



(12) 发明专利

(10) 授权公告号 CN 111177655 B

(45) 授权公告日 2022.04.22

(21) 申请号 201911403955.0

(22) 申请日 2019.12.30

(65) 同一申请的已公布的文献号

申请公布号 CN 111177655 A

(43) 申请公布日 2020.05.19

(73) 专利权人 联想(北京)有限公司

地址 100085 北京市海淀区上地信息产业基地创业路6号

(72) 发明人 金继民 杨帆 杨沛

(74) 专利代理机构 北京集佳知识产权代理有限公司 11227

代理人 李金

(51) Int.Cl.

G06F 17/18 (2006.01)

G06K 9/62 (2022.01)

权利要求书3页 说明书11页 附图5页

(54) 发明名称

一种数据处理方法、装置及电子设备

(57) 摘要

本申请公开了一种数据处理方法、装置及电子设备，方法包括：获得多个训练对象的工艺数据，所述工艺数据包含工艺标识和至少两项工艺参数，所述工艺标识表征所述训练对象是否为工艺合格的对象；以所述训练对象的工艺数据作为预先构建的工艺预测模型的训练样本，对所述工艺预测模型进行训练，得到训练完成的工艺预测模型；其中，所述工艺质量预测模型用于以待预测的目标对象的工艺数据作为输入，输出所述目标对象的预测结果，所述预测结果至少表征所述目标对象是否工艺合格。

(56) 对比文件

CN 110288199 A, 2019.09.27

CN 110288199 A, 2019.09.27

CN 110059738 A, 2019.07.26

CN 110334816 A, 2019.10.15

CN 109238610 A, 2019.01.18

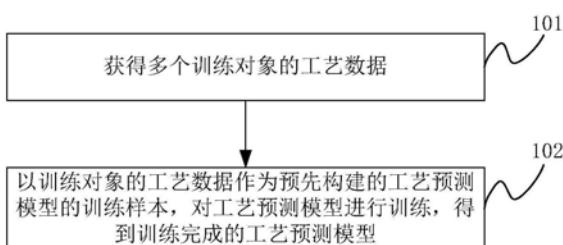
CN 107107237 A, 2017.08.29

CN 107107237 A, 2017.08.29

CN 109783898 A, 2019.05.21

CN 105279291 A, 2016.01.27

审查员 郑诗嘉



1.一种数据处理方法,包括:

获得多个对象的拉力测试数据,所述拉力测试数据包括对所述多个对象中的部分对象进行拉力测试所产生的数据;

根据所述拉力测试数据,获得所述多个对象中的处于第一区间的第一对象、处于第二区间的第二对象和被拉力测试的第三对象;所述第一区间为根据所述拉力测试数据确定为工艺合格的对象区间,所述第二区间为根据所述拉力测试数据确定为工艺不合格的对象区间;

从所述第一对象、所述第二对象和所述第三对象中,获得多个训练对象;

获得所述多个训练对象的工艺数据,所述工艺数据包含工艺标识和至少两项工艺参数,所述工艺标识表征所述训练对象是否为工艺合格的对象;

对所述训练对象的工艺数据添加质量参数,所述质量参数的参数值表征所述训练对象的工艺质量;所述质量参数用于对工艺预测模型进行训练;

其中,所述第一对象的质量参数为第一参数值,所述第二对象的质量参数为第二参数值;所述第二参数值根据所述第二对象在所述第二区间中的区间位置确定,其中,靠近所述第二区间的区间边缘的第二对象的第二参数值高于靠近所述第二区间的区间中间的第二对象的第二参数值;被拉力测试为工艺合格的第三对象的质量参数为所述第一参数值,被拉力测试为工艺不合格的第三对象的质量参数为第三参数值,所述第一参数值大于所述第二参数值,且所述第二参数值大于所述第三参数值;

以所述训练对象的工艺数据作为预先构建的工艺预测模型的训练样本,对所述工艺预测模型进行训练,得到训练完成的工艺预测模型;

其中,所述工艺预测模型用于以待预测的目标对象的工艺数据作为输入,输出所述目标对象的预测结果,所述预测结果至少表征所述目标对象是否工艺合格。

2.根据权利要求1所述的方法,以所述训练对象的工艺数据作为预先构建的工艺预测模型的训练样本,对所述工艺预测模型进行训练,得到训练完成的工艺预测模型,包括:

对所述训练对象的工艺数据进行预处理;

将所述工艺数据中的工艺参数作为预先构建的工艺预测模型的输入,将所述工艺数据中的质量参数作为所述工艺预测模型的输出,对所述工艺预测模型进行训练,得到训练完成的工艺预测模型;所述工艺预测模型的模型参数中至少包括质量阈值;

其中,在所述工艺预测模型输出的所述目标对象的预测质量值大于或等于所述质量阈值的情况下,表征所述目标对象为工艺合格的对象;在所述目标对象的预测质量值小于所述质量阈值的情况下,表征所述目标对象为工艺不合格的对象。

3.根据权利要求1所述的方法,还包括:

从所述第一对象和所述第三对象中,获得多个验证对象,所述验证对象与所述训练对象不同;

获得所述多个验证对象的工艺数据;

将所述多个验证对象的工艺数据作为所述工艺预测模型的输入,以得到所述工艺预测模型输出的所述验证对象的预测结果;

根据所述验证对象的预测结果,对所述工艺预测模型的模型参数进行修改。

4.根据权利要求3所述的方法,根据所述验证对象的预测结果,对所述模型参数进行修

改,包括:

将所述验证对象的预测结果和所述验证对象的工艺数据中的工艺标识进行比对,得到比对结果;

根据所述比对结果,对所述模型参数进行修改。

5.根据权利要求1所述的方法,以所述训练对象的工艺数据作为预先构建的工艺预测模型的训练样本,对所述工艺预测模型进行训练之前,所述方法还包括:

获得所述工艺数据中的工艺参数与所述训练对象的质量参数之间的相关度值;

根据所述相关度值,筛选所述工艺数据中的工艺参数。

6.根据权利要求1所述的方法,所述工艺预测模型为基于回归算法构建的模型。

7.一种数据处理装置,包括:

数据获得单元,用于获得多个对象的拉力测试数据,所述拉力测试数据包括对所述多个对象中的部分对象进行拉力测试所产生的数据;根据所述拉力测试数据,获得所述多个对象中的处于第一区间的第一对象、处于第二区间的第二对象和被拉力测试的第三对象;所述第一区间为根据所述拉力测试数据确定为工艺合格的对象区间,所述第二区间为根据所述拉力测试数据确定为工艺不合格的对象区间;从所述第一对象、所述第二对象和所述第三对象中,获得多个训练对象;获得所述多个训练对象的工艺数据,所述工艺数据包括工艺标识和至少两项工艺参数,所述工艺标识表征所述训练对象是否为工艺合格的对象;对所述训练对象的工艺数据添加质量参数,所述质量参数的参数值表征所述训练对象的工艺质量;所述质量参数用于对工艺预测模型进行训练;其中,所述第一对象的质量参数为第一参数值,所述第二对象的质量参数为第二参数值,所述第二参数值根据所述第二对象在所述第二区间中的区间位置确定,其中,靠近所述第二区间的区间边缘的第二对象的第二参数值高于靠近所述第二区间的区间中间的第二对象的第二参数值;被拉力测试为工艺合格的第三对象的质量参数为所述第一参数值,被拉力测试为工艺不合格的第三对象的质量参数为第三参数值,所述第一参数值大于所述第二参数值,且所述第二参数值大于所述第三参数值;

