



(12) 发明专利申请

(10) 申请公布号 CN 114429238 A

(43) 申请公布日 2022. 05. 03

(21) 申请号 202111477655.4

(22) 申请日 2021.12.06

(71) 申请人 石河子大学

地址 832003 新疆维吾尔自治区石河子市
北五路石河子大学北苑新区机械电气
工程学院

(72) 发明人 李春杨 夏博 王宇

(51) Int. Cl.

- G06Q 10/04 (2012.01)
- G06Q 10/06 (2012.01)
- G06Q 10/00 (2012.01)
- G06Q 50/06 (2012.01)
- G06N 3/04 (2006.01)
- G06N 3/08 (2006.01)

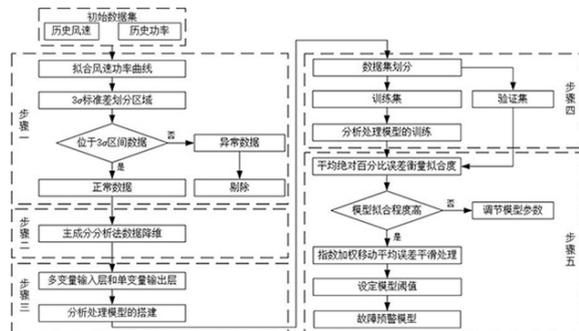
权利要求书2页 说明书3页 附图3页

(54) 发明名称

一种基于时空特征提取的风电机组故障预警方法

(57) 摘要

本发明公开了一种基于时空特征提取的风电机组故障预警方法,步骤如下:先通过风电机组监控和数据采集系统来获取风电机组数据集,进行数据预处理,检测并剔除异常点;再利用主成分分析方法对风电机组运行数据进行降维处理;然后将数据集分割为多变量的输入层和单变量的输出层,保留数据的时序相关性和空间相关性,进行风电机组监控和数据采集系统数据预测神经网络模型搭建与训练;最后使用平均绝对百分比误差指标衡量模型对数据的拟合程度,保留最优模型,设定模型阈值,得到故障预警模型。本发明综合考虑了风电机组数据的时间特征和空间特征,更加全面、准确的反映风电机组的潜在状态,提高了对风电机组故障预警能力。



1. 一种基于时空特征提取的风电机组故障预警方法,其特征在于,包括以下步骤:步骤1:风电机组监控和数据采集系统数据预处理,在切入风速和切出风速之间划出若干个风速区间,将所有数据点按照其风速大小归入各个风速区间,然后对各区间内的数据点进行统计分析,得到风速和功率的期望值,对风电机组风速-功率曲线进行拟合,根据风机运行数据统计差值分布的标准差 σ ,以 3σ 划分标准,对异常点进行检测和剔除;步骤2:风电机组监控和数据采集系统数据降维处理,为提高模型的计算效率,需要对风电机组运行数据进行降维处理,采用主成分分析方法,将风电机组数据参数通过正交变换为一组线性不相关的变量,保证降维后的信息贡献率在95%以上;步骤3:风电机组监控和数据采集系统数据预测神经网络模型搭建,首先在数据处理层面保留时序相关性,然后将数据分割为多变量的输入层和单变量的输出层,根据输入层的大小确定滑窗的宽度,进而保留数据层的空间相关性,将2维数据转变为4维,其中包含输入层,输出层,卷积核和步长,模型的卷积层为32层,选择ReLU为激活函数,池化层为最大池化层,然后展平为线性层,线性层的随机丢弃率为0.5,最后的输出层为1;步骤4:训练风电机组监控和数据采集系统数据预测神经网络模型,将数据分为60%的训练集和40%的验证集,用验证集判断模型的好坏,确定最优模型,最终确定为选用Xavier初始化方法初始化模型的参数,选用Adam优化器对模型的参数进行更新,确定损失函数为平方损失函数,采用小批量梯度下降法,批量大小为64,迭代50个周期,学习率为0.001;步骤5:使用平均绝对百分比误差指标衡量模型对数据的拟合程度,通过调整模型结构和超参数使模型得到最佳的效果,最后使用指数加权移动平均值方法对误差进行平滑处理,设定模型阈值,得到故障预警模型。

2. 如权利要求1所述的一种基于时空特征提取的风电机组故障预警方法,其特征在于:步骤1中所述的“ 3σ 划分标准”是指以拟合的风电机组风速功率曲线为基准,在 3σ 的范围内获取风电场数据,并剔除不在该范围内的数据。

