

doi: 10.7690/bgzdh.2025.11.015

基于深度学习的计算机设备在线监测与故障诊断方法

李懿琼

(上海工艺美术职业学院教务处, 上海 201808)

摘要: 针对计算机设备运行故障检测, 提出基于嵌入式技术检测方法。利用无线传感器采集计算机设备运行数据, 构建 MobileNet-SSD 网络模型, 通过卷积层与 Batch Normalization 层的融合实现构建网络模型的改进, 采用剪枝微调优化方法保证剪枝前模型信息保留剪枝后模型, 将其移植到嵌入式设备中, 将获取的 3 个故障特征量作为模型输入, 完成计算机设备运行故障检测。实验结果表明: 该方法能够完成计算机设备故障特征量的提取、获取设备故障检测结果; 收敛速率快、训练损失低于 0.01; 可实现故障检测模型的压缩, 参数量下降明显, mAP 指标波动较小。

关键词: 嵌入式; 深度学习; 计算机设备; 故障检测; 无线传感器; 剪枝微调优化

中图分类号: TP277.3 文献标志码: A

On-line Monitoring and Fault Diagnosis Method of Computer Equipment Based on Deep Learning

Li Yiqiong

(Academic Affairs Office of Shanghai Vocational College of Arts and Crafts, Shanghai 201808, China)

Abstract: In view of the fault detection of computer equipment, a detection method based on embedded technology is proposed. Acquire operation data of computer equipment by utilizing a wireless sensor, constructing a MobileNet-SSD network model, realizing that improvement of the constructed network model through the fusion of a convolution layer and a Batch Normalization layer, ensuring that model information before pruning is reserved by adopting a pruning fine-tuning optimization method, transplanting the pruned model to embedded equipment, The three fault characteristic quantities are used as the input of the model to complete the operation fault detection of the computer equipment. The experimental results show that the method can complete the extraction of computer equipment fault features and obtain equipment fault detection results; the convergence rate is fast, and the training loss is less than 0.01; the compression of fault detection model can be realized, the number of parameters decreases significantly, and the fluctuation of mAP index is small.

Keywords: embedded; deep learning; computer equipment; fault detection; wireless sensor; pruning and fine-tuning optimization

0 引言

计算机自问世以来, 为社会发展和人类进步作出了极大的贡献, 现已成为人们工作、生活中不可或缺的工具^[1]。计算机在长期使用过程中, 必然会出现设备运行故障问题, 但对设备的维护保养并非是任何人都能完成的, 需要具有一定技术水平的专业维修人员方能查出计算机设备的具体故障原因^[2]。计算机设备一旦出现运行故障, 必然会对工作、学习等产生一定影响, 轻则延误工作、学习进度, 重则可能会带来较大经济损失^[3]; 因此, 采用智能化手段对计算机设备运行故障进行检测, 是降低故障危害, 延长计算机使用寿命以及实现正常工作、学习的有效保证^[4~6]。技术变革推动了机器学习技术的快速发展, 利用传感器对设备运行数据进行

特征采集后, 通过机器学习能够实现设备故障特征的有效提取, 从而完成设备故障的分类识别与检测。深度学习作为其重要分支, 能够更深层次地获取数据特征, 在目标识别等领域受到高度关注^[7~8]。

包萍等^[9]利用生成对抗网络来扩充样本数据, 以解决故障检测过程中的数据不均衡问题, 通过样本数据训练自编码深度学习网, 以获取其深度特征, 完成设备故障诊断, 该方法可有效解决数据不对称等问题, 但因构建的故障诊断网络模型结构的不合理, 降低了设备故障诊断效果; 宋浏阳等^[10]为实现智能设备运行故障的高精度识别, 提出采用动态统计滤波对设备信号进行处理, 通过综合信息量指标对处理结果进行优化, 放大各类状态信号间的特征差异, 通过建立 DCNN 网络模型获取设备故

收稿日期: 2024-10-07; 修回日期: 2024-11-17

第一作者: 李懿琼(1982—), 女, 上海人。

障特征，实现故障类型的识别。该方法所用模型计算量大、参数过多，难以在嵌入式设备故障检测中取得较好的检测效果。计算机设备具有资源有限的缺陷，难以将结构复杂、计算量大的深度学习模型部署在其上，实现故障的智能化检测。

