

doi: 10.7690/bgzdh.2025.11.018

基于多传感器融合和改进 EfficientNetV2-B0 的电机故障诊断

朱思成¹, 孙宁^{1,2}, 王松雷³, 张佳宝¹, 陈津奥², 夏浩然²

(1. 南京信息工程大学自动化学院, 南京 210044; 2. 无锡学院江苏省工业环境危害要素监测与评估工程研究中心, 江苏 无锡 214105; 3. 江苏省特种设备安全监督检验研究院无锡分院, 江苏 无锡 214073)

摘要: 针对电机在发生故障时故障信号易被强噪声淹没、信号采集不全面且训练网络冗杂的问题, 将融合多传感器信号和改进 EfficientNetV2-B0 的迁移学习模型引入到电机故障诊断中。传感器融合方法通过格拉姆角场 (Gramian angular field, GAF) 将 1 维时间序列转换成图像, 保证特征信息的完整性, 没有时间依赖性, 并利用 Retinex 增强和拉普拉斯金字塔分解的图像融合算法实现多源传感器信号的图像融合。针对 EfficientNetV2-B0 网络提出了添加深度可分离卷积 (depthwise separable convolution, DSConv) 和高效多尺度注意力 (efficient multi-scale attention, EMA) 的改进, 并结合迁移学习 (transfer learning, TL) 技术建立电机故障诊断模型。对电机的各种工况进行分类和测试的结果表明: 该方法能有效地对设备故障进行分类, 对电机各种工况的识别平均准确率达 100%。

关键词: 传感器融合; 故障诊断; EfficientNetV2-B0; 迁移学习

中图分类号: TP212 **文献标志码:** A

Motor Fault Diagnosis Based on Multi-sensor Fusion and Improved EfficientNetV2-B0

Zhu Sicheng¹, Sun Ning^{1,2}, Wang Songlei³, Zhang Jiabao¹, Chen Jin'ao², Xia Haoran²

(1. School of Automation, Nanjing University of Information Science & Technology, Nanjing 210044, China;

2. Jiangsu Provincial Engineering Research Center for Monitoring and Evaluation of Industrial Environmental Hazard Elements, Wuxi University, Wuxi 214105, China;

3. Wuxi Branch of Jiangsu Province Special Equipment Safety Supervision Inspection Institute, Wuxi 214073, China)

Abstract: In order to solve the problems that the fault signal is easily submerged by strong noise, the signal acquisition is not comprehensive and the training network is complex when the motor fault occurs, the transfer learning model based on multi-sensor signal fusion and the improved EfficientNetV2-B0 is introduced into the motor fault diagnosis. The sensor fusion method converts the 1D time series into images through the gramian angular field (GAF) to ensure the integrity of the feature information without time dependence, and uses the image fusion algorithm of Retinex enhancement and Laplacian pyramid decomposition to realize the image fusion of multi-source sensor signals. Adding depthwise separable convolution (DSConv) and efficient multi-scale attention (EMA) are proposed for the EfficientNetV2-B0 network, and combined with transfer learning (TL) technology to establish a motor fault diagnosis model. Various working conditions of the motor are classified and tested, and the results show that the method can effectively classify the equipment faults, and the average accuracy of the identification of various working conditions of the motor reaches 100%.

Keywords: sensor fusion; fault diagnosis; EfficientNetV2-B0; transfer learning

0 引言

在电机发生故障时, 会产生各种各样的特征信号。采集和分析振动信号中包含的特征信息是一种有效的故障诊断方法。随着深度学习的不断发展, 基于振动信号的智能诊断技术不断涌现。单个传感器获得的数据有限, 难以反映设备在运行中的全部特性; 因此, 融合多传感器数据以获取有用的特征信息已经成为智能诊断发展的共识。

图像处理技术是融合传感器数据的新型方法,

可以细分为像素级融合、决策级融合和特征级融合。其中像素级融合在空间域技术下的计算成本较低, 仅依赖于根据数学公式或算法在多个图像数据的不同对应像素值之间进行选择, 避免信息丢失。

迁移学习技术是处理电机故障诊断样本数据的重要方法, 通过将相同或相关源领域的知识转移到目标领域来提高模型性能。Mao 等^[1]利用 VGG-16 模型, 在 Imagenet 大规模图像数据集上进行预训练, 以证明迁移学习技术在轴承故障诊断中的有效性。

