

Research
Intelligent Manufacturing—Article

数据驱动的加工过程异常诊断

Y.C. Liang^a, S. Wang^a, W.D. Li^{a,b,*}, X. Lu^a^a Faculty of Engineering, Environment and Computing, Coventry University, Coventry CV1 5FB, UK^b School of Logistics Engineering, Wuhan University of Technology, Wuhan 430070, China

ARTICLE INFO

Article history:

Received 17 July 2018

Revised 18 December 2018

Accepted 14 March 2019

Available online 19 June 2019

关键词

计算机数控加工

异常检测

果蝇优化算法

数据驱动方法

摘要

为了在计算机数控 (CNC) 加工过程中实现零缺陷生产, 开发有效的异常检测诊断系统势在必行。然而, 由于加工过程中机床和工装的动态条件限制, 目前在工业生产中采用的相关诊断系统所能发挥的作用往往非常有限。为了解决这个问题, 本文提出了一种全新的异常数据驱动的诊断系统。在该系统之中, 我们持续收集随动态加工过程而产生的状态监测功率数据, 并以此支持在线诊断分析。为了便于分析, 我们设计了预处理机制对所监视的数据进行去噪、标准化以及校准。随后我们即从监控数据中提取关键特征, 并定义阈值以识别异常。考虑到加工过程中机床和工装的动态条件, 用于识别异常的阈值可以调整。我们还可以基于历史数据利用果蝇优化 (FFO) 算法优化阈值, 以实现更准确的检测。通过实践验证, 我们证明了该系统在工业应用中的有效性和巨大前景。

© 2019 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. 引言

现代制造的特点是高价值、低产量和高客制化, 并且要求零缺陷的生产管理, 以最大限度地减少废料、提高产品质量和生产率。然而, 意外的异常 (如加工工具破损、机器主轴故障或严重的工具磨损) 可能阻碍我们对于零缺陷目标的追求。因此, 开发有效的诊断系统以便有效地检测加工过程中的意外异常至关重要, 这使得我们可以对设备进行适当的调整以解决异常[1,2]。为满足这一需求, 欧盟委员会在制造业中推动“零缺陷制造”这一概念。因此, 为了确定解决方案, 委员会亦资助了一批研究项目[如智能故障纠正和自优化制造系统 (IFaCOM) 项目等]。而从工业角度来看, 人们已经开发出一些诊断系统并将其部署在工

厂中。此类系统中所采用的一种较流行的策略是通过将关键性能指标 (KPI) 与由经验丰富的工程师所预设的静态阈值进行比较来识别异常。然而, 加工过程通常在不同的工作条件下进行, 这导致加工过程中高动态的特性。因此, 基于预设静态阈值的诊断系统无法有效地解决动态的问题。

近年来, 智能传感器和安全物理系统 (CPS) 越来越地被集成到工厂中, 以监控加工设备及工装的动态条件。因此, 数据驱动诊断系统的相关研究也已得到了积极的推进[3–5]。在此类系统中, 人们利用智能与深度学习算法, 通过时域、频域或时域/频域比来从大数据流中挖掘分析其中的异常[6,7]。为了在工业中更加有效地应用此类数据驱动系统, 我们有必要进行进一步的研究以提高数据处理和分析过程中的系统性能。

* Corresponding author.

E-mail address: weidong.li@coventry.ac.uk (W.D. Li).

本文提出了一种全新的用于计算机数控（CNC）加工过程的数据驱动诊断系统。基于该系统，我们可以实现对于加工过程的连续监控和数据收集。随后通过对监控数据的分析以动态检测机器和工装中的异常。该系统的创新特性如下：

(1) 对于所监测的数据设计了去噪、标准化以及校对机制，有助于进行异常分析。

(2) 定义了一组关键特征来表示所监视数据最重要的方面。我们将其与阈值进行比较以识别异常。同时，系统采用果蝇优化（FFO）算法对阈值进行优化，以实现动态加工过程更准确的诊断。

(3) 已通过工业应用实例验证了该系统在实际加工过程中的有效性。

2. 文献综述

过去，基于物理实体和模型的诊断方法是主流方法。而近年来，随着智能传感器、数据分析和深度学习技术的快速进步，人们开发了数据驱动算法以提高诊断的有效性及性能[如玻尔兹曼机（Boltzmann machines）、支持向量机(SVM)、卷积神经网络(CNN)等]。Hu等[8]开发了一种将深度玻尔兹曼机算法与多粒度级联森林算法相结合的工业设备故障挖掘方法。Tian等[9]设计改进了一种支持向量机（SVM）来诊断钢铁厂的故障，他们通过递归特征消除（RFE）算法来减少数据维数，以达到加快计算速度的目的。Zheng等[10]提出了通过复合多尺度模糊熵（CMFE）与集成支持向量机（ESVM）来提取非线性特征并对滚动轴承故障进行分类。然而，所使用的数据之中掺杂了其他冗余的、不相关的特征，这可能大大降低真实的检测率并增加计算时间。Wu和Zhao [11]提出了一种深度卷积神经网络（CNN）模型来检测化学过程的缺陷。然而，深度CNN往往需要较长的计算时间。Madhusudana等[12]开发了一种决策树方法（J48算法）来检测面铣刀的故障情况。通过该方法，人们可利用离散小波变换（DWT）方法从声信号中提取一组离散小波特征。而这项研究的局限在于决策树结构及其阈值难以被定义。Lu等[13]提出了一种双约化核极限学习机方法来诊断航空发动机故障。Wen等[14]提出了一种基于LeNet-5的新CNN；并对该CNN进行了电机轴承测试以及自吸式离心泵和轴向柱塞液压泵故障检测，其精度在99.481%~100%之间。此外，Wen等[15]还提出了一种基于稀疏自动编码器的新型深度迁移学习