针对以上问题，笔者引入了深度学习算法与卷积神经网络(CNN)，通过 CNN 从原始数据中提取高层次的特征，减少了特征工程的复杂度，提高了特征提取的准确性和效率，并且利用深度学习算法强大的非线性建模能力，能够学习复杂的故障模式和设备行为，适应多种故障类型和工况变化，搭建深度可分解网络模型，通过对其作剪枝处理实现模型参数等有效压缩后，再将其移植到嵌入式设备中，在保证模型压缩率的同时，提高计算机设备故障检测效果，确保其正常、安全运行。

1 计算机设备运行故障检测

1.1 计算机设备运行数据采集

计算机设备运行故障检测有效实施的前提是获取其运行数据信息。笔者利用无线传感器获取计算机设备运行数据信息，将其分别安装在各个计算机设备上，以并联方式与计算机设备电源电路建立连接，依据实际需求完成无线传感器关键参数的设定，如扫描周期、频率等，利用 I/S 通信协议实现计算机设备运行数据的传输，并保存在无线传感器数据库中^[11]。具体计算机设备运行数据采集系统框架如图 1 所示。

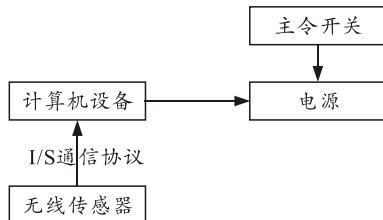


图 1 计算机设备运行数据采集系统框架

1.2 计算机设备故障样本提取

计算机设备在持续工作过程中，不可避免会产生故障，通常会表现为设备温度偏高、设备运行速率大幅降低，甚至会引发计算机设备电源短路或断路现象，故笔者将温度、运行速度、电源电流作为表征计算机设备故障的样本数据^[12]。由于采集到的样本数据可能会存在不对称的问题，因此，需对其进行预处理，具体步骤为：

第 1 步：通过无线传感器获取计算机设备运行数据后，确定需要采集的运行数据类型，例如 CPU

温度、内存使用率、磁盘 IO 等。选用合适的无线传感器，如 WiFi/Bluetooth 传感器、Zigbee 传感器、LoRa 传感器等。根据需要，选择传感器的工作频段和传输距离等参数。将无线传感器固定在计算机设备上，例如 CPU 散热器上或 USB 接口处，使其能够实时采集设备的运行数据。处理和过滤采集的运行数据，消除噪声、滤波以及数据压缩等操作，尽可能保证采集数据的准确性。并对其进行清洗和处理，以消除数据中的噪声和异常值^[13]，去除其中的无效、不完整以及冗余数据信息，再依据温度、运行速度、电流类别对余下数据作类别划分，完成计算机设备故障样本数据的提取，经过归一化处理，使计算机设备故障样本数据所属区间均为[-1, 1]。将已处理的数据通过无线网络传输到数据处理中心或云平台等，可以通过 MQTT 协议或 HTTP 等网络协议进行传输。在数据处理中心或云平台等设备上，可以使用各种机器学习方法和工具进行数据分析和故障检测。

第 2 步：对提取的 3 个计算机设备故障特征样本的上下限值进行设定，其中下限分别为 3 个特征量的初始值，上限为各特征量允许的最大值，依据实际需求，对 3 个特征量的阈值进行分别设置。

1.3 计算机设备运行故障检测

1.3.1 嵌入式深度学习网络结构

MobileNet 通过深度可分解卷积实现网络结构的优化^[14-15]，对标准卷积作分解处理，在获取深度卷积的同时，也可得到点卷积。设定 F 为轻量级卷积神经网络 MobileNet 的输入， $D_F \times D_F \times M$ 为其维度信息，对于 F ，其宽度、高度均为 D_F ，输入层中含有通道总量为 M 。在标准卷积操作中，通过 $D_K \times D_K \times M \times N$ 来描述标准卷积 K 中的总参数量， D_K 为其维数，输出通道总量表示为 N 。 C 为标准卷积下的计算代价：

$$C = D_K \times D_K \times M \times N \times D_F \times D_F \quad (1)$$

在深度可分解卷积下， K 分解成 $D_K \times D_K \times 1 \times M$ 的深度卷积和 $1 \times 1 \times M \times N$ 的点卷积。 C' 的表达式为：

$$C' = D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F \quad (2)$$

用式(2)除以式(1)，可得：

$$\frac{C'}{C} = \frac{D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F}{D_K \times D_K \times M \times N \times D_F \times D_F} = \frac{N^{-1} + (D_K^2)^{-1}}{N^{-1} + (D_K^2)^{-1}} \quad (3)$$