笔者提出一种新的基于多传感器融合和改进

收稿日期: 2024-10-09; 修回日期: 2024-11-17

基金项目: 国家自然科学基金(42275156); 江苏高校“青蓝工程”, 教育部产学研合作协同育人项目(202102224006); 江苏省大学生创新创业项目(202213982017Z)

第一作者: 朱思成(2000—), 男, 江苏人, 硕士。

EfficientNetV2-B0 的电机故障诊断方法, 首先利用格拉姆角场和基于 Retinex 增强和拉普拉斯金字塔分解的图像融合算法实现多源传感器信息的融合。通过设备故障检测与分类用例验证了该方法的有效性。使用改进的 EfficientNetV2-B0 迁移学习模型提取特征信息并实现图像分类, 达到对故障信息进行分类的目的并实现电机的故障诊断。

1 基于图像融合的传感器融合方法研究

1.1 基于格拉姆角场的信号预处理

格拉姆角场 (GAF) 保持了时间依赖性, 可以将 1 维时间序列转换为 2 维图像。用极坐标表示时间序列, 然后计算和/差三角函数。Gramian 角差场 (GADF) 分别表示 2 个角的和差。

考虑一个时间序列 $X = \{x_1, x_2, \dots, x_n\}$ 。规范化可以将每个元素映射到 $[-1, 1]$ 和 $[0, 1]$ 。

$$\tilde{x}_i = (x_i - x_{\min}) / x_{\max}; \quad (1)$$

$$\tilde{x}_i = ((x_i - x_{\min}) + (x_i - x_{\max})) / (x_{\max} - x_{\min})。 \quad (2)$$

式中 x_{\min} 和 x_{\max} 分别为数据序列的最小值和最大值。时间序列可以用角度 φ_i 和 r_i 半径在极坐标中表示, r_i 由时间戳 t_i 表示。当时间序列映射到 $[-1, 1]$ 上时, φ_i 映射到 $[0, \pi]$ 上, $[0, 1]$ 可划分为 n 个等价的子区间。除了初始点 0, 还可以得到 n 个分割点, 分割点与 $\tilde{X} = \{\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n\}$ 相关联。

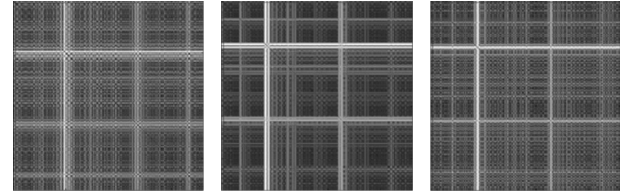
$$\left. \begin{aligned} \varphi_i &= \arccos(\tilde{x}_i), -1 \leq \tilde{x}_i \leq 1 \\ r_i &= t_i/n \end{aligned} \right\}; \quad (3)$$

$$\text{GADF} = \begin{bmatrix} \sin(\varphi_1 - \varphi_1) & \dots & \sin(\varphi_1 - \varphi_n) \\ \sin(\varphi_2 - \varphi_1) & \dots & \sin(\varphi_2 - \varphi_n) \\ \vdots & & \vdots \\ \sin(\varphi_n - \varphi_1) & \dots & \sin(\varphi_n - \varphi_n) \end{bmatrix} = \sqrt{1 - \tilde{X}^2}^T \times X - \tilde{X}^T \times \sqrt{1 - \tilde{X}^2}。 \quad (4)$$

由于函数的单调性, 每个序列都有一个唯一的极坐标映射。GAF 矩阵通过使用和差公式进行构造。由式(4)可知, 随着时间的推移, 矩阵元素由左上向右下移动, 时间维度被编码到矩阵的几何中, 矩阵对角元素为单角度值, 因此可以实现时间序列的近似重构。然而, 矩阵大小的扩大会增加计算复杂度。为了解决这一问题, 引入分段聚合近似, 既能保持序列趋势, 又能大大减小序列大小。

为了实现原始时间序列的重构, 避免特征信息的丢失, 笔者采用 GAF 对原始信号进行预处理, 生成 GADF 图像。图 1 展示了电机驱动端以 12 kHz

的采样频率和 1 797 的转速下产生 0.021 英寸的外圈故障, 由驱动端轴承故障所生成的电机基座 (BA)、驱动端 (DE) 和风扇端 (FE) 的振动数据 3 类的 GADF 图像。