模型训练单元,用于以所述训练对象的工艺数据作为预先构建的工艺预测模型的训练样本,对工艺预测模型进行训练,得到训练完成的工艺预测模型;

其中,所述工艺预测模型用于以待预测的目标对象的工艺数据作为输入,输出所述目标对象的预测结果,所述预测结果至少表征目标对象是否工艺合格。

8.一种电子设备,包括:

存储器,用于存储应用程序和所述应用程序运行所产生的数据;

处理器,用于执行所述应用程序,以实现:获得多个对象的拉力测试数据,所述拉力测试数据包括对所述多个对象中的部分对象进行拉力测试所产生的数据;根据所述拉力测试数据,获得所述多个对象中的处于第一区间的第一对象、处于第二区间的第二对象和被拉力测试的第三对象;所述第一区间为根据所述拉力测试数据确定为工艺合格的对象区间,所述第二区间为根据所述拉力测试数据确定为工艺不合格的对象区间;从所述第一对象、所述第二对象和所述第三对象中,获得多个训练对象;获得所述多个训练对象的工艺数据,所述工艺数据包括工艺标识和至少两项工艺参数,所述工艺标识表征所述训练对象是否为工艺合格的对象;对所述训练对象的工艺数据添加质量参数,所述质量参数的参数值表征

所述训练对象的工艺质量；所述质量参数用于对工艺预测模型进行训练；其中，所述第一对象的质量参数为第一参数值，所述第二对象的质量参数为第二参数值，所述第二参数值根据所述第二对象在所述第二区间中的区间位置确定，其中，靠近所述第二区间的区间边缘的第二对象的第二参数值高于靠近所述第二区间的区间中间的第二对象的第二参数值；被拉力测试为工艺合格的第三对象的质量参数为所述第一参数值，被拉力测试为工艺不合格的第三对象的质量参数为第三参数值，所述第一参数值大于所述第二参数值，且所述第二参数值大于所述第三参数值；以所述训练对象的工艺数据作为预先构建的工艺预测模型的训练样本，对工艺预测模型进行训练，得到训练完成的工艺预测模型；其中，所述工艺预测模型用于以待预测的目标对象的工艺数据作为输入，输出所述目标对象的预测结果，所述预测结果至少表征目标对象是否工艺合格。

一种数据处理方法、装置及电子设备

技术领域

[0001] 本申请涉及工艺检测技术领域,尤其涉及一种数据处理方法、装置及电子设备。

背景技术

[0002] 为了节能环保,电池已经成为核心的新能源,多个电池通常焊接成一个电池组为设备供电。

[0003] 而焊接点的牢固性和可靠性是电池产品质量的核心指标之一。在实际的生产过程中,一般每4小时会对流水线上的产品进行一次随机抽样,并对样本进行拉力测试,若焊接点需要64N以上(包含64N)的拉力才能拉开,则样品合格,否则样品不合格。若样品不合格,则通过排查的方式寻找不合格产品区间(异常区间),并将该区间内的产品统一默认为不合格产品。

[0004] 但是,拉力测试中不合格产品区间内是包含合格产品的,由此,对电池的工艺检测仍然存在不准确的情况,因此如何更加准确的对电池工艺进行检测,成为电池生产厂商所关心的问题。

发明内容

[0005] 有鉴于此,本申请提供一种数据处理方法、装置及电子设备,包括:

[0006] 一种数据处理方法,包括:

[0007] 获得多个训练对象的工艺数据,所述工艺数据包含工艺标识和至少两项工艺参数,所述工艺标识表征所述训练对象是否为工艺合格的对象;

[0008] 以所述训练对象的工艺数据作为预先构建的工艺预测模型的训练样本,对所述工艺预测模型进行训练,得到训练完成的工艺预测模型;

[0009] 其中,所述工艺质量预测模型用于以待预测的目标对象的工艺数据作为输入,输出所述目标对象的预测结果,所述预测结果至少表征所述目标对象是否工艺合格。

[0010] 上述方法,优选的,获得多个训练对象的工艺数据,包括:

[0011] 获得多个对象的拉力测试数据,所述拉力测试数据包括对所述多个对象中的部分对象进行拉力测试所产生的数据;

[0012] 根据所述拉力测试数据,获得所述多个对象中的处于第一区间的第一个对象、处于第二区间的第二个对象和被拉力测试的第三个对象;所述第一区间为根据所述拉力测试数据确定为工艺合格的对象区间,所述第二区间为根据所述拉力测试数据确定为工艺不合格的对象区间;

[0013] 从所述第一个对象、所述第二个对象和所述第三个对象中,获得多个训练对象;

[0014] 获得所述多个训练对象的工艺数据。

[0015] 上述方法,优选的,在以所述训练对象的工艺数据作为预先构建的工艺预测模型的训练样本,对所述工艺预测模型进行训练之前,所述方法还包括:

[0016] 对所述训练对象的工艺数据添加质量参数,所述质量参数的参数值表征所述训练

对象的工艺质量；所述质量参数用于对所述工艺预测模型进行训练；

[0017] 其中，所述第一对象的质量参数为第一参数值，所述第二对象的质量参数为第二参数值，所述第二参数值根据所述第二对象在所述第二区间中的区间位置确定，被拉力测试为工艺合格的第三对象的质量参数为所述第一参数值，被拉力测试为工艺不合格的第三对象的质量参数为第三参数值，所述第一参数值大于所述第二参数值，且所述第二参数值大于所述第三参数值。

[0018] 上述方法，优选的，以所述训练对象的工艺数据作为预先构建的工艺预测模型的训练样本，对所述工艺预测模型进行训练，得到训练完成的工艺预测模型，包括：

[0019] 对所述训练对象的工艺数据进行预处理；

[0020] 将所述工艺数据中的工艺参数作为预先构建的工艺预测模型的输入，将所述工艺数据中的质量参数作为所述工艺预测模型的输出，对所述工艺预测模型进行训练，得到训练完成的工艺预测模型；所述工艺预测模型的模型参数中至少包括质量阈值；

[0021] 其中，在所述工艺预测模型输出的所述目标对象的预测质量值大于或等于所述质量阈值的情况下，表征所述目标对象为工艺合格的对象；在所述目标对象的预测质量值小于所述质量阈值的情况下，表征所述目标对象为工艺不合格的对象。