图 2 为 MobileNet 的深度可分解卷积网络结构。

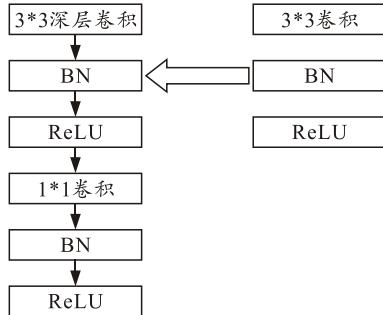


图 2 深度可分解卷积结构

分解后的深层卷积层、点卷积层均连接了一个 BN 层以及 ReLU 函数，其目的是降低计算难度，减小卷积参数。

1.3.2 合并 Batch Normalization 层设计

在 MobileNet-SSD 网络中，Batch Normalization 层的应用可起到提高网络训练效率，加快网络收敛的作用。对于传统卷积神经网络，通常在卷积层或全连接层后端连接 Batch Normalization 层，利用其对数据作归一化处理，可避免网络训练过程中产生的梯度不断降低问题，并有实现训练快速拟合的作用。该层的应用虽然能使传统卷积神经网络达到很好的训练效果，但增大了运算量，从而使得网络运算效率降低，另外，Batch Normalization 层的使用，加大了内存占用率。为此，笔者对 Batch Normalization 层进行优化处理，将其部署于卷积层中，以改善 MobileNet-SSD 网络对计算机设备运行故障的检测效率。

用 X 表示各卷积层的输入，权重通过 W 表示、偏置量为 b ，执行卷积操作，其计算公式为：

$$z = W \times X + b. \quad (4)$$

式中 z 为卷积处理提取的计算机设备运行故障特征。

卷积结果输入到 Batch Normalization 层后，需完成 2 个处理过程。

1) 归一化过程，其公式为：

$$z' = (X - \mu) / \sqrt{\sigma^2 + \varepsilon}. \quad (5)$$

式中： μ 、 σ 分别为均值、方差； ε 为常数，设定其值为很小值，其目的是避免分母等于 0； z' 为归一化的特征。

2) 缩放处理过程为：

$$f(z') = \gamma z' + \beta. \quad (6)$$

式中： γ 为缩放系数； β 为偏置量。

将卷积层与 Batch Normalization 层作融合处理，可得：

$$G = \gamma \times \frac{(W \times X + b) - \mu}{\sqrt{\sigma^2 + \varepsilon}} + \beta. \quad (7)$$

融合后的权重通过下式计算：

$$W' = \frac{\gamma}{\sqrt{\sigma^2 + \varepsilon}} \times W. \quad (8)$$

融合后的偏置计算公式为：

$$b' = \frac{\gamma}{\sqrt{\sigma^2 + \varepsilon}} (b - \mu) + \beta. \quad (9)$$

1.3.3 MobileNet-SSD 网络模型剪枝

为实现计算机设备运行故障的高精度检测，需对部署于嵌入式设备的改进 MobileNet-SSD 网络模型进行剪枝处理，以减少模型参数以及复杂运算对计算机内存的占用。

对于改进后 MobileNet-SSD 原始网络模型， W' 为第 i 个卷积层的权重， n_i 为该层卷积核总数。对其作剪枝处理后，对于第 i 个卷积层， W''_i 为其权重通过， n'_i 为剪枝后的卷积核总数，设定 θ_i 为该层剪枝率，通过下式进行计算：

$$\theta_i = (n_i - n'_i) / n_i. \quad (10)$$

设定 α 为阈值，通过将其与 θ_i 进行比较，以判断剪枝处理后的改进 MobileNet-SSD 网络模型中含有原网络模型信息量的大小，当 $\theta_i > \alpha$ 时，表明原始 MobileNet-SSD 网络模型中过多重要信息被剪除，将严重影响剪枝后改进 MobileNet-SSD 网络模型对计算机设备运行故障检测效果，因此需通过信息优化处理，以保证原始改进 MobileNet-SSD 网络模型中重要信息的完整性。