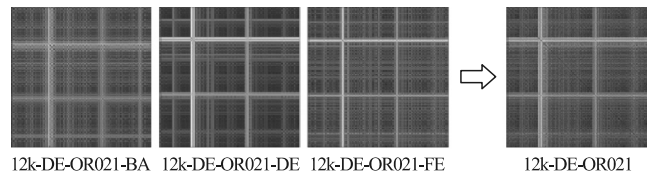


12k-DE-OR021-BA 12k-DE-OR021-DE 12k-DE-OR021-FE
(a) 电机基座(BA) (b) 驱动端(DE) (c) 风扇端(FE)

图 1 振动数据 3 类 GADF 图像

1.2 基于 SSR-拉普拉斯的图像融合算法优化

为了更加全面地获得被测对象的准确描述, 笔者使用传感器图像融合算法, 通过整合多个传感器的数据, 提高检测结果的精确度。笔者提出一种改进的 SSR-拉普拉斯图像融合算法: 首先, 采用单尺度 Retinex(SSR)算法^[2]对可见光图像进行增强; 然后, 利用拉普拉斯金字塔算法^[3]将增强后的可见光图像分解成多层图像。最后, 采用基于区域特征测量的方法对分解后的图像进行融合。实验结果表明: 即使在低对比度、低照度的源图像中, 该算法也能获得对比度增强程度较高的融合图像, 能较好地地区分目标和背景, 并保留源图像细节, 如图 2 所示。



12k-DE-OR021-BA 12k-DE-OR021-DE 12k-DE-OR021-FE 12k-DE-OR021
图 2 基于 Retinex 增强和拉普拉斯金字塔分解的图像融合

2 EfficientNetV2-B0 故障诊断算法的优化

2.1 EfficientNetV2-B0 基础模型

EfficientNetV2-B0 的网络结构如表 1 所示。该网络共分为 8 个 Stage: Stage0 为一个卷积核尺寸为 3×3 , 步长为 2 的普通卷积层, 包含 BN 和 Swish 激活函数; Stage1~Stage3 是 Fused-MBConv 结构的重复; Stage4~Stage6 是 MBConv 结构的重复; Stage7 由一个 1×1 的普通卷积层、平均池化层和全连接层组成。Channels 为特征通道数; SE0.25 表示 Squeeze-and-Excitation(SE)模块^[4]当中第一个全连接层的节点个数是输入该 MBConv 模块特征通道数的 1/4; Layers 代表每个模块重复执行多少次。

表 1 EfficientNetV2-B0 网络结构表

Stage	Operator	Stride	Channels	Layers
0	Conv3×3	2	32	1
1	Fused-MBConv1, k3×3	1	16	1
2	Fused-MBConv4, k3×3	2	32	2
3	Fused-MBConv4, k3×3	2	38	2
4	MBConv4, k3×3, SE0.25	2	96	3
5	MBConv6, k3×3, SE0.25	1	112	5
6	MBConv6, k3×3, SE0.25	2	192	8
7	Conv1×1 & Pooling & FC	—	960	1

2.2 改进的 EfficientNetV2-B0

2.2.1 添加深度可分离卷积 (DSConv)

深度可分离卷积 (DSConv)^[5] 因其在降低神经网络中卷积层的计算成本方面的有效性而受到广泛关注，核心思想是将一个完整的卷积运算分解为逐深度卷积 (depthwise convolution) 和逐点卷积 (pointwise convolution) 2 步进行。如图 3 所示，DSConv 先经过逐通道卷积将每个卷积核应用到每个输入通道生成 3 个特征图，然后通过逐点卷积，用 n 个大小为 $1 \times 1 \times 3$ 的卷积核生成所需数量的特征图。与传统卷积相比，深度可分离卷积使用更少的参数获得相似的特征提取结果。采用深度可分离卷积之后，在特征提取方面与标准卷积提取效果差不多，但在参数量和计算量上却有大幅度下降。笔者在 Fused-MBConv 网络结构中采用 DSConv 深度可分离卷积，代替传统卷积进行特征提取。