[0022] 上述方法，优选的，还包括：

[0023] 从所述第一对象和所述第三对象中，获得多个验证对象，所述验证对象与所述训练对象不同；

[0024] 获得所述多个验证对象的工艺数据；

[0025] 将所述多个验证对象的工艺数据作为所述工艺预测模型的输入，以得到所述工艺预测模型输出的所述验证对象的预测结果；

[0026] 根据所述验证对象的预测结果，对所述模型参数进行修改。

[0027] 上述方法，优选的，根据所述验证对象的预测结果，对所述模型参数进行修改，包括：

[0028] 将所述验证对象的预测结果和所述验证对象的工艺数据中的工艺标识进行比对，得到比对结果；

[0029] 根据所述比对结果，对所述模型参数进行修改。

[0030] 上述方法，优选的，以所述训练对象的工艺数据作为预先构建的工艺预测模型的训练样本，对所述工艺预测模型进行训练之前，所述方法还包括：

[0031] 获得所述工艺数据中的工艺参数与所述训练对象的质量参数之间的相关度值；

[0032] 根据所述相关度值，筛选所述工艺数据中的工艺参数。

[0033] 上述方法，优选的，所述工艺预测模型为基于回归算法构建的模型。

[0034] 一种数据处理装置，包括：

[0035] 数据获得单元，用于获得多个训练对象的工艺数据，所述工艺数据包括工艺标识和至少两项工艺参数，所述工艺标识表征所述训练对象是否为工业合格的对象；

[0036] 模型训练单元，用于以所述训练对象的工艺数据作为预先构建的工艺预测模型的训练样本，对工艺预测模型进行训练，得到训练完成的工艺预测模型；

[0037] 其中，所述工艺预测模型用于以待预测的目标对象的工艺数据作为输入，输出所述目标对象的预测结果，所述预测结果至少表征目标对象是否工艺合格。

- [0038] 一种电子设备,包括:
- [0039] 存储器,用于存储应用程序和所述应用程序运行所产生的数据;
- [0040] 处理器,用于执行所述应用程序,以实现:获得多个训练对象的工艺数据,所述工艺数据包括工艺标识和至少两项工艺参数,所述工艺标识表征所述训练对象是否为工业合格的对象;以所述训练对象的工艺数据作为预先构建的工艺预测模型的训练样本,对工艺预测模型进行训练,得到训练完成的工艺预测模型;其中,所述工艺预测模型用于以待预测的目标对象的工艺数据作为输入,输出所述目标对象的预测结果,所述预测结果至少表征目标对象是否工艺合格。
- [0041] 从上述技术方案可以看出,本申请公开的一种数据处理方法、装置及电子设备,通过获得多个训练对象的工艺数据,进而以这些工艺数据中的工艺标识和至少两项工艺参数作为预先构建的工艺预测模型的训练样本,对模型进行训练,进而训练完成的工艺预测模型就可以用来对待预测的目标对象进行预测,以预测出目标对象是否工艺合格。可见,本申请的技术方案中无需对目标对象进行破坏性的测试操作,就可以利用根据训练对象的工艺数据所训练出的工艺预测模型实现对任一对象的工艺检测,而不会存在不合格产品区间包含合格产品的情况,因此,本申请的技术方案能够提高对对象工艺检测的准确性。

附图说明

[0042] 为了更清楚地说明本申请实施例的技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本申请的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

- [0043] 图1为本申请实施例一提供的一种数据处理方法的流程图;
- [0044] 图2为本申请实施例一的部分流程图;
- [0045] 图3为本申请实施例一关于对象区间的示例图;
- [0046] 图4及图5分别为本申请实施例一提供的一种数据处理方法的另一流程图;
- [0047] 图6为本申请实施例二提供的一种数据处理装置的结构示意图;
- [0048] 图7为本申请实施例三提供的一种电子设备的结构示意图;
- [0049] 图8为本申请适用于电池焊接工艺检测的示例图。

具体实施方式

[0050] 下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本申请一部分实施例,而不是全部的实施例。基于本申请中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本申请保护的范围。

[0051] 参考图1,为本申请实施例一提供的一种数据处理方法的实现流程图,该方法可以适用于能够进行数据处理的电子设备中,如计算机或服务器等。本实施例中的技术方案主要用于训练出能够对任一对象进行工艺检测的工艺预测模型,以便于提高对象工艺检测的准确性,同时也能够提高对象工艺检测的效率。

- [0052] 具体的,本实施例中的方法可以包括以下步骤:

[0053] 步骤101:获得多个训练对象的工艺数据。

[0054] 其中,工艺数据中包含有工艺标识,该工艺标识表征训练对象是否为工艺合格的对象。具体的,工艺标识可以为数字、字母、文字或符号等字符或字符串实现,例如,以1作为工艺标识来表征训练对象工艺合格,以0作为工艺标识来表征训练对象工艺不合格。

[0055] 另外,工艺数据中还包括有该工艺标识下的至少两项工艺参数,这里的工艺参数可以为训练对象在生产工艺中的相关参数,如以训练对象为电池为例,工艺参数可以包括有:电池的焊接压力参数、焊接时间参数、焊接振幅参数和焊接能量参数中的一种或任意多种的组合。

[0056] 具体的,本实施例中可以在工艺生产的信息系统数据库中读取到已经完成工艺生产的对象的工艺数据,这些对象即为训练对象。

[0057] 步骤102:以训练对象的工艺数据作为预先构建的工艺预测模型的训练样本,对工艺预测模型进行训练,得到训练完成的工艺预测模型。

[0058] 其中,工艺预测模型可以基于回归算法预先构建,这里的回归算法可以为梯度提升迭代决策树GBDT(Gradient Boosting Decision Tree)、神经网络、随机森林等回归算法,相应的,在构建完成工艺预测模型之后,将训练对象的工艺数据中的工艺参数作为工艺预测模型的输入,将训练对象的工艺数据中的工艺标识作为工艺预测模型的输出,对工艺预测模型中的模型参数进行训练及优化,进而得到训练完成的工艺预测模型。

[0059] 相应的,训练完成的工艺预测模型可以用于对待预测的目标对象进行工艺检测,具体的,在存在需要进行工艺检测的目标对象的情况下,将目标对象的工艺数据如工艺标识和至少一项工艺参数等输入到工艺预测模型中,并运行工艺预测模型,以得到工艺预测模型输出的目标对象的预测结果,该预测结果至少表征目标对象是否工艺合格。

[0060] 例如,将待检测焊接工艺的目标电池的焊接工艺数据和是否为焊接合格的工艺标识输入到工艺预测模型中,以输出目标电池是否为焊接工艺合格的电池。

[0061] 需要说明的是,在工艺预测模型的训练中,可以利用K-折交叉验证的方法进行模型优化,以提高模型的准确性。

[0062] 由上述方案可知,本申请实施例一提供的一种数据处理方法,通过获得多个训练对象的工艺数据,进而以这些工艺数据中的工艺标识和至少两项工艺参数作为预先构建的工艺预测模型的训练样本,对模型进行训练,进而训练完成的工艺预测模型就可以用来对待预测的目标对象进行预测,以预测出目标对象是否工艺合格。可见,本实施例的技术方案中无需对目标对象进行破坏性的测试操作,就可以利用根据训练对象的工艺数据所训练出的工艺预测模型实现对任一对象的工艺检测,而不会存在不合格产品区间包含合格产品的情况,因此,本申请的技术方案能够提高对对象工艺检测的准确性。

[0063] 在一种实现方式中,本实施例中在获得多个训练对象的工艺数据时,可以首先在多个对象中选出训练对象,再获得训练对象的工艺数据,具体如图2中所示:

[0064] 步骤201:获得多个对象的拉力测试数据。

[0065] 其中,拉力测试数据中包括有对多个对象中的部分对象进行拉力测试所产生的数据。拉力测试数据是指,对多个对象中的部分对象施加64N以上(包含64N)的拉力后所得到的拉力测试结果后对多个对象标记是否工艺合格的数据。具体的,本实施例中可以在测试数据库中读取到拉力测试数据。