模型，并将其应用于电机轴承故障检测，其检测精度达到了99.82%。Wen等[16]还提出了一种新的分层卷积神经网络（HCNN），其精度在96.1%~99.82%之间。表1总结了上述研究工作。

根据García等[17]以及Pan和Yang [18]的调研，在进一步提高数据驱动算法的效率方面存在以下研究空白：

(1) 我们有必要为检测数据设计合理的预处理技术，以确保最佳的诊断准确性及诊断效率。

(2) 深度学习算法通常需要很长的训练时间才能达到高精度。获取足够的错误数据模式用于算法训练也是困难和昂贵的。

(3) 对不同故障进行分类的阈值通常由经验丰富的工程师预先设定。这对于现代生产中日益动态化的环境而言并非最佳解决方案。

3. 系统结构

计算机数控（CNC）机床中控制电机的功率数据可以反映出机床与工装的工作条件[19,20]。此外，与振动传感器或声学传感器[21,22]相比，功率传感器在实际应用中更具成本效益。因此，在本系统之中，我们选用由安装在计算机数控（CNC）设备上的无线传感器网络（WSN）作为数据来源，以此收集功率数据以支持生产设备和工装的异常诊断[5]。系统结构如图1所示。其功能说明如下：

(1) 数据存储库：我们配备并部署了一个大型数据基础设施，在生产过程中用于收集、存储及可视化实时监控数据[5]。

(2) 数据预处理：考虑到监控数据的准确性，我们设计了数据预处理机制。这些机制包括：①根据各个加工过程将数据划分为时间序列数据集；②利用高斯核模型[23,24]对来自监测数据源的波动信息进行去噪，以便进一步处理；③对数据进行标准化以确保监测数据的规模适合分析；④基于协方差交叉算法[5]进行数据校对，以标准和预设的参考模式重新调整功率数据，以便进行异常识别。

(3) 特征表示与异常识别：预设一组关键特征用于生产过程中的异常分析与诊断。关键特征的阈值则通过与实时数据比较用于异常识别。该系统对过程中产生的新异常保持开放，并在加工过程中动态更新。

(4) 阈值优化：我们设计了一种优化算法，以历史监测的数据为基础确定优化的阈值。

表1 综述研究方法总结

Method	De-noising, normalization and alignment	Application	Accuracy	Disadvantage	Ref.
Deep Boltzmann machine	No	Industrial diagnosis	29.85%–93.67%	High computation time	[8]
Modified SVM and RFE	No	Steel plates diagnosis	80.74%	Crippled detection rate	[9]
CMFE and ESVMs	No	Rolling bearing diagnosis	100.00%	Crippled detection rate	[10]
CNN	No	Chemical process diagnosis	91.00%	High computation time	[11]
Decision tree	No	Milling Tool diagnosis	81.00%	Structure and threshold difficult to define	[12]
Extreme learning machine	No	Aero-engine diagnosis	90.00%	High computation time	[13]
CNN based on LeNet-5	No	Motor bearing, etc.	99.481%–100.00%	High computation time	[14]
Sparse auto-encoder	No	Motor bearing	99.82%	High computation time	[15]
HCNN	No	Motor bearing, etc.	96.10%–99.82%	High computation time	[16]