3. 如权利要求1所述的一种基于时空特征提取的风电机组故障预警方法,其特征在于:步骤2中所述的“主成分分析方法”是指首先对数据进行标准化处理,计算样本指标的均值和标准差,然后计算每个样本指标的相关系数矩阵,然后根据相关系数矩阵的特征值和特征向量,组成新的主成分,最后计算主成分的信息贡献率,累计计算所有主成分的累计贡献率,选择相应的主成分作为输入变量。

4. 如权利要求1所述的一种基于时空特征提取的风电机组故障预警方法,其特征在于:步骤3中所述的“多变量的输入层和单变量的输出层”是指对于同一台风电机组而言,在数据降维后,具有一定数量的特征属性,其特征属性中含有风电机组这段时间内的时间信息和不同空间位置的特征信息,它们作为模型多变量的输入层,我们将风电机组温度信息作为模型的输出值(预测值),这个输出值与前一段时间内的温度信息有关,所以为单变量的输出层。

5. 如权利要求1所述的一种基于时空特征提取的风电机组故障预警方法,其特征在于:步骤4中所述的“训练风电机组监控和数据采集系统数据神经网络模型”具体分为五步,分别为随机初始化权重,向前传播得到所有样本,计算损失函数,反向传播计算偏导,更新权重,其使用的方法分别对应为Xavier初始化,小批量梯度下降,平方损失函数,Adam优化器,覆盖更新。

6. 如权利要求1所述的一种基于时空特征提取的风电机组故障预警方法,其特征在于:

步骤5中所述的“平均绝对百分比误差”是指用这组数据的实际值减去模型的预测值,然后除以实际值,最后取这组比值的平均数来判断模型拟合的好坏。

7. 如权利要求1所述的一种基于时空特征提取的风电机组故障预警方法,其特征在于:步骤5中所述的“指数加权移动平均值”是指以指数式递减加权的移动平均,温度预测值的加权影响力随时间成指数式递减,规定上一时刻的预测值的影响大小为0.95。

一种基于时空特征提取的风电机组故障预警方法

技术领域

[0001] 本发明提供一种基于时空特征提取的风电机组故障预警方法,属于质量监测领域。

背景技术

[0002] 风能作为重要的可再生能源,近几十年来,全球风能使用规模迅速增长,陆上和海上风力发电机组发电容量的不断增加,由于风力发电机组潜在故障的维修成本巨大,因此必须开发有效且可靠的风力发电机组故障预警方法,以便降低风电场的运营和维护成本。然而传统的风力机故障预警技术大多只使用单一类型的信号作为故障特征,其性能可能受到此类信号特征的限制。为解决这一问题,设计了一种基于时空特征提取的风电机组故障预警方法,从而更加全面、准确的反映风电机组的潜在状态,提高对风电机组故障预警能力。

发明内容

[0003] 本发明针对传统风电机组故障预警技术故障特征单一的问题,设计了一种基于时空特征提取的风电机组故障预警方法。

[0004] 为实现上述目的,本发明型采用以下五个步骤,如附图1所示。

[0005] 步骤1:风电机组监控和数据采集系统数据预处理,在切入风速和切出风速之间划出若干个风速区间,将所有数据点按照其风速大小归入各个风速区间,然后对各区间内的数据点进行统计分析,得到风速和功率的期望值,对风电机组风速-功率曲线进行拟合,根据风机运行数据统计差值分布的标准差 σ ,以 3σ 划分标准,对异常点进行检测和剔除。

步骤2:风电机组监控和数据采集系统数据降维处理,为提高模型的计算效率,需要对风电机组运行数据进行降维处理,采用主成分分析方法,将风电机组数据参数通过正交变换为一组线性不相关的变量,保证降维后的信息贡献率在95%以上。

[0006] 步骤3:风电机组监控和数据采集系统数据预测神经网络模型搭建,首先在数据处理层面保留时序相关性,然后将数据分割为多变量的输入层和单变量的输出层,根据输入层的大小确定滑窗的宽度,进而保留数据层的空间相关性。将2维数据转变为4维,其中包含输入层,输出层,卷积核和步长。模型的卷积层为32层,选择ReLU为激活函数,池化层为最大池化层,然后展平为线性层,线性层的随机丢弃率为0.5,最后的输出层为1。