[0066] 需要说明的是,拉力测试数据中包含有测试工艺合格的对象区间的数据、测试工艺异常的对象区间的数据和被拉力测试的对象的数据,如图3中所示。

[0067] 步骤202:根据拉力测试数据,获得多个对象中处于第一区间的第一对象、处于第二区间的第二对象和被拉力测试的第三对象。

[0068] 其中,第一区间为根据某个对象批次中经过拉力测试后所得到的拉力测数据确定为工艺合格的对象区间,如对电池的焊接点进行拉力测试后确定的正常区间,该第一区间中的对象即为第一对象;第二区间为根据经过拉力测试后所得到的拉力测试数据确定为工艺不合格的对象区间,如对电池的焊接点进行拉力测试后确定的异常区间,该第二区间中的对象即为第二对象,而被拉力测试的对象即为第三对象,如图3中所示。

[0069] 步骤203:从第一对象、第二对象和第三对象中,获得多个训练对象。

[0070] 也就是说,本实施例中从测试工艺合格的对象区间、测试工艺异常的对象区间和被拉力测试的对象中选取全部或部分作为训练对象。

[0071] 步骤204:获得多个训练对象的工艺数据。

[0072] 其中,在选取出训练对象之后,在工艺生产的信息系统数据库中读取到这些训练对象的工艺数据。此时,训练对象的工艺数据中既包含测试工艺合格的对象的工艺数据,也包含测试工艺异常的对象的工艺数据,还包含有被拉力测试后确定为工艺合格或不合格的对象的工艺数据,由此,经过训练对象的工艺数据训练的工艺预测模型能够更加准确的检测目标对象的生产工艺是否合格。

[0073] 基于以上实现,本实施例中在训练工艺训练模型之前,本实施例中的方法还可以包括以下步骤,如图4中所示:

[0074] 步骤103:对训练对象的工艺数据添加质量参数。

[0075] 其中,训练对象的质量参数的参数值表征训练对象的工艺质量,如焊接质量或拼接质量等。这里的参数值可以以大于或等于0,且小于或等于1的值来表示,参数值越大,相应训练对象的工艺质量越高,相应训练对象工艺合格的可能性就越高。而质量参数能够与工艺数据中的工艺标识和工艺参数一起用于对工艺预测模型进行训练。

[0076] 需要说明的是,在对训练对象的工艺数据参加质量参数时,可以根据训练对象所属的对象区间进行参数值的设置,例如,第一对象的质量参数设置为第一参数值,如1;

[0077] 第二对象的质量参数设置为第二参数值,第二参数值为大于0且小于1的值来标识,以表征第二对象为工艺合格或不合格的可能性。具体的,第二参数值根据第二对象在第二区间中的区间位置来确定,例如,越靠近第二区间的区间边缘的第二对象的质量参数设置为趋近于1的值,越靠近第二区间的区间中间的第二对象的质量参数设置为趋近于0的值,处于第二区间的区间边缘的第二对象A的质量参数为0.8,靠近第二区间的区间中间的第二对象B的质量参数为0.5,而处于第二区间的区间中间的第二对象C的质量参数为0.1,等等;

[0078] 而第三对象中,被拉力测试为工艺合格的第三对象的质量参数设置为第一参数值,如1,被拉力测试为工艺不合格的第三对象的质量参数设置为第三参数值,如0,由此,第二参数值是大于第三参数值且小于第一参数值的。

[0079] 相应的,在工艺数据中添加质量参数之后,对工艺预测模型进行训练时,具体可以通过以下方式实现:

[0080] 首先,对训练对象的工艺数据进行预处理,其中,预处理可以包括有:对工艺数据中的各项工艺参数进行数值归一化、对工艺参数进行缺失补全、对冗余的工艺参数进行去除、对工艺参数中的噪声数据和异常数据进行删除等预处理中的一种或任意多种的组合,进而提高工艺数据的准确率,并保障工艺数据之间的可比性;

[0081] 之后,将工艺数据中的工艺参数作为预先构建的工艺预测模型的输入,将工艺数据中的质量参数作为工艺预测模型的输出,对工艺预测模型进行训练,得到训练完成的工艺预测模型,而工艺预测模型的模型参数中至少包括质量阈值,其中,该质量阈值能够对工艺预测模型输出的预测质量值进行判断,进而确定被预测的对象是否为工艺合格的对象,例如,如果经过工艺预测模型预测出的预测质量值大于或等于质量阈值,则表明工艺合格,如果预测质量值小于质量阈值,那么表明工艺不合格。

[0082] 相应的,在工艺预测模型用于对目标对象进行工艺检测时,在工艺预测模型输出的目标对象的预测质量值大于质量阈值的情况下,表征目标对象为工艺合格的对象;在目标对象的预测质量值小于质量阈值的情况下,表征目标对象为工艺不合格的对象。

[0083] 可见,由于即使工艺标识相同但工艺质量可能不同,如工艺质量参数为1和工艺质量参数值为0.8对应的工艺标识均为工艺合格,因此本实施例中进一步添加与工艺标识相关联的质量参数来参与模型的训练,并以工艺质量来作为模型的输出,进一步再对工艺质量进行阈值判断,而区别于单用工艺标识来进行训练的方式,本实施例中通过细化作为训练样本的精度,来提高模型训练的准确性及后续进行对象工艺检测的准确性。

[0084] 在一种实现方式中,本实施例中在对工艺预测模型进行训练完成之后,还可以对工艺预测模型进行验证,如图5中所示,本实施例中的方法还可以包括以下步骤:

[0085] 步骤104:从第一对象和第三对象中,获得多个验证对象。

[0086] 其中,验证对象与训练对象不同,也就是说,验证对象为从第一对象和第三对象中选取的区别于训练对象的对象,一个对象只能是训练对象,或者只能是验证对象,且验证对象只能从确定为工艺合格或工艺不合格的第一对象和第三对象中选取。

[0087] 步骤105:获得多个验证对象的工艺数据。

[0088] 其中,本实施例中可以在生产工艺的信息系统数据库中读取到多个验证对象的工艺数据。

[0089] 步骤106:将多个验证对象的工艺数据作为工艺预测模型的输入,以得到工艺预测模型输出的验证对象的预测结果。

[0090] 其中,该预测结果表征经过工艺预测模型检测后所得到的表征验证对象可能工艺合格或不合格的结果。

[0091] 步骤107:根据验证对象的预测结果,对模型参数进行修改。

[0092] 其中,本实施例中可以根据验证对象的预测结果和验证对象的工艺数据中的工艺标识来实现对模型参数如质量阈值等的修改,由此,本实施例中通过已知工艺是否合格的验证对象的工艺数据对工艺预测模型进行验证,从而使得经过模型参数修改优化的工艺预测模型的准确性更高,进而提高后续进行其他对象的工艺检测的准确性。

[0093] 具体的,本实施例中可以将验证对象的预测结果与验证对象的工艺数据中的工艺标识进行比对,得到比对结果,该比对结果表征验证对象的预测结果与对应的工艺标识是否相一致,如果比对结果中验证对象的预测结果和对应的工艺标识均表征验证对象工艺合

格或均表征验证对象工艺不合格,那么表征工艺预测模型的模型参数是准确度较高的,如果比对结果中验证对象的预测结果和对应的工艺标识不一致,如预测结果表征验证对象工艺合格而实际上验证对象的工艺标识表征验证对象不合格,那么表明工艺预测模型中的模型参数需要进行优化,此时可以对模型参数进行修改,或者以该验证对象的工艺数据作为训练样本,再次对工艺预测模型进行优化训练,使得模型参数被修改,进而得到优化后的工艺预测模型。