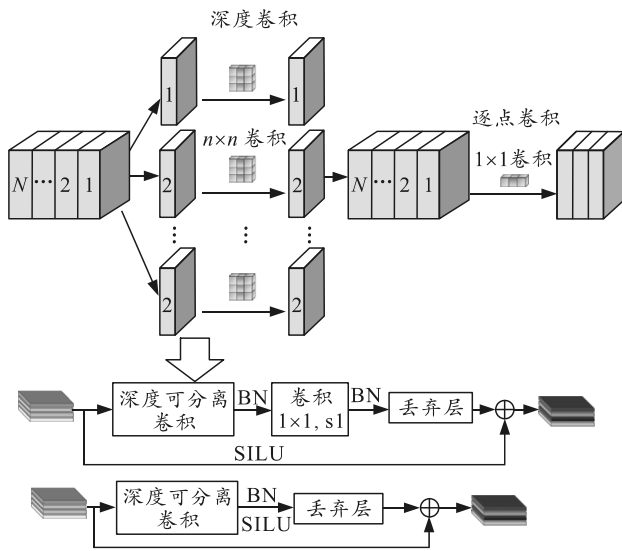


图 3 深度可分离卷积 (DSConv) 结构和改进的 Fused-MBConv 结构

2.2.2 引入 EMA 注意力机制

在不受噪声干扰的情况下提取有效的故障特征一直是重点关注的问题。注意机制可以通过注意图增加潜在特征的接受度。EfficientNetV2-B0 网络中

使用的注意力机制为 SE 模块，其中该模块的 2 个全连接层旨在捕获非线性跨通道交互，这样会产生降维导致的通道与权重不对应，丢失率一部分特征细节，并且通道之间的依存关系是效率不高的。

为了对原始 EfficientNetV2 中的注意力机制进行改进，笔者在 EfficientNetV2-B0 中使用一种新的高效的多尺度注意力 (EMA) 模块。以保留每个通道上的信息和降低计算开销为目标，将部分通道重塑为批量维度，并将通道维度分组为多个子特征，使空间语义特征在每个特征组中均匀分布。其中并行子网块有助于有效地捕获跨维度交互并建立维度间依赖关系。EMA 的整体结构如图 4 所示。注意因素仅由每组中全局和局部特征描述符之间的相似性来指导。考虑到跨空间信息聚合方法，在建模远程依赖关系的同时，将精确的位置信息嵌入到 EMA 中。融合不同尺度的上下文信息使 CNN 能够对高级特征图产生更好的像素级关注。通过使用跨空间学习方法，卷积核的并行化是一种更强大的结构来处理短期和长期依赖。与有限接受域形成的渐进式行为相反，在中间特征映射中并行使用 3×3 和 1×1 卷积可以获得更多的上下文信息。

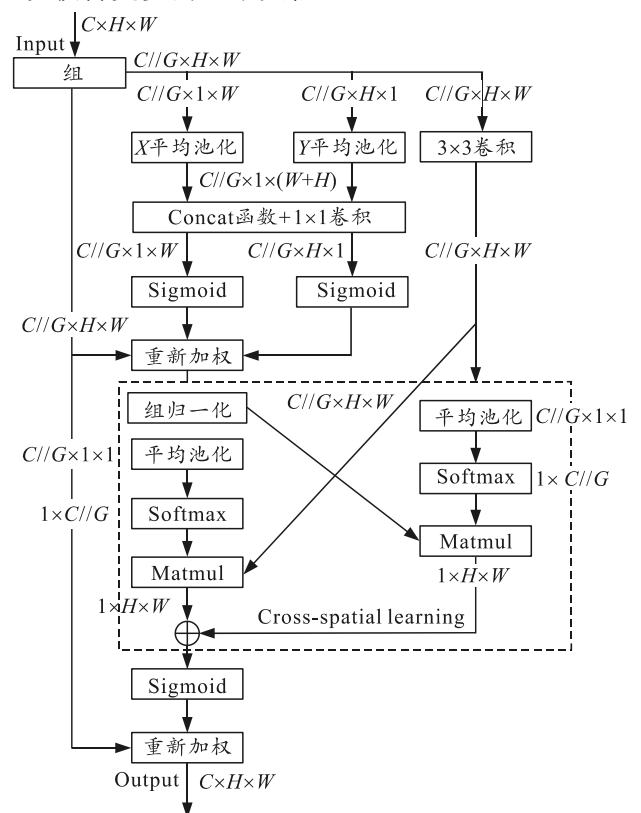


图 4 EMA 注意力机制

笔者还将 MBConv 中 SE 模块替换为 EMA 模块。EMA 模块使特征图不会因为维度缩减的操作而

丢失信息，同时网络也能对图片特征进行更充分地提取。改进后的 MBConv 模块结构如图 5 所示。

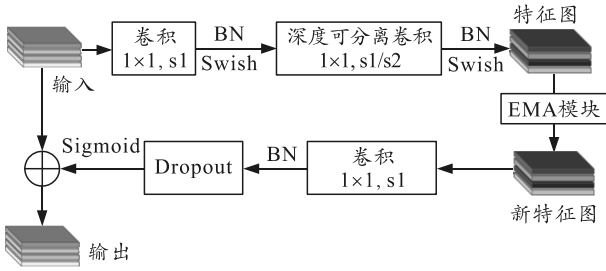


图 5 改进的 MBConv 结构

2.3 迁移学习技术的结合

迁移学习是一种机器学习 (ML) 方法, 通过将相同或相关源领域的知识转移到目标领域来提高模型性能。与从头开始训练模型不同, 预先训练的模型可以用于改进特定的目标任务。