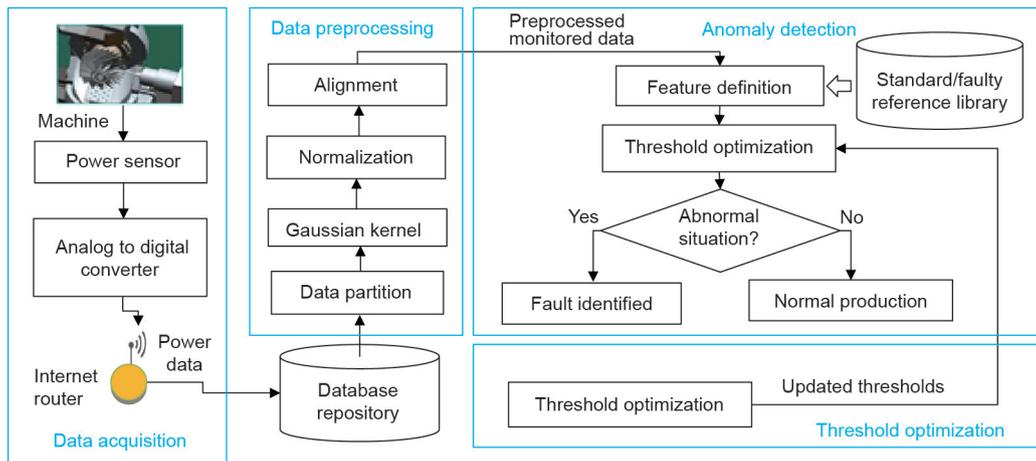


图1. 数控加工过程的系统框架。

4. 监控数据预处理

4.1. 监控数据分区处理

在加工过程中获得的监测功率数据将用于故障诊断。功率的计算公式如下：

$$P(i) = [I_1(i) + I_2(i) + I_3(i)] \times V \times Factor \quad (1)$$

式中， $P(i)$ 为沿时间轴（ x 轴）的功率数据的第 i 个点； $I_1(i)$ 、 $I_2(i)$ 以及 $I_3(i)$ 代表三相电流； V 为电源的电压； $Factor$ 为电源的质量系数。对加工过程中采集到的所有功率数据进行分析是耗时且无效的。为便于分析，我们首先根据机器特定的功率水平对监控数据进行分区，以表示加工过程的各道工序。然后对已分区的监视数据应用下述步骤以便于进一步分析。

4.2. 监控数据的降噪及平滑处理

通常而言，由于信号中的噪声干扰，监控的功率数据往往存在一定波动。为有效提取关键特征，必须对所

监视的数据进行去噪与平滑处理。在本研究当中，我们设计了一种基于高斯核的模型用于数据降噪处理。Feng等[23]以及Rimpault等[24]的研究已经证明了高斯核的鲁棒性。在此处，我们通过高斯核卷积计算来平滑处理所监测数据。在第 i 点处去噪及平滑处理后的功率数据 $P_\sigma(i)$ 计算如下：

$$P_\sigma(i) = \frac{\sum_{j=1}^n [P(i) \times g_\sigma(x_j)]}{\sum_{j=1}^n [g_\sigma(x_j)]} \quad (2)$$

式中， n 为 P （功率数据）中的总点数； x_j 为沿 x 轴（时间）的 P 中的第 j 个点；而 $g_\sigma(x_j)$ 则为内核宽度为 σ 的第 j 个点的高斯核。

上述过程的一个示例如图2所示。

4.3. 监控数据的标准化

为确保数据具有恰当的比例，我们对监测数据进行标准化，以便从数据之中提取关键特征（例如，在第5节中描述的峰值在没有标准化的情况下极高）：

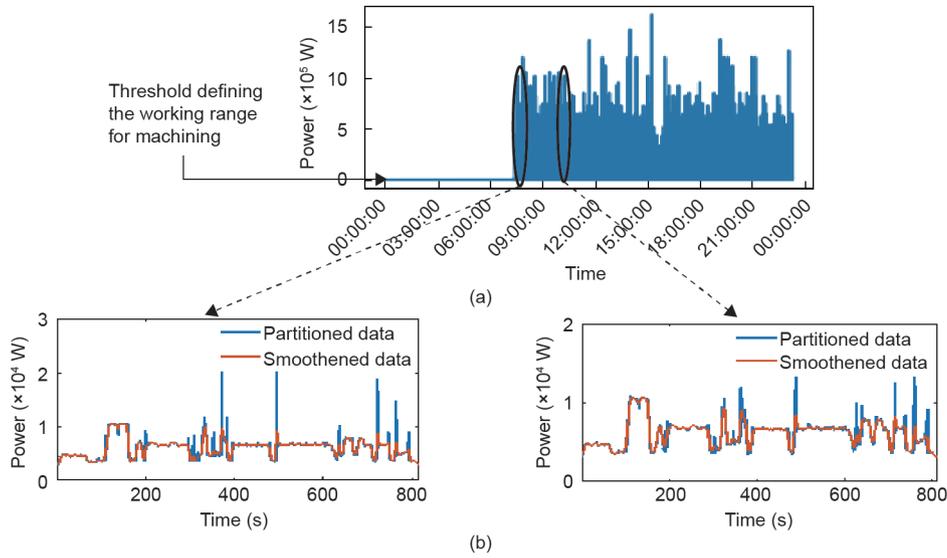


图2. 监控数据的数据分区与去噪示例。(a) 在一天内（2016年5月31日）获得的功率数据；(b) 分区与去噪后两个分区进程的功率模式（红色）。

$$NP = P_{\sigma}/P_{ref} \quad (3)$$

$$P_{cutting_alignment} = P_{cutting}(t + Difference) \quad (7)$$

式中，NP为标准化的功率数据； P_{σ} 为原始功率数据； P_{ref} 为机器设备的参考功率数据。

4.4. 监控数据校对

在实际的制造条件下，当加工部件时，分区监控数据可能存在时间延迟或偏差，这将会导致与标准模式（即在正常工作条件下加工相同部件时的功率模式）的不一致。我们将监测数据与参考标准（ $\sigma_{S_{cutting}}$ ）进行协方差交叉运算以识别时间延迟[5]。

$$\sigma_{S_{cutting}S_{standard}}(T) = \frac{1}{N-1} \times \sum_{t=1}^N [P_{standard}(t) - \mu_{cutting}] \times [P_{cutting}(t+T) - \mu_{standard}] \quad (4)$$

