF在多参数异常检测场景中，如制造质量检测过程，LOF具有较好的适用性。

参　考　文　献

[1] ÖZCAN A M, AKDOĞAN A, DURAKBASA N M. Improvements in manufacturing processes by measurement and evaluation studies according to the quality management system standard in automotive industry[C]//International Symposium for Production Research 2020, 2020.

[2] PSAROMMATIS F, PROUVOST S, MAY G, et al. Product quality improvement policies in industry 4.0: characteristics, enabling factors, barriers, and evolution toward zero defect manufacturing[J]. Frontiers in Computer Science, 2020, 2: 1 - 15.

[3] KUSIAK A. Smart manufacturing[J]. International Journal of Production Research, 2017, 56(1-2): 508 - 517.

[4] SAHOO S, LO C Y. Smart manufacturing powered by recent technological advancements: a review [J]. Journal of Manufacturing Systems, 2022, 64: 236 - 250.

[5] JAVAID M, HALEEM A, SINGH R P, et al. Significance of sensors for industry 4.0: roles, capabilities, and applications[J]. Sensors International, 2021, 2: 100110.

[6] BABIC M, FARAHANI M A, WUEST T. Image based quality inspection in smart manufacturing systems: a literature review [J]. Procedia CIRP, 2021, 103: 262 - 267.

[7] 周亮，王振环，孙东辰，等. 现代精密测量技术现状及发展[J]. 仪器仪表学报，2017，38(8)：1869 - 1878.
ZHOU Liang, WANG Zhenhuan, SUN Dongchen, et al. Present situation and development of modern precision measurement technology [J]. Chinese Journal of Scientific Instrument, 2017, 38(8): 1869 - 1878. (in Chinese)

[8] TANELLI M, PANZANI G, SAVARESI S M, et al. Transmission control for power-shift agricultural tractors: design and end-of-line automatic tuning[J]. Mechatronics, 2011, 21(1): 285 - 297.

[9] MATTETTI M, MOLARI G, VERTUA A. New methodology for accelerating the four-post testing of tractors using wheel hub displacements[J]. Biosystems Engineering, 2015, 129: 307 - 314.

[10] PARAFOROS D S, GRIEPENTROG H W, VOUGIOUKAS S G. Methodology for designing accelerated structural durability tests on agricultural machinery[J]. Biosystems Engineering, 2016, 149: 24 - 37.

[11] 苑严伟, 方宪法, 荣耀, 等. 联合收割机脱粒滚筒焊接质量在线检测系统[J]. 农业工程学报, 2015, 31(4): 11 - 16.
YUAN Yanwei, FANG Xianfa, RONG Yao, et al. Online diagnosis system of threshing cylinder welding quality of combine harvester[J]. Transactions of the CSAE, 2015, 31(4): 11 - 16. (in Chinese)

[12] 高瑞遥, 王丽丽, 牛康, 等. 基于激光扫描的螺旋输送器焊接质量在线检测[J]. 农业机械学报, 2020, 51(增刊1): 292 - 297.
GAO Ruiyao, WANG Lili, NIU Kang, et al. Online detection of welding quality for screw conveyor of combine harvester based on laser scanning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(Supp.1): 292 - 297. (in Chinese)

[13] 张亚伟, 王冬, 陈度, 等. 大型收获机械发动机孔组位置度误差在线检测方法[J]. 农业机械学报, 2017, 48(增刊): 71 - 78.
ZHANG Yawei, WANG Dong, CHEN Du, et al. Online method for large-scale harvester engine punch combination position accuracy measurement[J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(Supp.): 71 - 78. (in Chinese)

[14] SONG D, OH Y G, KIM N. Study on correlation-based feature selection in an automatic quality inspection system using support vector machine (SVM)[J]. Journal of Korean Institute of Industrial Engineers, 2016, 42(6): 370 - 376.

[15] KAMM S, VEEKATI S S, MÜLLER T, et al. A survey on machine learning based analysis of heterogeneous data in industrial automation [J]. Computers in Industry, 2023, 149: 103930.

[16] ZOU D, XIANG Y, ZHOU T, et al. Outlier detection and data filling based on KNN and LOF for power transformer operation data classification[J]. Energy Reports, 2023, 9: 698 - 711.

[17] QIU Y, DONG T, LIN D, et al. Fault diagnosis for lithium-ion battery energy storage systems based on local outlier factor[J]. Journal of Energy Storage, 2022, 55: 105470.

[18] CHEN Z, XU K, WEI J, et al. Voltage fault detection for lithium-ion battery pack using local outlier factor[J]. Measurement, 2019, 146: 544 - 556.

[19] GB 16151.12—2008 农业机械运行安全技术条件 第12部分:谷物联合收割机[S]. 2008.

[20] GB 36886—2018 非道路移动柴油机械排气烟度限值及测量方法[S]. 2018.

[21] GBT/14248—2008 收获机械 制动性能测定方法[S]. 2008.

[22] GBT/20949—2007 农林轮式拖拉机 照明和灯光信号装置的安装规定[S]. 2007.

[23] JBT/6268—2015 自走式收获机械 噪声测定方法[S]. 2015.

[24] JBT/7316—2015 谷物联合收割机 液压系统试验方法[S]. 2015.

[25] LEE Q. How to balance a manufacturing work cell[C]//Institute of Industrial Engineers - IE Solutions Conference, 2000.

[26] 陶飞, 张萌, 程江峰, 等. 数字孪生车间——一种未来车间运行新模式[J]. 计算机集成制造系统, 2017, 23(1): 1 - 9.
TAO Fei, ZHANG Meng, CHENG Jiangfeng, et al. Digital twin workshop: a new paradigm for future workshop[J]. Computer Integrated Manufacturing Systems, 2017, 23(1): 1 - 9. (in Chinese)

[27] 倪昕东, 陈度, 王书茂, 等. 基于数字车间的联合收获机制造品质终检系统研究[J]. 农业机械学报, 2020, 51(12): 166 - 174.
NI Xindong, CHEN Du, WANG Shumao, et al. End-of-line inspection system of combine harvester manufacturing quality based on digital workshop[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(12): 166 - 174. (in Chinese)

[28] JAIN A K, MURTY M N, FLYNN P J. Data clustering: a review[J]. ACM Computing Surveys, 1999, 31(3): 264 - 323.

[29] BREUNIG M M, KRIEGEL H P, NG R T, et al. LOF: identifying density-based local outliers [C]//The 2000 ACM SIGMOD International Conference on Management of Data, 2000.