对于正式定义: 让定义一个 ML 分类问题的域, 该问题由 2 部分组成 $D=\{X, P(X)\}$, 其中 X 表示特征空间, $P(X)$ 表示边缘概率分布。特征向量 $x=\{x_1, \dots, x_n\} \in X$ 是特征空间的特定元素, y 是属于标签空间 Y 的对应类标号。对于域 D , 任务可以定义为 $T=\{Y, f(\cdot)\}$, 其中 $f(\cdot)$ 是从 $\{x_i, y_i\}$ 对学习到的预测函数。源域数据集可以定义为 $D_S\{(x_{s1}, y_{s1}), \dots, (x_{sn}, y_{sn})\}$, 其中 $x_{s1} \in X_S$ 的 D_S 和 $y_{s1} \in Y_S$ 为对应的类标号。类似地, 目标域可以定义为 $D_T\{(x_{t1}, y_{t1}), \dots, (x_{tn}, y_{tn})\}$ 。源域和目标域的任务分别为 T_S 和 T_T 。如果源任务和目标任务预测函数分别为 $f_S(\cdot)$ 和 $f_T(\cdot)$, 则可以将迁移学习正式定义为利用 D_S 和 T_S 在 $D_S \neq D_T$ 或 $T_S \neq T_T$ 时获得的知识来改进 $f_T(\cdot)$ 。

笔者使用流行的 ImageNet^[5]数据集获得 D_S 和 T_S , 并使用改进的 EfficientNetV2-B0 模型进行图像分类训练。该模型在 ImageNet 数据集的图像上进行了预训练, 并在开始使用融合传感器图像进行训练之前进行了冻结。训练过程中, 只有最后的全连接层与分类输出的最终全连接层进行了修改, 以与我们实验中的故障类别数量相匹配。经过优化的源模型可以识别超过 100 万张图像的特征, 旨在为目标域的故障诊断任务做出贡献。

3 实验与结果分析

深度学习模型的所有权重都是从之前利用著名的 ImageNet 数据集进行优化和预训练的模型中进行迁移学习所转移过来的。模型训练将融合的时频图像整合到先前训练的网络之上。所有的模型都使用相同的超参数进行训练, 使用相同的数据预处理。

采用指数衰减策略调整学习率大小, Batch Size 为 64, Epoch 的最大数目限制为 100, 设初始学习率为 0.01, 每经过 100 轮迭代后学习率缩减为 0.000 1, 选择损失函数收敛稳定的模型作为最终的分类模型。探索的预训练模型为改进的 EfficientNetV2-B0 和其他对比模型。

3.1 实验设置和数据收集

为了验证基于迁移学习模型的传感器融合方法在故障诊断中的有效性, 笔者采用美国凯斯西储大学 (CWRU) 的轴承数据中心的数据进行分析。如图 6 所示, 该轴承数据采集平台具体组成包含: 1) 功率 1.5 kw 的电动机; 2) 扭矩传感器 (连接电动机与功率测试机); 3) 功率测试机; 4) 电动机的电子控制器 (在图中未显示)。实验轴承为 DE 轴承, 该轴承有 4 种健康状态, 分别为健康 (NS)、内圈故障 (IR)、外圈故障 (OR) 和滚动体故障 (B), 每种故障位置包含 3 种故障尺寸, 分别为 0.007、0.014 和 0.021 英寸。待测轴承为 BA 轴承、DE 轴承和 FE 轴承。实验中, 将振动传感器放置在驱动电机 3 个轴承座之上, 以 12 kHz 的采样频率和 1 797 的电机转速采集由驱动端 (DE) 故障轴承影响的基座、驱动端和风扇端 3 类振动加速度信号。每截取一次故障模式, 获取长度为 864 的 1 590 个样本, 共计 $1\ 590 \times 10 = 15\ 900$ 个样本。此外, 所有数据集的 70% 用作训练数据集, 20% 用作模型选择和交叉验证的验证数据集, 10% 用作最终测试的测试数据集。在每次训练和测试中, 随机划分数据集, 以保证对模型性能的综合评价。