式中， $P_{standard}$ 和 $P_{cutting}$ 分别为参考标准和分区监控数据； $\mu_{standard}$ 和 $\mu_{cutting}$ 为时间序列的均值； N 为两个数据集当中较小的数；而 t 和 T 分别为时间偏差与标准时间。时间延迟可以通过以下公式计算：

$$X_{coef} = \frac{\sigma_{S_{cutting}S_{standard}}(T, T = 1 : N)}{\sqrt{\sigma_{S_{cutting}S_{cutting}}(0)\sigma_{S_{standard}S_{standard}}(0)}} \quad (5)$$

当 X_{coef} 取最大值时，时间延迟可计算如下：

$$Difference = T, \text{ when } X_{coef} \text{ is maximum} \quad (6)$$

因此，经校准的监控数据为：

将监测数据与故障参考 P_{fault} （即在异常条件下加工相同部件时的功率模式）进行校准时，亦是相同的步骤。只需将上述公式中的 $P_{standard}$ 换为 P_{fault} 。

5. 异常检测过程

在加工过程中，我们将定义一些关键特征用以表示预处理的监测数据与参考标准（即在正常工作条件下加工相同部件的数据模式）间的差异。在良好的工作条件下，我们在部件加工的过程收集数据以形成参考标准。如图3所示的诊断程序包括以下步骤：

(1) 关键特征是基于数据的绝对平均值、峰值以及峰值因子矩阵来表示的。相关定义详见表2。在表2当中， $Feature_m$ 是根据每条经预处理的监测数据及其标准参考值计算而得的。 $Feature_2 \sim Feature_m$ 是根据每条经预处理的监测数据及其故障参考值计算而得的（其中， m 的含义是异常类型）。

(2) 定义了一系列的阈值。我们设立 $Threshold_1$ ，通过比较 $Feature_1$ 和 $Threshold_1$ 就可以确定情况为正常或异常。我们设立 $Threshold_2 \sim Threshold_m$ ，可分别通过比较 $Feature_2 \sim Feature_m$ 和 $Threshold_2 \sim Threshold_m$ 来对异常类型进行分类。如果不存在现有的异常类型进行匹配，系统则会将新的异常类型更新到数据库当中。

(3) 基于最新的历史数据，通过FFO算法周期性地优化上述阈值。

在本研究当中，我们根据下述规则定义异常的工作

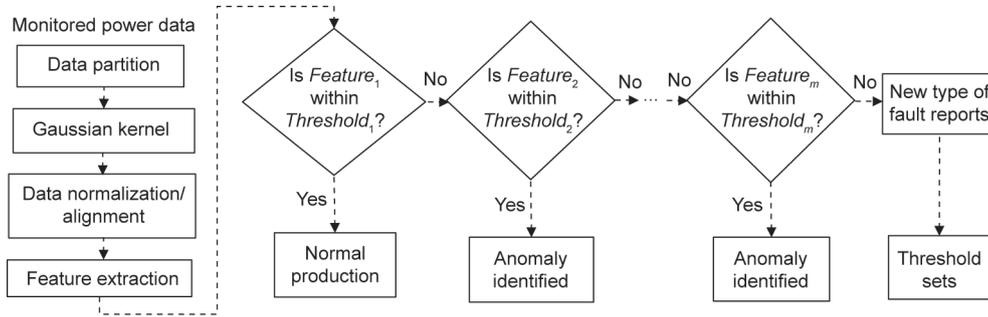


图3. 异常诊断过程。

表2 标准参考、错误参考和监控数据的特征和阈值的定义[25]

Variables	Difference between	
	Preprocessed monitored data and standard reference	Preprocessed monitored data and fault reference
Absolute mean	$Mean_{\text{standard}} = \frac{1}{N} \sum_{t=1}^N P_{\text{standard}} - P_{\text{cutting}} $	$Mean_{\text{fault}} = \frac{1}{N} \sum_{t=1}^N P_{\text{fault}} - P_{\text{cutting}} $
Kurtosis	$k_{\text{standard}} = \frac{1}{N} \sum_{t=1}^N (P_{\text{standard}} - P_{\text{cutting}})^4$	$k_{\text{fault}} = \frac{1}{N} \sum_{t=1}^N (P_{\text{fault}} - P_{\text{cutting}})^4$
Crest factor	$C_{\text{standard}} = \max(P_{\text{standard}} - P_{\text{cutting}}) / \sqrt{\frac{1}{N} \sum_{t=1}^N (P_{\text{standard}})^2}$	$C_{\text{fault}} = \max(P_{\text{fault}} - P_{\text{cutting}}) / \sqrt{\frac{1}{N} \sum_{t=1}^N (P_{\text{fault}})^2}$
Features	$Feature_1 = (Mean_{\text{standard}} k_{\text{standard}} C_{\text{standard}})$	$Feature_{2-n} = (Mean_{\text{fault}} k_{\text{fault}} C_{\text{fault}})$
Thresholds	$Threshold_1 = (V_{Mean_{\text{standard}}} V_{k_{\text{standard}}} V_{C_{\text{standard}}})$	$Threshold_{2-n} = (V_{Mean_{\text{fault}}} V_{k_{\text{fault}}} V_{C_{\text{fault}}})$