满足 $\theta_j > \alpha$ 条件的卷积层， W'_j 为其剪枝前的权重，有 $W'_j \in R^{m \times n}$ ， W''_j 为剪枝后的权重，对其作信息优化，以满足条件 $W'_j W'^T \approx W''_j W''^T$ ，即：

$$\arg \min \|W'_j W'^T - W''_j W''^T\|_2. \quad (11)$$

通过对式(11)进行优化计算，即可实现剪枝后改进 MobileNet-SSD 网络模型信息完整性问题的有效解决。笔者采用优化后的 Matrix Sketching 算法对式(11)进行寻优计算，采用基于全局信息的剪枝微调优化方法对改进 MobileNet-SSD 网络模型进行微调，具体过程如下：

第 1 步：将原始改进 MobileNet-SSD 网络模型的第 i 个卷积层的权重 $W'_i \in R^{m \times n}$ 作为算法的输入信息。

第 2 步：对 $\mathbf{W}'_j \in R^{l \times n}$ 的全零矩阵进行构造。

第 3 步：在 \mathbf{W}'_j 矩阵中加入 \mathbf{W}'^k_i ，当满足无全零行条件后停止添加。

第 4 步：对处理后的矩阵作奇异值分解，即 $[U, S, V] \leftarrow SVD(\mathbf{W}'_j)$ 。

第 5 步：对 S 矩阵中的上 1 半数值进行提取，则有 $\delta = S_{1/2}^2$ 。

第 6 步：对 $\bar{S} \leftarrow \max \sqrt{(S^2 - I_l \delta, 0)}$ 进行运算。

第 7 步：对 $\mathbf{W}'_j \leftarrow U\bar{S}$ 进行运算。

第 8 步：确定满足 $p_i > \alpha$ 的全部卷积层，按步骤 1 至 7 顺序逐步对其进行处理。

第 9 步：微调改进 MobileNet-SSD 网络模型。

获取状态估计值 x' 和理想模型状态 x_k^m 作差得到所要的残差，对利用改进 MobileNet-SSD 网络模型处理分析得到所要的状态信息。

$$\gamma(k) = x' - x_k^m. \quad (12)$$

已经知道噪声都是白噪声，所以测量值是无偏的。状态估计值只是测量值的滤波处理值，偏差的均值也应该是 0，所以残差只会在 0 附近上下小幅波动。由于当出现突变故障时，残差幅值会迅速增大，所以设计一个故障评价函数：

$$J_s = \frac{1}{N} \sum_{j=k-N}^k \gamma'(j) \gamma'(j). \quad (13)$$

式中： N 为时间窗口长度；当 $J_s \geq v_1$ 时，被认为有故障。 v_1 为人为选定的阈值。一般根据经验确定，有时可以根据一些先验信息和数据分布特点确定，就此实现故障检测。

笔者在 SSD 网络框架下对计算机设备运行故障检测的嵌入式深度学习模型进行构建，利用轻量级卷积神经网络 MobileNet 代替原有的 VDD-16 网络，将 Batch Normalization 层与卷积层作融合处理，实现 MobileNet-SSD 网络的改进。将提取的计算机设备温度、运行速度以及电流运行数据作为故障训练样本，将其输入到改进后的 MobileNet-SSD 网络模型中，在 PC 机上对模型进行训练，对训练后的改进 MobileNet-SSD 网络模型作剪枝处理，以减少模型参数，达到压缩计算机设备运行故障检测模型的目的。将压缩后的模型植入到嵌入式设备中，通过对测试样本进行检测，获得计算机设备运行故障检测结果。具体计算机设备运行故障检测流程如图 3 所示。