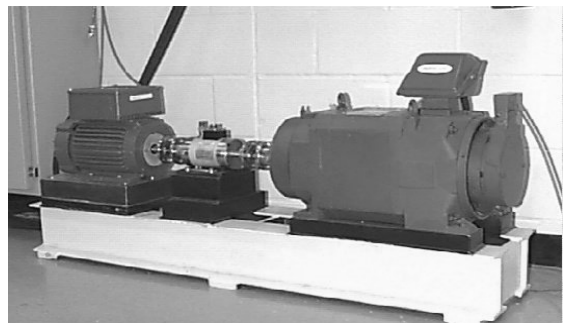


图 6 凯斯西储大学轴承故障试验台

3.2 数据处理

GAF 对原始数据进行预处理, 选择不重叠的滑动窗口, 得到各类信号的 GADF 图像, 图像大小为 224×224 。完成全部 GADF 图像的生成后, 通过基于 Retinex 增强和拉普拉斯金字塔分解的图像融合算法将各类信号的 GADF 图像进行融合。不同故障

模式下的融合图像如图 7 所示，这些图像将作为提出的故障诊断模型的输入。

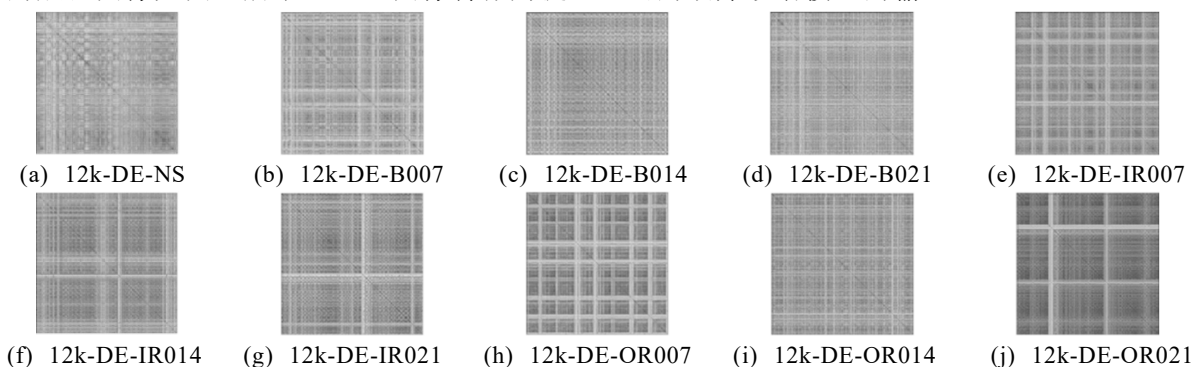


图 7 不同故障模式下的融合图像

3.3 模型评估指标

在结合机器学习技术时，精度评估在检查模型性能中是不可分割的一部分。本实验选用准确率 (Accuracy)，模型参数量 (Params num) 以及计算复杂度 (Flops) 3 个参数对模型进行评价。其中 accuracy 是所有的预测正确 (正类负类) 的占总的比重；Params num 是指网络模型中需要训练的参数总数，在这里用于度量模型的大小；Flops 是指每秒浮点运算次数，是考量一个网络模型的计算量的标准。

笔者使用验证精度和损失进一步评估训练模型。此外，当涉及到验证准确性和损失时，使用连续监控来发现训练和验证性能中任何明显的偏差。

3.4 训练结果

如图 8 所示，从训练精度、训练损失和验证精度、验证损失方面展示所提出模型的性能。提出的模型进行了 100 次迭代训练。该模型的训练准确率大幅提高。在 100 次迭代过程中，模型的训练准确率最高达到了 99.68%。经过大约 20 次迭代后，训练数据集的准确率和损失值已经基本稳定，模型已经开始收敛。模型的验证准确率在第 3 轮迭代中就超过了 98%，在 100 次迭代过程中，模型的准确率最高达到了 100%。并且在第 3 轮迭代中，验证数据集的准确率和损失值就已经基本确定，这说明 EfficientNetV2-B0 模型的收敛能力很强。