环境[5]。

- 刀具等工具磨损：功率范围明显垂直移动，但在设备空转阶段的功率范围保持不变。
- 刀具等工具破损：功率先是增加到峰值，随后又回到切割空气的功率范围。
- 主轴故障：在加工过程与设备空转阶段均突然出现功率峰值，同时功率范围增大。

根据上述规则与历史数据，我们可以对上述异常条件下的三个阈值做出定义：判断刀具等工具磨损的 $Threshold_2$ ，判断刀具等工具破损的 $Threshold_3$ 和判断主轴故障的 $Threshold_4$ 。确定最佳阈值的过程将在下一节中介绍。

6. 阈值优化

正如我们之前的研究中[5]所讨论的那样，整体检测准确度可以由四个因素决定：真阳性 (TP)、假阳性 (FP)、真阴性 (TN) 和假阴性 (FN)。 TP 表示异常情况被正确识别为异常； FP 表示正常情况被错误识别为异常； TN 表示正常情况被正确识别为正常；而 FN 表示异常情况被错误识别为正常。基于上述四要素，我们引入 $Precision$, $Recall$ 和 F 来评估整体检测精度[26]：

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

$$F = 2 \frac{Precision \times Recall}{Precision + Recall} \quad (10)$$

式中， $Precision$ 为正确识别的异常条件与所有参与识别的异常条件的比例； $Recall$ 为正确识别的异常情况与所有实际异常条件的比例； F 为检测的总体准确性。 F 得分越高（即越接近1），表明检测的整体准确度越高。

TP , FP , TN 和 FN 受四个阈值（即 $Threshold_1$, $Threshold_2$, $Threshold_3$, $Threshold_4$ ）的影响。因此，阈值的选择会影响最终的 F 分数。

在本研究中，阈值通过历史监测数据使用 FFO 算法而非依赖于专家的经验进行优化。 FFO 能够避免局部最优，并且具有比其他一些主流优化算法更优异的性能[27,28]。在该算法当中，我们对群体中心进行初始化以进行搜索（在本研究中，每个中心被模型化为四个阈值的向量，即 $Threshold_1 \sim Threshold_4$ ），随后围绕着每一个群体中心，将生成所谓的“果蝇”随机解决方案，我们利用基于气味与视觉的策略分别对最优化程度和群体中心的选择进行计算（详见步骤3和步骤4）。最终计算

迭代以达到优化目的。

优化目标是确定使得分数 F 可以取到最大的最佳阈值，我们取向量：

$$\text{Vector}(\text{Threshold}_1, \text{Threshold}_2, \text{Threshold}_3, \text{Threshold}_4) \rightarrow F \quad (11)$$

优化过程如下所述（步骤2和6中提供了对典型FFO算法的改进方案）：

步骤1 设最大迭代次数为 T_{\max} ，群体中心的种群大小为 v ，每个群体中心周围的果蝇数量为 k 。

步骤2 根据以下公式在每个群体中心周围随机生成果蝇：

$$\text{Vector}_{\text{sub}} = \text{Vector}_{\text{center}} \pm \alpha \times \text{rand} \quad (12)$$

式中， $\text{Vector}_{\text{center}}$ 和 $\text{Vector}_{\text{sub}}$ 分别为每个群体中心以及该群体中心周围的果蝇子群的向量； α 为确定果蝇在每个群体中心周围搜索边界大小的决定变量；而 rand 代表一个随机数。

步骤3 进行基于气味的搜索以计算每个果蝇对应的气味浓度（即最优化程度）。

步骤4 进行基于视觉的搜索，用具有相对最优效果的子群中的果蝇代替原来的群体中心，并指导子群进一步进行搜索。

步骤5 在典型的FFO算法中，搜索距离总是恒定的，所以当果蝇接近最终的最优结果时，搜索往往难以收敛。因此为了改进算法，当优化结果在五次迭代后仍没有得到进一步改善时，我们将缩短搜索距离。（这提高了收敛速度，因为当果蝇群体接近解决方案时，可以更容易地靠近该解决方案[29]。）

$$\alpha_{\text{new}} = \alpha - \frac{\alpha \times (T_{\max} - 1) \times 0.8}{T_{\max}} \quad (13)$$

式中， α_{new} 为接近最佳结果时所缩短的搜索距离。

步骤6 重复上述步骤，直到达到收敛的解或最大迭代次数 T_{\max} 。

7. 案例研究

在“欧盟智能及云端流量建构项目”（EU Smarter and Cloudflow projects）的赞助下，我们在英国一家公司的车间内开发和部署了一个无线传感器网络。该公司专门从事汽车、航空航天以及模具应用的高精度加工。