[0007] 步骤4:训练风电机组监控和数据采集系统数据预测神经网络模型,将数据分为60%的训练集和40%的验证集。用训练集训练模型,用验证集判断模型的好坏,确定最优模型。最终确定为选用Xavier初始化方法初始化模型的参数,选用Adam优化器对模型的参数进行更新,确定损失函数为平方损失函数,采用小批量梯度下降法,批量大小为64,迭代50个周期,学习率为0.001。

[0008] 步骤5:使用平均绝对百分比误差指标衡量模型对数据的拟合程度,通过调整模型结构和超参数使模型得到最佳的效果,最后使用指数加权移动平均值方法对误差进行平滑

处理,设定模型阈值,得到故障预警模型。

[0009] 其中,在步骤1中所述的“ 3σ 划分标准”如附图2所示。是指根据拟合曲线确定 3σ 区间,对数据进行扩充,不保证风电机组运行数据的准确性。

[0010] 其中,在步骤2中所述的“主成分分析方法”,具体方法如下:

2-1:对原始数据进行标准化处理,将各指标值 x_{ij} 转换成标准化指标 \tilde{x}_{ij} 。

$$\tilde{x}_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}, (i = 1, 2, \dots, n; j = 1, 2, \dots, m)$$

[0011] 其中 $\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}$, $s_j = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}$, ($j = 1, 2, \dots, m$)。

[0012] 即 \bar{x}_j , s_j 分别为第 j 个指标的样本均值和标准差。

[0013] 2-2:计算相关系数矩阵 R 。

[0014] 相关系数矩阵 $R = (r_{ij})_{m \times m}$ 。

$$r_{ij} = \frac{\sum_{k=1}^n \tilde{x}_{ki} \cdot \tilde{x}_{kj}}{n-1}, (i, j = 1, 2, \dots, m)$$

[0015] 式中 $r_{ii} = 1$, $r_{ij} = r_{ji}$, r_{ij} 是第 i 个指标与第 j 个指标的相关系数。

[0016] 2-3:计算特征值和特征向量。

[0017] 特征值要由大到小排序。

[0018] 计算相关系数矩阵 R 的特征值 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$,及对应的特征向量

u_1, u_2, \dots, u_m , 其中 $u_j = (u_{1j}, u_{2j}, \dots, u_{nj})^T$, 由特征向量组成 m 个新的指标变量。

$$\begin{cases} y_1 = u_{11}\tilde{x}_1 + u_{21}\tilde{x}_2 + \dots + u_{n1}\tilde{x}_n \\ y_2 = u_{12}\tilde{x}_1 + u_{22}\tilde{x}_2 + \dots + u_{n2}\tilde{x}_n \\ \dots \dots \\ y_m = u_{1m}\tilde{x}_1 + u_{2m}\tilde{x}_2 + \dots + u_{nm}\tilde{x}_n \end{cases}$$

[0019] 式中 y_1 是第一主成分, y_2 是第二主成分, \dots , y_m 是第 m 个主成分。

[0020] 2-4:选择 P ($p \leq m$)个主成分,计算综合评价值。

[0021] 计算特征值 λ_j ($j = 1, 2, \dots, m$)的信息贡献率和累计贡献率。称为主成分 y_j 的信息贡献率。

$$b_j = \frac{\lambda_j}{\sum_{k=1}^m \lambda_k} (j = 1, 2, \dots, m)$$

[0022] 计算综合得分。

$$Z = \sum_{j=1}^p b_j y_j$$

[0023] 其中 b_j 为第 j 个主成分的信息贡献率。

[0024] 其中,在步骤3中所述的“风电机组监控和数据采集系统数据神经网络模型”,其构成如附图3所示。所述的“多变量的输入层和单变量的输出层”,其构成如附图4所示。

[0025] 其中,在步骤4中所述的“训练风电机组监控和数据采集系统数据神经网络模型”,其步骤如附图5所示。

[0026] 其中,在步骤5中所述的“平均绝对百分比误差”,其具体计算方法如下:

平均绝对百分比误差(Mean absolute percentage error,MAPE)用于计算预测结果与真实结果之间的偏离度。

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{\text{actual}(t) - \text{forecast}(t)}{\text{actual}(t)} \right| \times 100\%$$