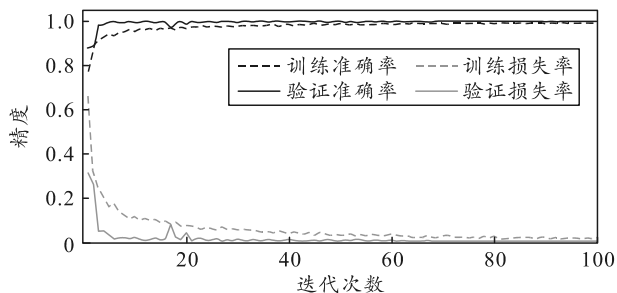


图 8 改进的 EfficientNetV2-B0 训练

为了展示传感器融合方法中的图像融合技术在电机故障诊断中的优势，笔者使用基座、驱动端和风扇端 3 类振动加速度信号融合图像和单个驱动端振动加速度信号图像做对比。使用 EfficientNetV2-B0 模型架构在 100 次迭代中获得的准确率百分比。如图 9 所示，基座、驱动端和风扇端 3 类振动加速度信号融合图像在第 3 轮迭代中的准确率就超过了 98%，并且在第 3 轮迭代后就开始收敛，并且准确率最高达到了 100%。综上证明笔者所提出的方法不仅收敛性好，而且预测效果更稳定，诊断效率更高。

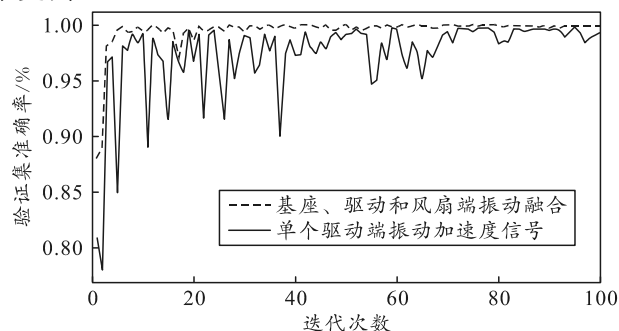


图 9 3 类振动加速度信号融合和单个驱动端振动加速度信号图像训练

为了证明笔者所改进的模型的优越性和可行性，以及对电机故障识别的鲁棒性，笔者还将该模型与一些常见的分类网络模型进行对比实验，将训练好的模型用于诊断测试集数据，以评估诊断性能。为了减少实验的偶然性，每种方法测试 10 次。实验结果如表 2 所示。可以明显看出：本文中方法在电机故障传感器融合图像测试集上的分类检测结果优于其他方法。在分类精度方面，本文中方法与 ResNet-50 相比精度提高了约 4.8%。结果表明，该方法对电机故障信号进行分类识别是有效的。同时，该方法模型参数量从 EfficientNetV2-B0 的 7.1 降低到 6.1 M，大大降低了模型的复杂性。实验结果

表明, 笔者所提出的改进 EfficientNetV2-B0 有利于移动端的模型部署。

表 2 不同模型运行时的对比

Method	Accuracy/%	Params num/M	Flops/G
Improved EfficientNetV2-B0	100.00	6.1	0.821
EfficientNetV2-B0	99.87	7.1	0.737
MobileNetV3	98.17	8.2	0.326
ShuffleNet	97.33	5.4	0.524
VGG-19	97.04	20.5	19.643
ResNet-50	95.23	47.3	3.798
AlexNet	96.76	60.9	0.727

为了展示所改进的各模块的有效性, 笔者对改进的 EfficientNetV2-B0 模型做消融实验, 针对改进的各个模块设计了 4 种不同的模型, 如表 2 所示。并分别生成模型的混淆矩阵, 如图 10 所示, 在混淆矩阵中, 颜色越深值越高, 其中标签 0 为 12k-DE-NS, 标签 1 为 12k-DE-B007, 标签 2 为

12k-DE-B014, 标签 3 为 12k-DE-B021, 标签 4 为 12k-DE-IR007, 标签 5 为 12k-DE-IR014, 标签 6 为 12k-DE-IR021, 标签 7 为 12k-DE-OR007, 标签 8 为 12k-DE-OR014, 标签 9 为 12k-DE-OR021。

表 3 4 种不同模型进行消融实验

Expt	Efficient-NetV2-B0	DSCConv	EMA	Accuracy/%	Params num
1	✓			99.87	7, 139, 704
2	✓	✓		99.90	6, 989, 224
3	✓		✓	100.00	6, 214, 820
4	✓	✓	✓	100.00	6, 064, 340