在本实践研究当中，我们对五轴铣床MX520进行了监控。6个月以来，我们收集到了超过10 GB的功率数据并将其储存在本地数据库当中。随后，基于开源平台Hadoop，我们开发了一个大型数据处理基础设施，用于管理海量数据并加速数据处理。

生产线的一部分如图4所示。三台电流传感器（每相一个）夹在数控机床的主电源上，每个传感器的数据搜集速率为每秒一个样本；随后，每秒产生的样本将通过工厂车间里的Wi-Fi传输到Hadoop的数据服务器。此后我们会根据三相电流、220 V电压和0.82的电源质量系数计算功率。

在本案例研究中，FFO算法旨在根据历史数据确定可以使 F 取得最高分数的最佳阈值。表3显示了此优化过程的基准测试结果。 Threshold_1 的优化阈值为(0.192, 0.032, 0.287)，代表刀具等生产工具磨损的阈值 Threshold_2 为(0.632, 0.410, 0.652)，代表刀具等生产工具破损的阈值 Threshold_3 为(3.698, 75.363, 10.737)，代表主轴故障的阈值 Threshold_4 为(2.412, 1.081, 0.921)。FFO算法可以在23次迭代中实现最优结果，与其他基准算法相比，收敛速度最快。同时，它可以使 F 的得分取值为1，这意味着优化的阈值可以基于历史数据实现100%的真实检测率。下文将介绍一些异常检测和识别的示例。

7.1. 正常生产

图5显示了用于异常检测的监测数据的分析过程。所取的关键特征 Feature_1 为(0.147, 0.004, 0.113)，小于 Threshold_1 (0.192, 0.032, 0.287)（特征和阈值的定义见

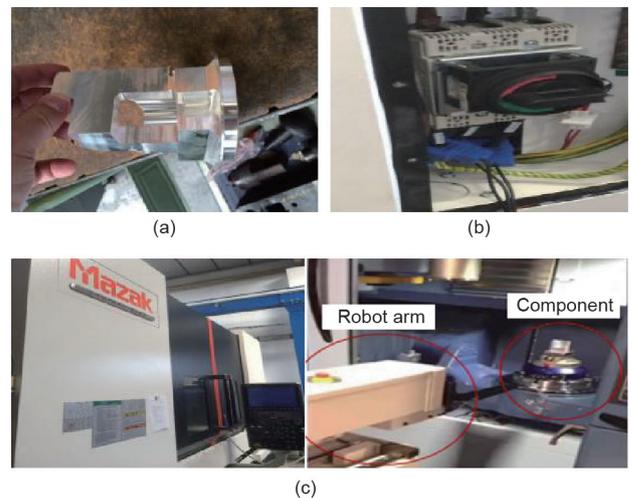


图4. CNC加工过程。(a) 机加工零件；(b) 功率测量；(c) 马扎克机床及其加工工艺。

表2)。因此，它可以被归类为正常生产。

7.2. 异常情况——刀具等生产工具磨损

对于图6(a)中所示的监测数据，其 $Feature_1$ 为(0.206, 0.042, 0.295)，高于阈值 $Threshold_1$ (0.192, 0.032, 0.287)。因此，此生产过程被归类为存在异常。随后，我们对其进行异常诊断[图6(b)]。 $Feature_2$ 为(0.171, 0.058, 0.250)，小于 $Threshold_2$ (0.632, 0.410, 0.652)。因此，该生产异常可以被分类为工具磨损。

7.3. 异常情况——刀具等生产工具破损

对于图7(a)所示的监测数据， $Feature_1$ 为(0.460, 41.532, 2.303)，高于 $Threshold_1$ (0.192, 0.032, 0.287)。因

此，此生产过程被归类为存在异常。随后，我们对其进行异常诊断[图7(b)]。 $Feature_3$ 为(1.039, 61.512, 1.744)，小于阈值 $Threshold_3$ (3.698, 75.363, 10.737)。因此，该生产异常可以被分类为工具破损。

7.4. 新的异常情况——设备长时间空转

图8显示了用于异常检测的监测数据的分析过程。 $Feature_1$ 为(0.492, 0.441, 0.379)，高于 $Threshold_1$ (0.192, 0.032, 0.287)。因此，此生产过程存在异常。但是，数据库中没有任何类似于此数据模型的故障参照。因此，该情况下产生的数据被报告给了车间工程师。最终发现设备意外地一直在空转。随后，该数据模型被保存至数据库以更新故障参照。

表3 优化算法的比较

	FFO	GA	SA
Iterations to reach optimization	23	44	51
Optimized F	1	1	1

GA: genetic algorithms; SA: simulated annealing.