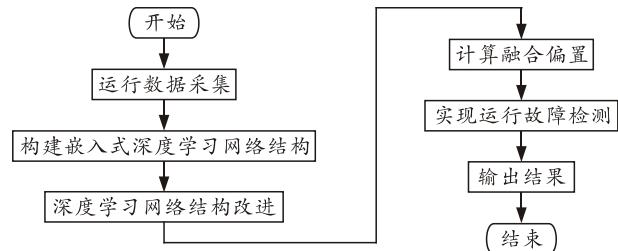


图 3 设备运行故障检测流程

2 实验分析

将某大型公司的计算机运行网络作为研究对象，网络中计算机以并行计算方式完成业务的处理与分析，计算机设备总数为 60 台，主要由服务器和处理器构成。由于计算机设备的长期运行，已处于性能衰退状态。布置 5 个无线传感器以 2 ms 的数据采集间隔、3.76 Hz 频率对计算机设备运行数据进行采集，构建样本数据集，样本总数为 8 000，以 4:1 比例进行训练、测试样本划分。采用笔者所提方法对计算机设备运行故障进行检测，验证笔者方法的检测效果，监测系统界面如图 4 所示。

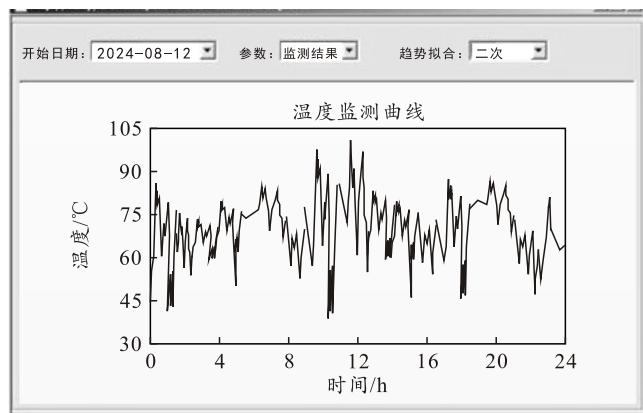
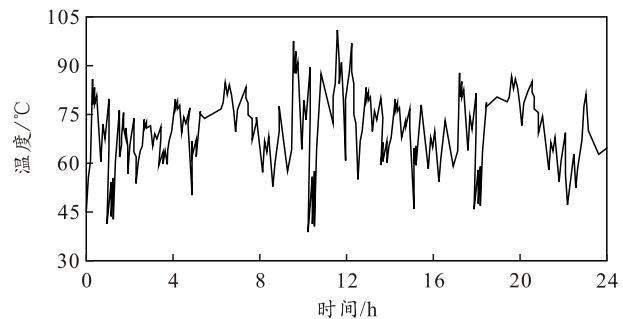


图 4 监测系统界面

确定表征计算机设备运行故障的特征量是实现各个计算机设备故障高精度识别的关键。采用笔者方法对样本数据集进行预处理，通过计算机设备的故障特征量提取结果分析笔者方法的数据处理能力，实验结果如图 5 所示。



(a) 温度特征量提取结果

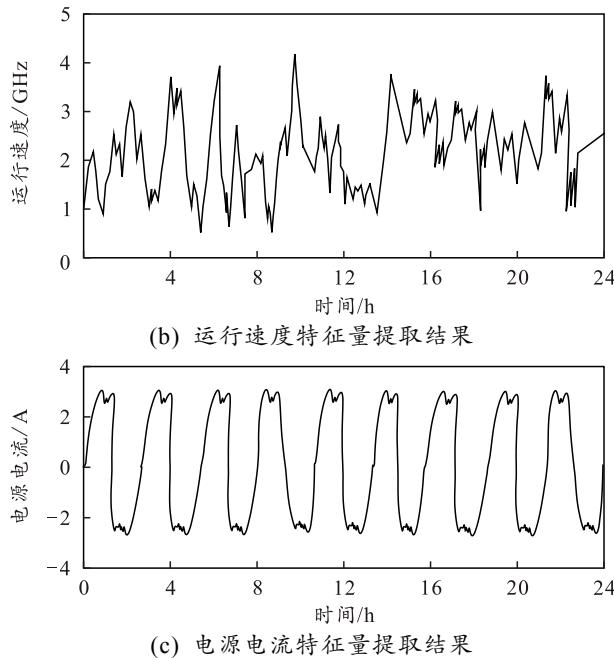


图 5 计算机设备的故障特征量提取结果分析

分析图 5 可知:采用本文中方法对样本数据集进行处理,能够实现不同时间下计算机设备温度、运行速度、电源电流特征量的提取,各特征量曲线呈现不断变化趋势,温度特征量基本在 45~80 °C 范围内变化,也有部分时间点设备温度达到 90 °C 以上。运行速度曲线整体在 1~4 GHz 范围内变化,少量时间点计算机设备运行速度低于 1 GHz; 电源电流曲线走势不平滑,波峰处有明显波动。实验结果表明,本文中方法可实现计算机设备运行故障特征量的提取。