[0094] 在一种实现方式中,为了减轻工艺预测模型的训练量,进而提高工艺预测模型的训练效率,本实施例中可以在对工艺预测模型进行训练之前,首先对工艺数据中的工艺参数进行筛选,例如,首先获得工艺数据中的各个工艺参数与训练对象的质量参数之间的相关度值,该相关度值表征工艺参数对训练对象的质量参数的影响程度,由此,在得到相关度值之后,根据相关度值筛选工艺数据中的工艺参数,例如,筛选相关度值从大到小排名靠前P(P为大于或等1的正整数)位的工艺参数作为训练样本。

[0095] 具体的,本实施例中对于工艺数据中的工艺参数进行相关度值进行计算时,可以利用皮尔逊相关系数算法和/或卡方检验算法对工艺参数与训练对象的质量参数之间的相关度值进行计算,以得到工艺参数与训练对象的质量参数之间的相关度值,再对相关度值进行排序,选取出从大到小排名靠前P位的工艺参数,以剔除工艺数据中的其他工艺参数,再将工艺数据作为训练样本对工艺预测模型进行训练,由此,剔除对训练对象的质量参数无关或影响较小的工艺参数,避免增加工艺预测模型的训练计算量,从而提高工艺预测模型的训练效率。

[0096] 参考图6,为本申请实施例二提供的一种数据处理装置的结构示意图,该装置可以配置在能够进行数据处理的电子设备中,如计算机或服务器等。本实施例中的技术方案主要用于训练出能够对任一对象进行工艺检测的工艺预测模型,以便于提高对象工艺检测的准确性,同时也能够提高对象工艺检测的效率。

[0097] 具体的,本实施例中的装置可以包括以下功能单元:

[0098] 数据获得单元601,用于获得多个训练对象的工艺数据,所述工艺数据包括工艺标识和至少两项工艺参数,所述工艺标识表征所述训练对象是否为工业合格的对象;

[0099] 模型训练单元602,用于以所述训练对象的工艺数据作为预先构建的工艺预测模型的训练样本,对工艺预测模型进行训练,得到训练完成的工艺预测模型;

[0100] 其中,所述工艺预测模型用于以待预测的目标对象的工艺数据作为输入,输出所述目标对象的预测结果,所述预测结果至少表征目标对象是否工艺合格。

[0101] 由上述方案可知,本申请实施例二提供的一种数据处理装置,通过获得多个训练对象的工艺数据,进而以这些工艺数据中的工艺标识和至少两项工艺参数作为预先构建的工艺预测模型的训练样本,对模型进行训练,进而训练完成的工艺预测模型就可以用来对待预测的目标对象进行预测,以预测出目标对象是否工艺合格。可见,本实施例的技术方案中无需对目标对象进行破坏性的测试操作,就可以利用根据训练对象的工艺数据所训练出的工艺预测模型实现对任一对象的工艺检测,而不会存在不合格产品区间包含合格产品的情况,因此,本申请的技术方案能够提高对对象工艺检测的准确性。

[0102] 在一种实现方式中,数据获得单元601在获得多个训练对象的工艺数据时,可以通过以下方式实现:

[0103] 获得多个对象的拉力测试数据,所述拉力测试数据包括对所述多个对象中的部分对象进行拉力测试所产生的数据;根据所述拉力测试数据,获得所述多个对象中的处于第一区间的第一对象、处于第二区间的第二对象和被拉力测试的第三对象;所述第一区间为根据所述拉力测试数据确定为工艺合格的对象区间,所述第二区间为根据所述拉力测试数据确定为工艺不合格的对象区间;从所述第一对象、所述第二对象和所述第三对象中,获得多个训练对象;获得所述多个训练对象的工艺数据。

[0104] 可选的,模型训练单元602在以所述训练对象的工艺数据作为预先构建的工艺预测模型的训练样本,对所述工艺预测模型进行训练之前,还用于:

[0105] 对所述训练对象的工艺数据添加质量参数,所述质量参数的参数值表征所述训练对象的工艺质量;所述质量参数用于对所述工艺预测模型进行训练;

[0106] 其中,所述第一对象的质量参数为第一参数值,所述第二对象的质量参数为第二参数值,所述第二参数值根据所述第二对象在所述第二区间中的区间位置确定,被拉力测试为工艺合格的第三对象的质量参数为所述第一参数值,被拉力测试为工艺不合格的第三对象的质量参数为第三参数值,所述第一参数值大于所述第二参数值,且所述第二参数值大于所述第三参数值。

[0107] 可选的,模型训练单元602以所述训练对象的工艺数据作为预先构建的工艺预测模型的训练样本,对所述工艺预测模型进行训练,得到训练完成的工艺预测模型,具体可以通过以下方式实现:

[0108] 对所述训练对象的工艺数据进行预处理;

[0109] 将所述工艺数据中的工艺参数作为预先构建的工艺预测模型的输入,将所述工艺数据中的质量参数作为所述工艺预测模型的输出,对所述工艺预测模型进行训练,得到训练完成的工艺预测模型;所述工艺预测模型的模型参数中至少包括质量阈值;

[0110] 其中,在所述工艺预测模型输出的所述目标对象的预测质量值大于或等于所述质量阈值的情况下,表征所述目标对象为工艺合格的对象;在所述目标对象的预测质量值小于所述质量阈值的情况下,表征所述目标对象为工艺不合格的对象。

[0111] 在一种实现方式中,模型训练单元602还用于:

[0112] 从所述第一对象和所述第三对象中,获得多个验证对象,所述验证对象与所述训练对象不同;获得所述多个验证对象的工艺数据;将所述多个验证对象的工艺数据作为所述工艺预测模型的输入,以得到所述工艺预测模型输出的所述验证对象的预测结果;根据所述验证对象的预测结果,对所述模型参数进行修改。

[0113] 可选的,模型训练单元602根据所述验证对象的预测结果,对所述模型参数进行修改,具体包括:

[0114] 将所述验证对象的预测结果和所述验证对象的工艺数据中的工艺标识进行比对,得到比对结果;根据所述比对结果,对所述模型参数进行修改。

[0115] 在一种实现方式中,模型训练单元602以所述训练对象的工艺数据作为预先构建的工艺预测模型的训练样本,对所述工艺预测模型进行训练之前,还用于:

[0116] 获得所述工艺数据中的工艺参数与所述训练对象的质量参数之间的相关度值;根据所述相关度值,筛选所述工艺数据中的工艺参数。

[0117] 可选的,所述工艺预测模型为基于回归算法构建的模型。

[0118] 需要说明的是,本实施例中的装置中各单元的具体实现可以参考前文中相应内容,此处不再详述。

[0119] 参考图7,为本申请实施例三提供的一种电子设备的结构示意图,该电子设备可以为能够进行数据处理的电子设备,如计算机或服务器等。本实施例中的技术方案主要用于训练出能够对任一对象进行工艺检测的工艺预测模型,以便于提高对象工艺检测的准确性,同时也能够提高对象工艺检测的效率。

[0120] 具体的,本实施例中的电子设备可以包括以下结构:

[0121] 存储器701,用于存储应用程序和所述应用程序运行所产生的数据;

[0122] 处理器702,用于执行所述应用程序,以实现:获得多个训练对象的工艺数据,所述工艺数据包括工艺标识和至少两项工艺参数,所述工艺标识表征所述训练对象是否为工业合格的对象;以所述训练对象的工艺数据作为预先构建的工艺预测模型的训练样本,对工艺预测模型进行训练,得到训练完成的工艺预测模型;其中,所述工艺预测模型用于以待预测的目标对象的工艺数据作为输入,输出所述目标对象的预测结果,所述预测结果至少表征目标对象是否工艺合格。