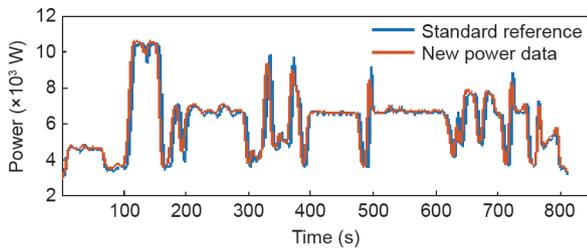


图5. 监测数据表明生产条件正常。

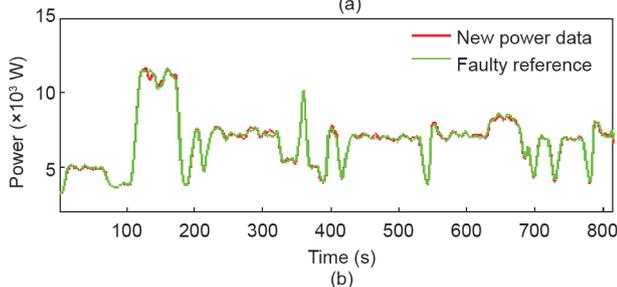
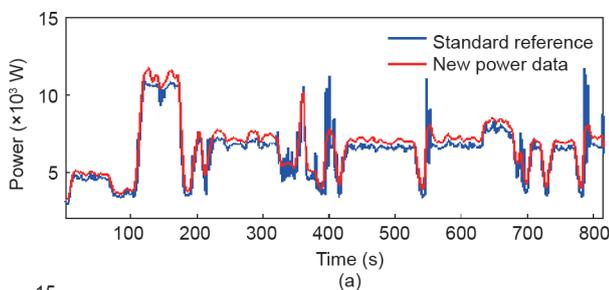


图6. 刀具等生产工具磨损检测。(a) 故障识别；(b) 故障分类。

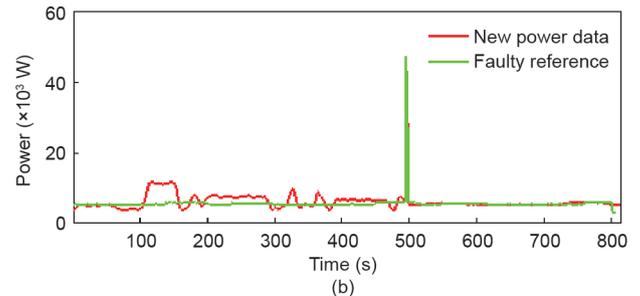
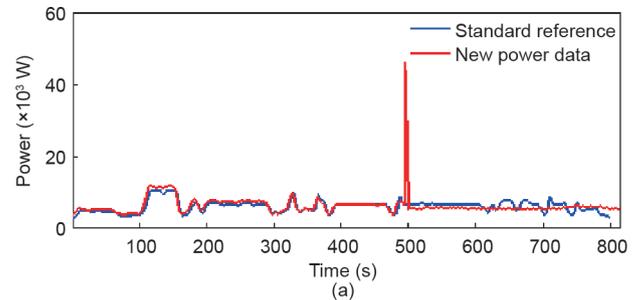


图7. 刀具等生产工具破损检测。(a) 故障识别；(b) 故障分类。

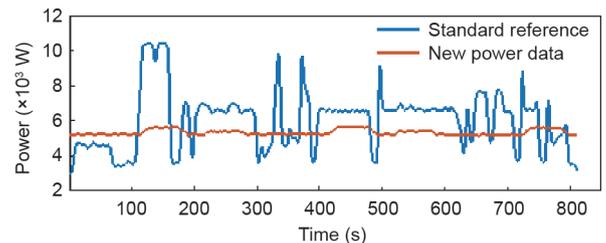


图8. 新的异常数据检测。

8. 总结

在本研究中，我们开发了数据驱动的异常情况分析方法。该系统在某机加工企业进行了实际加工条件下的

验证。本研究的创新点如下：

(1) 开发了预处理机制，包括去噪、数据标准化和校准，解决了监控数据的准确性问题。

(2) 设计了FFO算法来识别最佳异常阈值，以便在动态加工过程中实现更准确的检测。今后，我们将进行进一步的调查，以提升该系统的可靠性，调查将包括以下内容：①我们将测试不同的数据采样率，以找到最佳的系统精度和效率。此外，我们将尝试通过使用数据融合来增强预测结果，并考虑不同的数据源（如振动、力数据等）。②我们将考虑设计高效的深度学习算法以及有效的计算架构[如用于递归神经网络（RNN）、长短期记忆递归神经网络（LSTM RNN）等的迁移学习算法和边缘计算架构]，以进一步提高系统性能。

Acknowledgement

The authors acknowledge the funding from the EU Smarter project (PEOPLE-2013-IAPP-610675).

Compliance with ethics guidelines

Y.C. Liang, S. Wang, W.D. Li, and X. Lu declare that they have no conflict of interest or financial conflicts to disclose.