在获取计算机设备运行故障特征量的基础上,将其作为本文中计算机设备故障检测模型的输入,对模型进行训练,并与改进前的 SSD 网络模型进行对比,通过对本文中检测模型和改进前后训练损失曲线进行对比分析,验证本文中方法的检测效果,实验结果如图 6 所示。

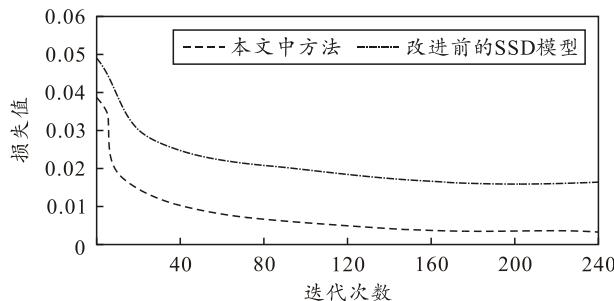


图 6 损失曲线变化对比分析

分析图 6 可知:分别利用本文中方法与改进前的 SSD 网络模型对训练样本进行学习,随着迭代次

数的不断增加,损失曲线均呈下降趋势变化,本文中方法在迭代初期损失值下降幅度较大,当迭代 20 次后,损失值降低幅度减缓,当迭代 80 次后其值逐渐收敛,且低于 0.01; 改进前的 SSD 模型在迭代 160 次后方可慢慢收敛,损失值达到 0.02 左右,其值远高于本文中方法。实验结果表明:对 SSD 网络模型进行改进处理,可有效降低训练损失,加快模型的收敛效率,对计算机设备运行故障检测精度的提升具有推动作用。

在对计算机设备运行故障进行检测时,难点在于部署于嵌入式设备上的改进 MobileNet-SSD 网络模型参数量仍然较大、对嵌入式设备资源占用较高,为了提高故障检测效率以及检测效果,笔者对该模型进行剪枝处理,设定剪枝率阈值 α 分别为 0.2、0.3、0.5, 利用平均精度均值(mAP)指标反映故障检测效果,将剪枝处理前的模型参数量以及 mAP 值作为对比,通过分析不同 α 下模型参数量的变化以及 mAP 的变化验证本文中方法的有效性,实验结果如表 1 所示。

表 1 剪枝效果对比分析

模型	模型参数量	mAP 指标值
剪枝前	15.26	29.4
本文中方法	$\alpha=0.2$	7.00
本文中方法	$\alpha=0.3$	7.00
本文中方法	$\alpha=0.5$	7.00
		28.7
		28.5
		28.2

分析表 1 可知:采用本文中方法对部署于嵌入式设备上的改进 MobileNet-SSD 网络模型进行剪枝处理,可实现模型的有效压缩,模型参数量大幅降低,当 α 取值逐渐增大后,计算机设备运行故障检测的 mAP 指标仅有小幅度的降低,说明虽然有更多的卷积层受到剪枝处理,但是通过微调优化使得原始模型的信息得以完整保留。实验结果表明:本文中方法在高压缩率条件下,也获得较好的计算机设备运行故障检测效果。

采用本文中方法对测试样本集中的计算机设备运行数据进行检测,通过对部分计算机设备故障检测结果进行分析验证本文中方法的检测效果,实验结果如表 2 所示。

表 2 计算机设备运行故障检测结果分析

设备名称	故障时间	故障检测结果
A1	9:12:20	供电故障
A2	8:36:52	测温装置失灵
A3	9:50:15	散热不良, 设备温度升高
A4	10:20:36	针脚接触不好, 会对计算机启动产生影响
A5	13:14:48	Windows 服务不启动
A6	14:23:54	找不到域服务器

分析表 2 可知：应用本文中方法能够根据计算机设备运行数据样本实现其故障的识别与检测，确定故障设备、故障发生时间，并查找出故障发生原因；因此，本文中方法可实现计算机设备运行故障检测。