[0123] 由上述方案可知,本申请实施例三提供的一种电子设备,通过获得多个训练对象的工艺数据,进而以这些工艺数据中的工艺标识和至少两项工艺参数作为预先构建的工艺预测模型的训练样本,对模型进行训练,进而训练完成的工艺预测模型就可以用来对待预测的目标对象进行预测,以预测出目标对象是否工艺合格。可见,本实施例的技术方案中无需对目标对象进行破坏性的测试操作,就可以利用根据训练对象的工艺数据所训练出的工艺预测模型实现对任一对象的工艺检测,而不会存在不合格产品区间包含合格产品的情况,因此,本申请的技术方案能够提高对对象工艺检测的准确性。

[0124] 需要说明的是,本实施例中处理器的具体实现可以参考前文中相应内容,此处不再详述。

[0125] 以对象为电池,工艺数据为焊接数据为例,对本申请的技术方案进行举例说明,本申请在模型训练的具体实现中有以下流程,如图8中所示:

[0126] 步骤一:数据获取

[0127] 本步骤主要从已有的生产工艺的信息系统中获取m条焊接数据信息,m为大于或等于2的正整数,每条焊接数据包括:焊接压力、焊接时间、焊接振幅、焊接能量等信息,可以用:X1、X2、…、Xn等表示,n为大于或等于2的正整数,还包括:产品合格信息,可以用Y表示。如表1中影响焊接质量因素与产品是否合格的数据所示:

[0128] 表1影响焊接质量因素与产品是否合格数据

X1	X2	...	Xn	Y
x1_1	X2_1	...	Xn_1	Y_1
X1_2	X2_2	...	Xn_2	Y_2
X1_3	X2_3	...	Xn_3	Y_3
X1_4	X2_4	...	Xn_4	Y_4
...
X1_m	X2_m	...	Xn_m	Y_m

[0130] 其中X1、X2、…、Xn主要为影响焊接质量的潜在因素,产品是否合格Y通过质检获

得。由于在实际的生产过程中,工人通过抽样并进行拉力测试的方式进行质检,通过寻找异常区间排查不合格产品,因此生产数据包括异常区间、正常区间(即未进行拉力测试的非异常区间样品)和拉力数据。在这里,可以认为正常区间和拉力数据可以获得焊接质量信息,如表1所示;异常区间不仅包含不合格产品还包含合格产品,因此无法确定其焊接质量信息,如表2的异常区间数据所示,最后一列为空表明无法确定其焊接质量:

[0131] 表2异常区间数据

[0132]

X1	X2	...	Xn	
x1_1	X2_1	...	Xn_1	
X1_2	X2_2	...	Xn_2	
X1_3	X2_3	...	Xn_3	
X1_4	X2_4	...	Xn_4	
...	
X1_m	X2_m	...	Xn_m	

[0133] 步骤二:数据预处理

[0134] 本步骤主要对步骤一种所获取的焊接数据进行预处理,以得到高质量的焊接数据。主要包括以下部分:

[0135] (1) 数据清洗:

[0136] 对存在的缺失、重复、噪音、异常等缺陷的焊接数据进行处理,保证数据的正确性和有效性。

[0137] (2) 异常区间数据处理

[0138] 异常区间两段的产品不合格的概率较高,其余产品的不合格的概率具有很强的不确定性,因此对其分别设定不同的焊接质量信息,例如0.7(靠近区间边缘的焊接质量参数值较大)和0.2(靠近区间中间的焊接质量参数值较小)。处理后的数据形式如表1所示。

[0139] (3) 特征筛选

[0140] 本步骤主要获取焊接数据对于焊接质量相对有效的特征数据,主要包括:

[0141] 1) 删除部分意义重复或无效的特征;

[0142] 2) 根据行业经验筛选特征,选择焊接机理中对焊接质量有重要影响的因素作为特征;

[0143] 3) 通过使用皮尔逊相关系数或卡方检验计算特征与焊接质量的相关性进行特征过滤。

[0144] 步骤三:焊接质量判定

[0145] 1、模型选择:

[0146] 判定焊接产品是否合格为分类问题,为了充分利用异常区间数据,本方案将其转化为回归问题,如首先选择合适的回归算法,例如GBDT、神经网络、随机森林等经典的回归算法,获取焊接质量预测模型(工艺预测模型)。

[0147] 2、模型训练:

[0148] 本步骤主要是生成焊接质量预测模型。主要包括:

[0149] 1) 数据拆分,从正常区间和拉力数据中抽取一定比例的样本构建验证集,将其余的数据作为训练集。

[0150] 2) 使用训练集中通过步骤二的(3)筛选得到的特征作为模型的输入,对应单一的焊接质量作为模型的输出,训练获得焊接质量预测模型M。根据构建模型的特点,使用K-折交叉验证的方法搜索寻优得到较优的模型参数,或通过经验设置的方法对参数进行调整。

[0151] 3、产品合格性的判定:

[0152] 本步骤主要是根据焊接质量实现产品合格性的判定。主要包括:

[0153] 1) 使用步骤三的2得到的模型M,输入训练集和验证集数据预测对应的焊接质量。

[0154] 2) 根据模型对训练集的焊接质量预测结果,查找合适阈值t实现合格产品与不合格产品的判定。

[0155] 3) 根据阈值t和模型对验证集的焊接质量预测,对产品是否合格进行判定,例如焊接质量大于t的判定为合格产品,焊接质量小于t的判定为不合格产品。

[0156] 可见,对于判定焊接产品如电池等是否合格属于分类任务,在样本量不足的情况下,本方案将焊接质量判定问题转化为回归问题。相对于传统焊接质量评估方式,本方案中使用机器学习构建焊接质量判定模型,减少了拉力测试对产品的破坏,降低损失和节约人力成本。而且,本方案中结合排查异常区间的业务经对异常区间数据进行标注,充分利用数据信息;并且结合业务经验与数据挖掘方法分析并筛选影响焊接质量的有效特征,更加符合实际情况。

[0157] 本说明书中各个实施例采用递进的方式描述,每个实施例重点说明的都是与其他实施例的不同之处,各个实施例之间相同相似部分互相参见即可。对于实施例公开的装置而言,由于其与实施例公开的方法相对应,所以描述的比较简单,相关之处参见方法部分说明即可。

[0158] 专业人员还可以进一步意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、计算机软件或者二者的结合来实现,为了清楚地说明硬件和软件的可互换性,在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本申请的范围。

[0159] 结合本文中所公开的实施例描述的方法或算法的步骤可以直接用硬件、处理器执行的软件模块,或者二者的结合来实施。软件模块可以置于随机存储器(RAM)、内存、只读存储器(ROM)、电可编程ROM、电可擦除可编程ROM、寄存器、硬盘、可移动磁盘、CD-ROM、或技术领域内所公知的任意其它形式的存储介质中。

[0160] 对所公开的实施例的上述说明,使本领域专业技术人员能够实现或使用本申请。对这些实施例的多种修改对本领域的专业技术人员来说将是显而易见的,本文中所定义的一般原理可以在不脱离本申请的精神或范围的情况下,在其它实施例中实现。因此,本申请将不会被限制于本文所示的这些实施例,而是要符合与本文所公开的原理和新颖特点相一致的最宽的范围。