References

- [1] Bayar N, Darmoul S, Hajri-Gabouj S, Pierreval H. Fault detection, diagnosis and recovery using artificial immune systems: a review. *Eng Appl Artif Intell* 2015;46:43–57.
- [2] Venkatasubramanian V, Rengaswamy R, Yin K, Kavuri SN. A review of process fault detection and diagnosis: part I: quantitative model-based methods. *Comput Chem Eng* 2003;27(3):293–311.
- [3] Aydin I, Karakose M, Akin E. Chaotic-based hybrid negative selection algorithm and its applications in fault and anomaly detection. *Expert Syst Appl* 2010;37(7):5285–94.
- [4] Yang H, Li T, Hu X, Wang F, Zou Y. A survey of artificial immune system based intrusion detection. *Sci World J* 2014;2014:156790.
- [5] Wang S, Liang YC, Li WD, Cai XT. Big data enabled intelligent immune system for energy efficient manufacturing management. *J Clean Prod* 2018;195:507–20.
- [6] Gao R, Wang L, Teti R, Dornfeld D, Kumara S, Mori M, et al. Cloud-enabled prognosis for manufacturing. *CIRP Ann* 2015;64(2):749–72.
- [7] Lee J, Wu F, Zhao W, Ghaffari M, Liao L, Siegel D. Prognostics and health management design for rotary machinery systems—reviews, methodology and applications. *Mech Syst Signal Process* 2014;42(1–2):314–34.
- [8] Hu G, Li H, Xia Y, Luo L. A deep Boltzmann machine and multi-grained scanning forest ensemble collaborative method and its application to industrial fault diagnosis. *Comput Ind* 2018;100:287–96.
- [9] Tian Y, Fu M, Wu F. Steel plates fault diagnosis on the basis of support vector machines. *Neurocomputing* 2015;151:296–303.
- [10] Zheng J, Pan H, Cheng J. Rolling bearing fault detection and diagnosis based on composite multiscale fuzzy entropy and ensemble support vector machines. *Mech Syst Signal Process* 2017;85:746–59.
- [11] Wu H, Zhao J. Deep convolutional neural network model based chemical process fault diagnosis. *Comput Chem Eng* 2018;115:185–97.
- [12] Madhusudana C, Kumar H, Narendranath S. Fault diagnosis of face milling tool using decision tree and sound signal. *Materials Today Proc* 2018;5(5):12035–44.
- [13] Lu F, Jiang J, Huang J, Qiu X. Dual reduced kernel extreme learning machine for aero-engine fault diagnosis. *Aerosp Sci Technol* 2017;71:742–50.
- [14] Wen L, Li X, Gao L, Zhang Y. A new convolutional neural network-based datadriven fault diagnosis method. *IEEE Trans Ind Electron* 2018;65(7):5990–8.
- [15] Wen L, Gao L, Li X. A new deep transfer learning based on sparse auto-encoder for fault diagnosis. *IEEE Trans Syst Man Cybern Syst* 2019;49(1):136–44.
- [16] Wen L, Li X, Gao L. A new two-level hierarchical diagnosis network based on convolutional neural network. *IEEE Trans Instrum Meas*. Forthcoming 2019.
- [17] García S, Luengo J, Herrera F. Tutorial on practical tips of the most influential data preprocessing algorithms in data mining. *Knowl Base Syst* 2016;98:1–29.
- [18] Pan S, Yang Q. A survey on transfer learning. *IEEE Trans Knowl Data Eng* 2010;22(10):1345–59.
- [19] Liu Z, Guo Y, Sealy M, Liu Z. Energy consumption and process sustainability of hard milling with tool wear progression. *J Mater Process Technol* 2016;229:305–12.
- [20] Sealy M, Liu Z, Zhang D, Guo Y, Liu Z. Energy consumption and modeling in precision hard milling. *J Clean Prod* 2016;135:1591–601.
- [21] Stoney R, Donohoe B, Geraghty D, O'Donnell G. The development of surface acoustic wave sensors (SAWs) for process monitoring. *Procedia CIRP* 2012;1:569–74.
- [22] García Plaza E, Núñez López PJ. Application of the wavelet packet transform to vibration signals for surface roughness monitoring in CNC turning operations. *Mech Syst Signal Process* 2018;98:902–19.
- [23] Feng Z, Zuo M, Chu F. Application of regularization dimension to gear damage assessment. *Mech Syst Signal Process* 2010;24(4):1081–98.
- [24] Rimpault X, Bitar-Nehme E, Balazinski M, Mayer J. Online monitoring and failure detection of capacitive displacement sensor in a Capball device using fractal analysis. *Measurement* 2018;118:23–8.
- [25] Xia M, Li T, Xu L, Liu L, De Silva C. Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks. *IEEE/ASME Trans Mechatron* 2018;23(1):101–10.
- [26] Zheng X, Wang L, Wang S. A novel fruit fly optimization algorithm for the semiconductor final testing scheduling problem. *Knowl Base Syst* 2014;57:95–103.
- [27] Liang Y, Lu X, Li W, Wang S. Cyber physical system and big data enabled energy efficient machining optimisation. *J Clean Prod* 2018;187:46–62.
- [28] Powers D. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation. *J Mach Learn Technol* 2011;2(1):37–63.
- [29] Du T, Ke X, Liao J, Shen Y. DSLC-FOA: improved fruit fly optimization algorithm for application to structural engineering design optimization problems. *Appl Math Model* 2018;55:314–39.