为了验证本文中方法下计算机设备运行故障诊断效果，采用 LSTM 方法作为对比方法，并进行计算机设备运行故障诊断准确率对比，得到结果如表 3 所示。

表 3 计算机设备运行故障诊断准确率 %

设备数量	LSTM 方法	本文中方法
50	68.9	90.5
100	82.0	98.3
150	78.3	96.9
200	80.5	99.8

分析表 3 可知：当设备数量为 50 台时，LSTM 方法下故障诊断准确率为 68.9%，本文中方法下故障诊断准确率可达 90.5%；当设备数量为 200 台时，LSTM 方法下故障诊断准确率为 80.5%，本文中方法下故障诊断准确率可达 99.8%；上述结果表明，本文中方法能够有效提升计算机设备运行故障诊断准确率。

3 结论

将本文中方法应用于某大型企业的计算机运行网络中，通过分析计算机设备故障特征量提取结果、故障检测模型改进前后损失曲线的变化以及模型剪枝前后的 mAP 等指标的变化，验证了本文中方法的故障检测能力。实验结果表明：1) 可完成计算机设备故障特征量的提取；2) 可有效降低训练损失，加快模型的收敛效率；3) 可降低模型参数量，在高剪枝率下，也能取得较好的故障检测效果。

参考文献：

- [1] 许剑, 蔡贵龙. 计算机常见故障检测与维护 [J]. 信息与电脑(理论版), 2019, 31(20): 11-12, 20.
- [2] 田心宇, 姚英. 一种新型 BIT 技术在机载计算机设计中的应用研究 [J]. 微电子学与计算机, 2019, 36(4): 29-32.
- [3] 唐颜, 徐天蒙, 周伟, 等. 基于 PNN 某型弹载计算机故障诊断 [J]. 弹箭与制导学报, 2021, 41(3): 84-87.
- [4] 周剑飞, 刘晨. 基于 LSTM 的设备故障在线检测方法 [J]. 计算机工程与应用, 2020, 56(1): 272-278.
- [5] 文成林, 吕菲亚. 基于深度学习的故障诊断方法综述 [J]. 电子与信息学报, 2020, 42(1): 234-248.
- [6] 汪晓臣, 段鹏宇, 李樊, 等. 物联网设备的深度学习故障预测方法 [J]. 小型微型计算机系统, 2020, 41(12): 2631-2635.
- [7] 陈志强, 陈旭东, 李川, 等. 深度学习在设备故障预测与健康管理中的应用 [J]. 仪器仪表学报, 2019, 40(9): 206-226.
- [8] 刘鑫月, 陈瑞, 王坤峰, 等. 基于平行图像与深度学习的绝缘子目标检测法 [J]. 计算机仿真, 2021, 38(1): 61-66, 202.
- [9] 包萍, 刘运节. 不均衡数据集下基于生成对抗网络的改进深度模型故障识别研究 [J]. 电子测量与仪器学报, 2019, 33(3): 176-183.
- [10] 宋浏阳, 李石, 王芃鑫, 等. 基于动态统计滤波与深度学习的智能故障诊断方法 [J]. 仪器仪表学报, 2019, 40(7): 39-46.
- [11] 吴天成, 王晓荃, 蔡艺军, 等. 基于特征融合的轻量级 SSD 目标检测方法 [J]. 液晶与显示, 2021, 36(10): 1437-1444.
- [12] 张骏, 朱标, 吉涛. 基于 MobileNet-SSD 的红外人脸检测算法 [J]. 激光与红外, 2021, 51(1): 107-113.
- [13] 施思雨, 魏继增. 基于 MobileNet-SSD 目标检测算法的硬件加速器设计 [J]. 微电子学与计算机, 2022, 39(6): 99-107.
- [14] 孙长迪, 潘志松, 张艳艳. 基于低开销可变形卷积的 MobileNet 再轻量化方法 [J]. 计算机科学, 2022, 49(12): 312-318.
- [15] 汝承印, 张仕海, 张子森, 等. 基于轻量级 MobileNet-SSD 和 MobileNetV2-DeeplabV3+ 的绝缘子故障识别方法 [J]. 高电压技术, 2022, 48(9): 3670-3679.