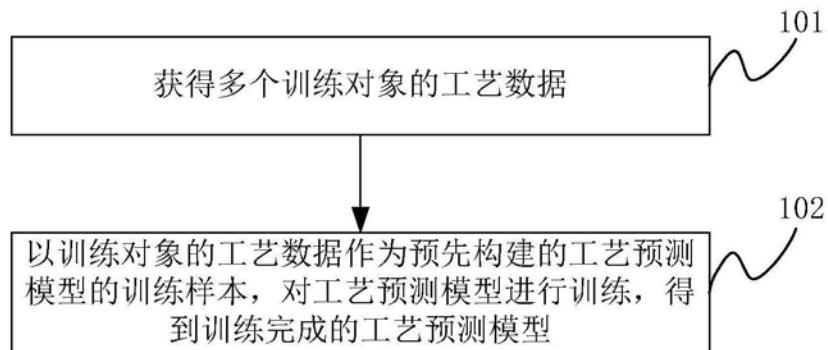


图1

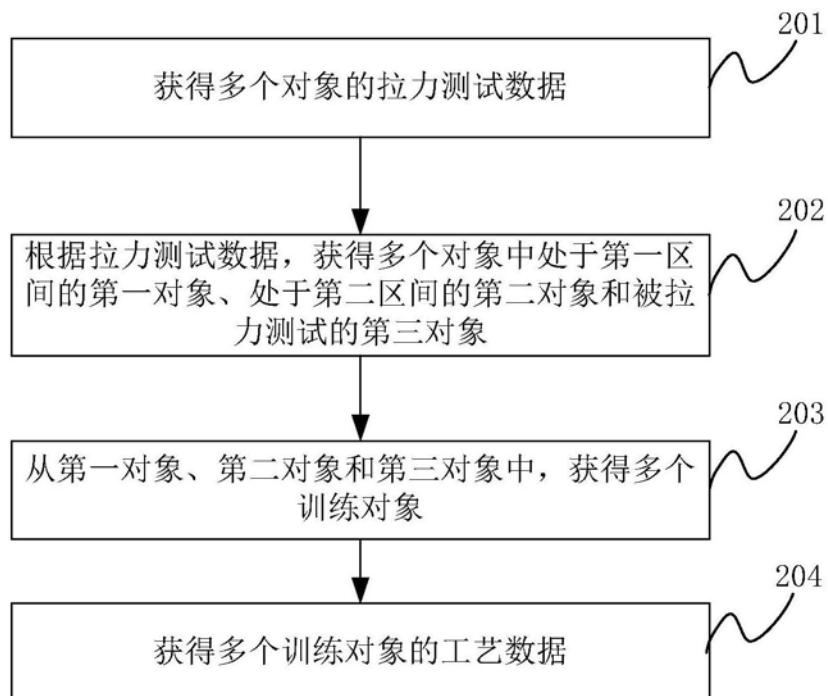


图2

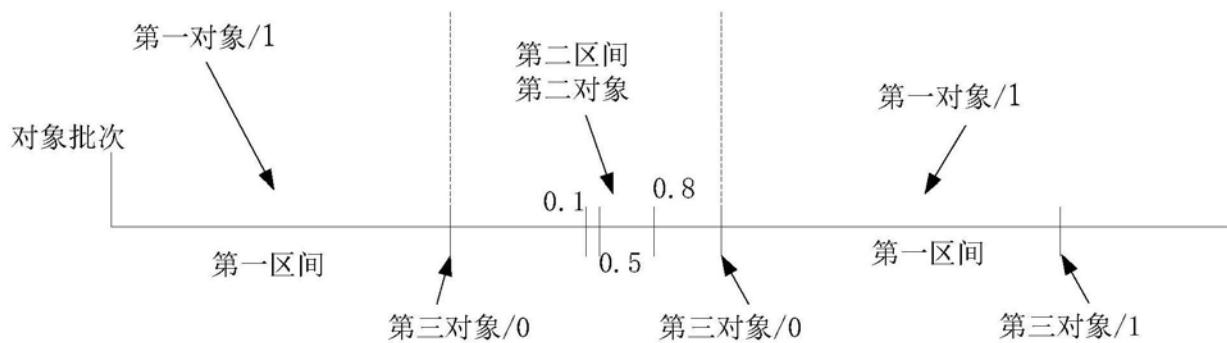


图3

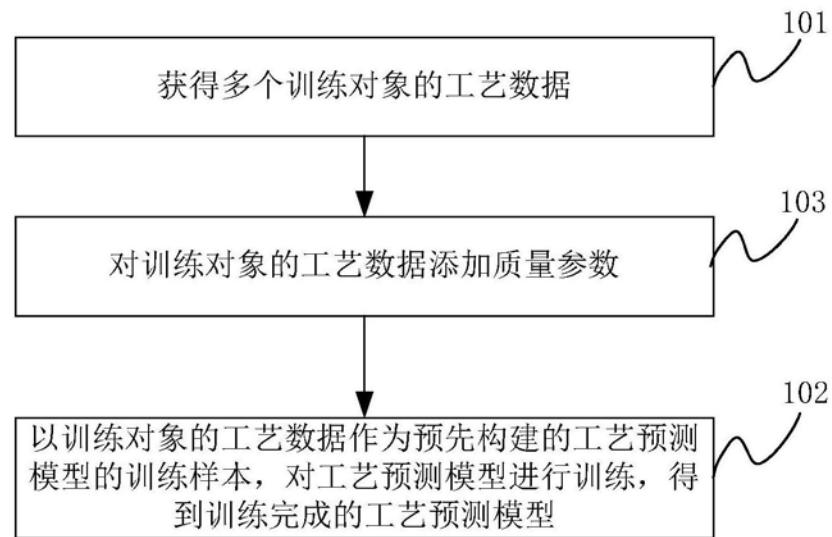


图4