由表 3 可知, 笔者针对 EfficientNetV2-B0 所改进的每个模块都是有效的, 引入深度可分离卷积 (DSCConv) 替换普通卷积显著降低了模型参数量且略微提高了模型预测的准确率, 引入 EMA 注意力机制在提高准确率的同时减少了模型参数量。

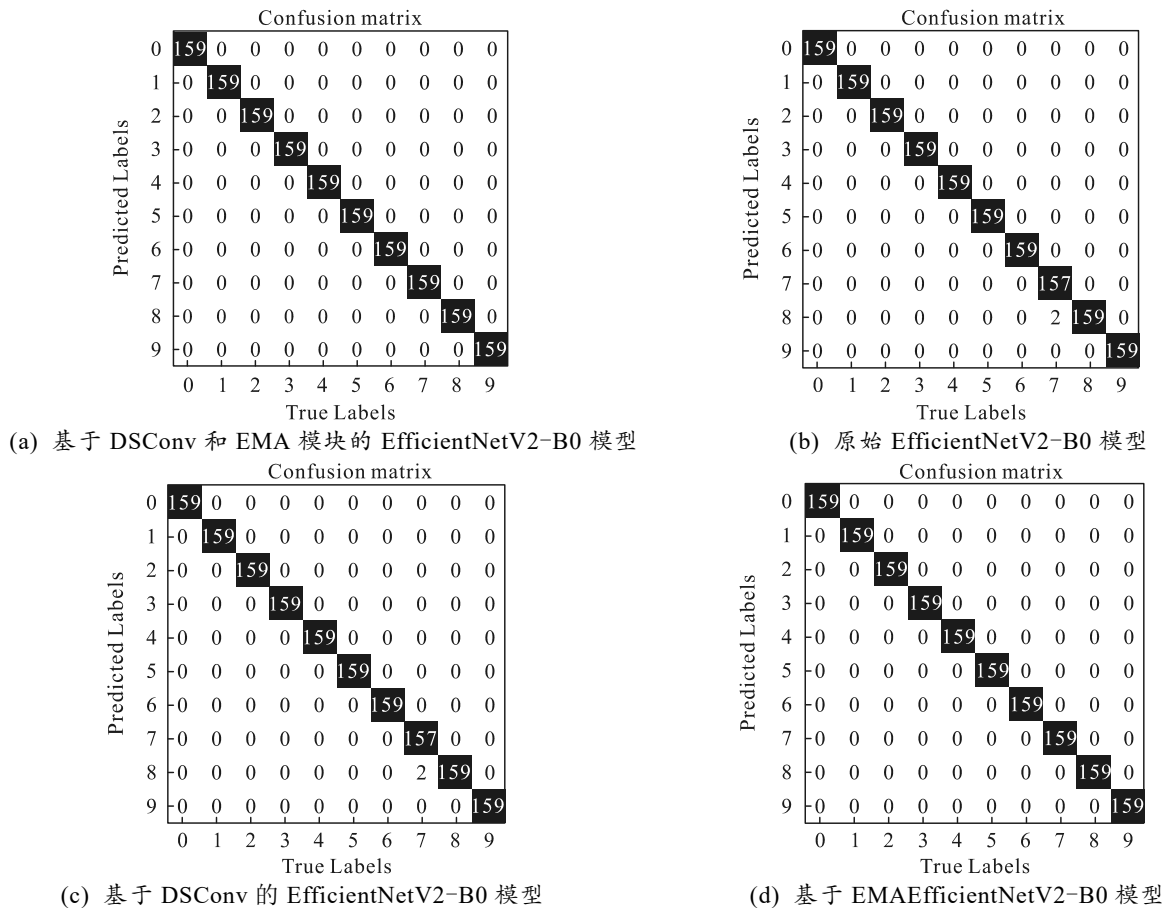


图 10 4 种不同模型的混淆矩阵

从图 10 的混淆矩阵可以直观地看出: 原先的模型在驱动端外圈故障下故障尺寸为 0.007、0.014 英寸的分类并不精确, 而改进的 EfficientNetV2-B0 可以实现对各类电机故障信号的精确分类, 且模型在每个类别的准确率都达到了最大值。

综合以上实验结果分析, 笔者提出的基于多传

感器融合和改进 EfficientNetV2-B0 的电机故障诊断方法在故障信号分类识别方法在测试集中效果很好, 识别准确率高达 100%, 且该方法的复杂性低于大多数其他模型。