[0027] 以判断风电机组模型的拟合度为例, $\text{actual}(t)$ 表示实际的风电机组温度值, $\text{forecast}(t)$ 代表模型预测的风电机组温度值。其中MAPE的值越小,说明预测模型拥有更好的精确度。

[0028] 其中,在步骤5中所述的“指数加权移动平均值”,其具体算法如下:

“指数加权移动平均值”是以指数式递减加权的移动平均,各数值的加权影响力随时间呈指数式递减,时间越靠近当前时刻的数据加权影响力越大。

$$v_t = \beta v_{t-1} + (1 - \beta) \theta_t$$

[0029] 以风电机组温度数据预测为例,公式中 θ_t 为 t 时刻的实际温度值;系数 β 表示加权下降的快慢,值越小权重下降的越快; v_t 为 t 时刻的指数加权移动平均值。一般情况下取 $\beta \geq 0.9$,此算法中我们取 $\beta = 0.95$ 。

附图说明

图1是本发明所述方法的流程框图

图2是步骤1中所述的“ 3σ 划分标准”图

图3是风电机组监控和数据采集系统数据神经网络模型图

图4是神经网络模型多变量的输入层和单变量的输出层图

图5是训练风电机组监控和数据采集系统数据神经网络模型图。

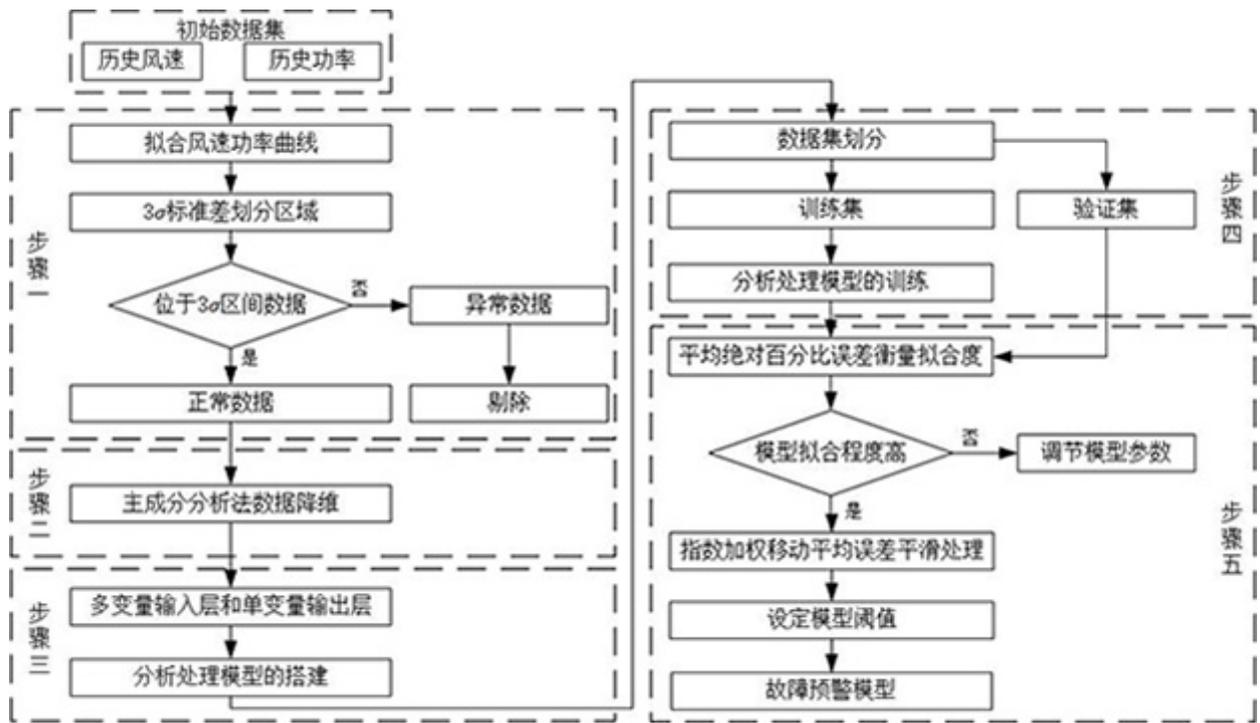


图1

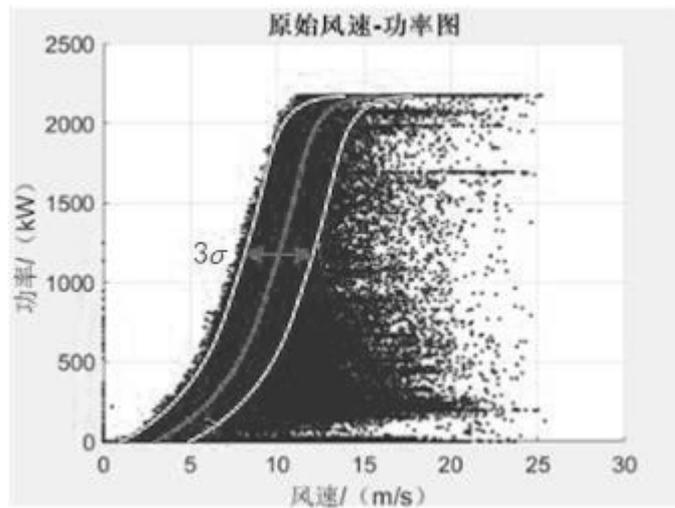


图2

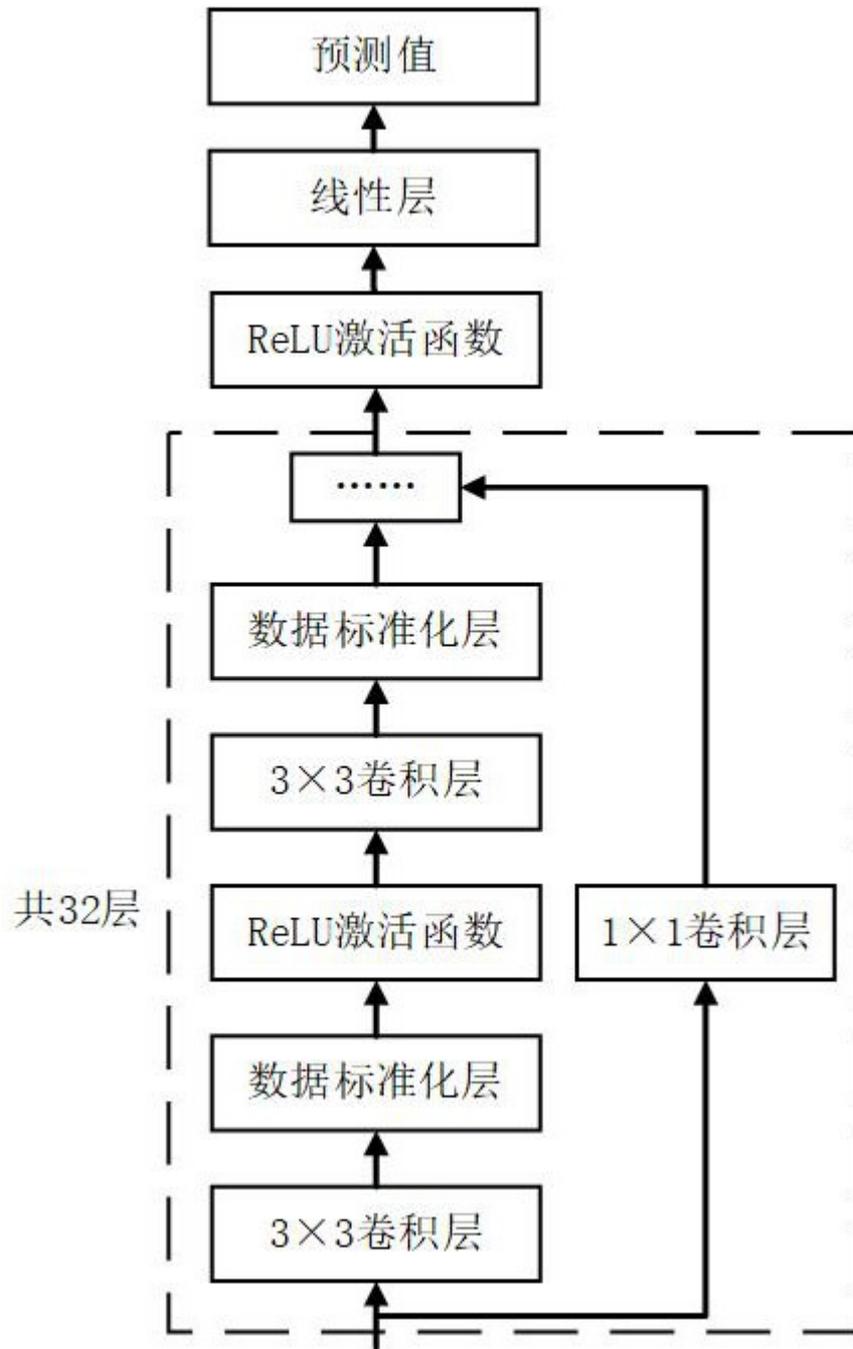


图3



图4

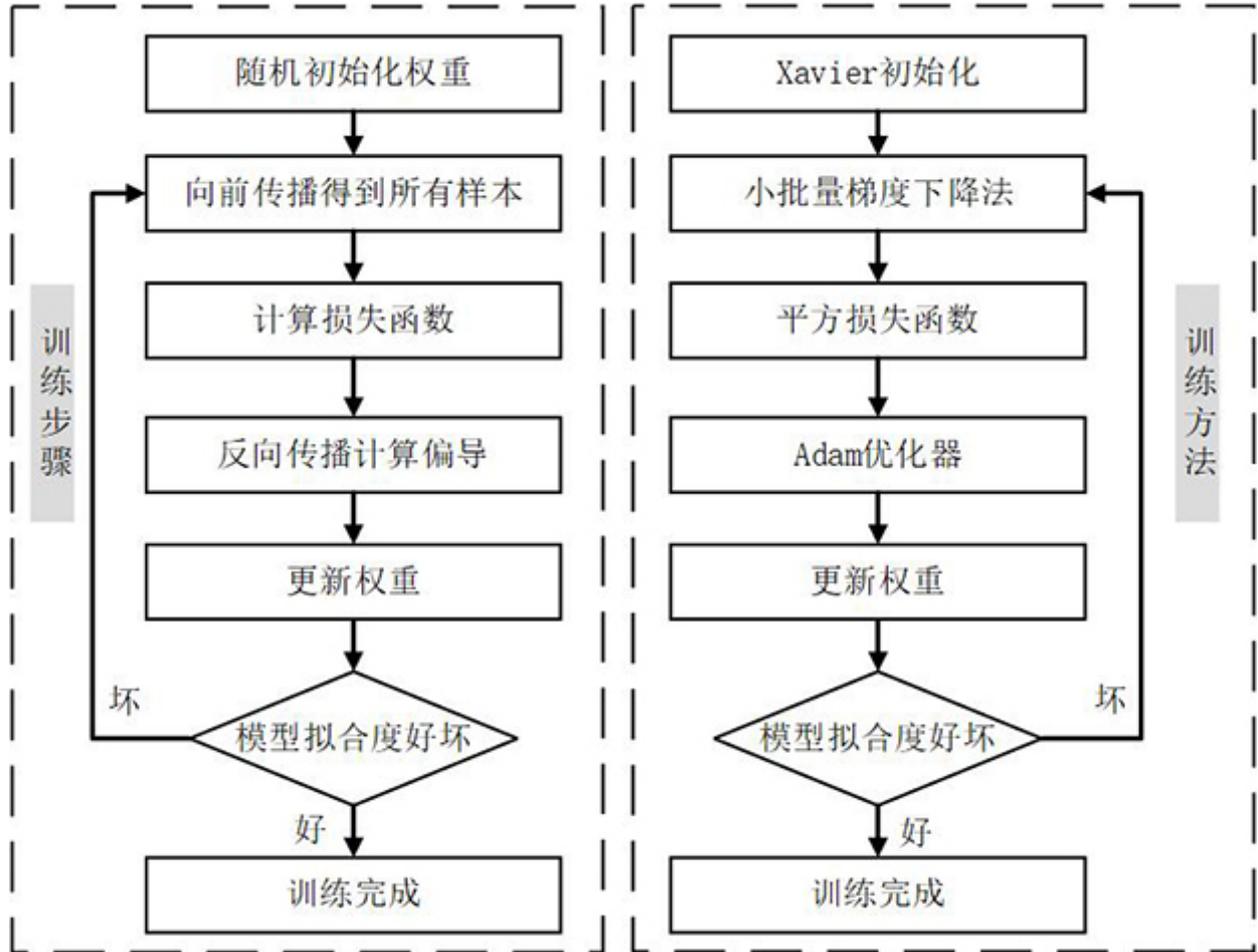


图5