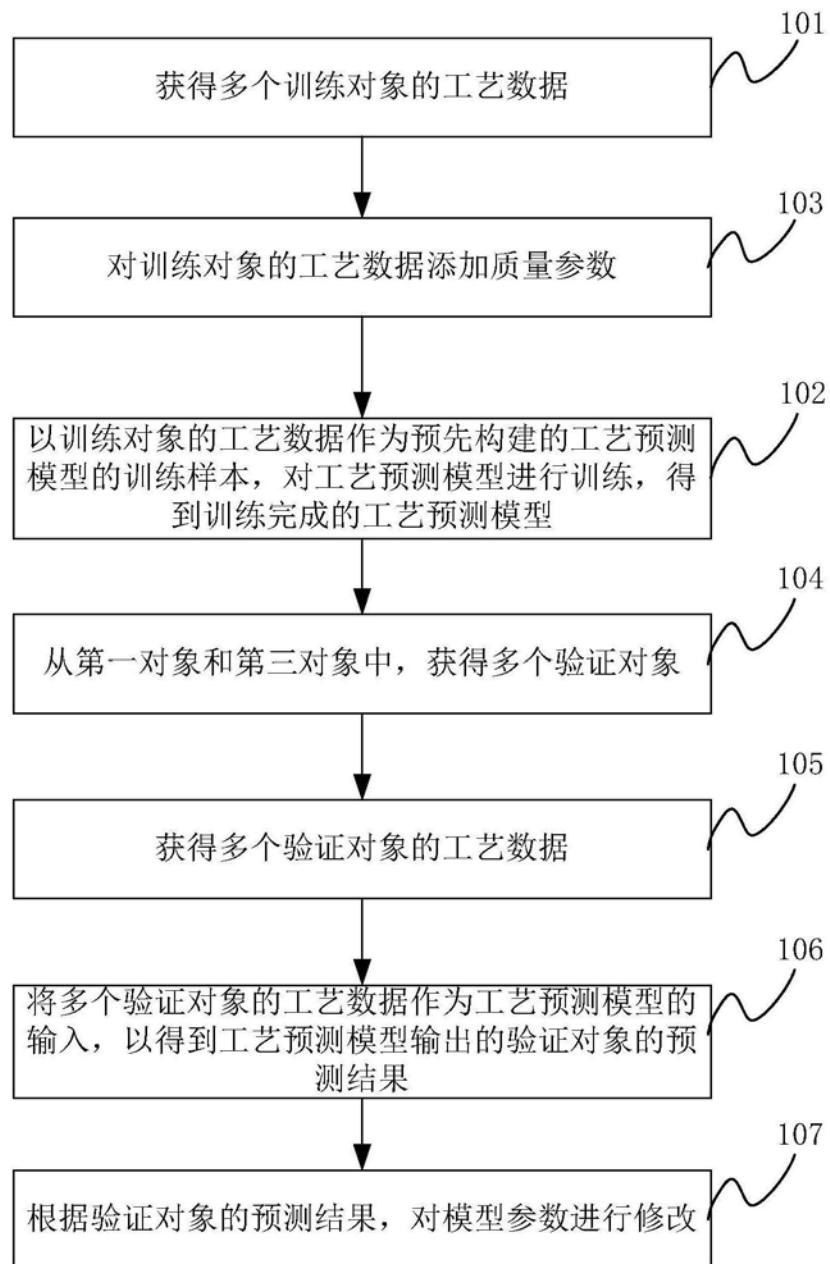


图5

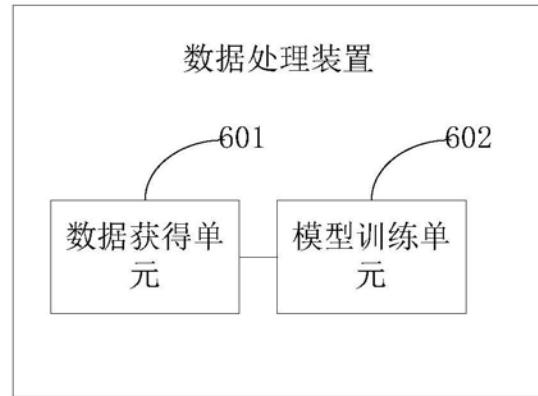


图6

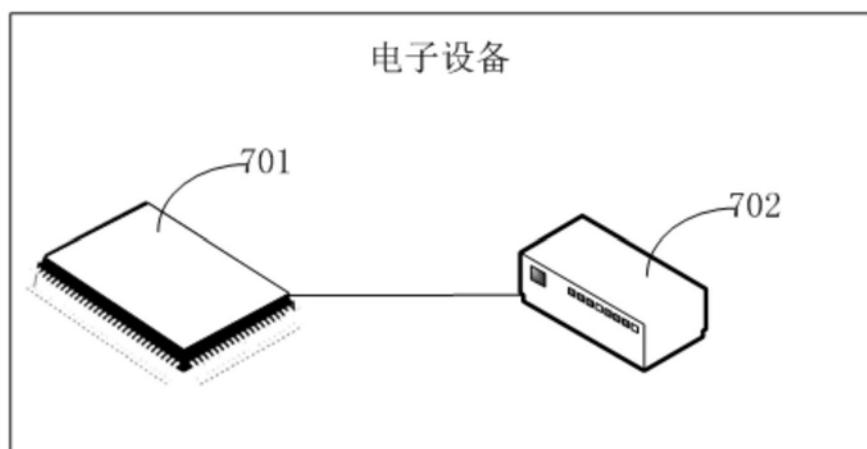


图7

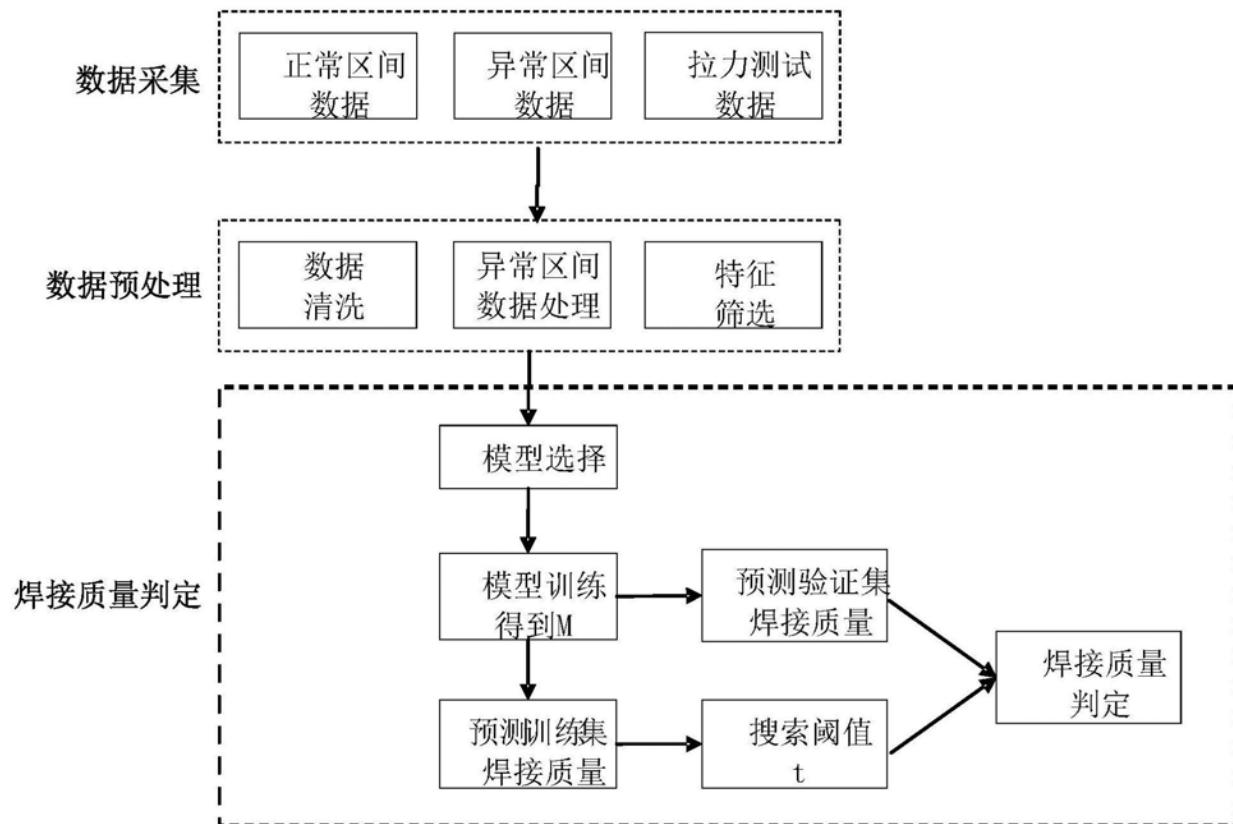


图8