

# 基于 CNN 特征提取及模型融合的飞机液压系统故障诊断

李时奇,赵东标,申珂楠,丰赢政  
(南京航空航天大学 机电学院,江苏 南京 210016)

**摘要:**为了对飞机液压系统进行有效故障诊断,采用 CNN 对飞机液压系统的压力信号进行特征提取。用提取到的特征输入线性模型、决策树、支持向量机、k 邻近等算法对其进行故障诊断,并使用 Stacking 模型融合技术将多个模型融合。结果表明,相比于直接用 CNN 训练进行故障诊断,使用 CNN 提取出的特征进行训练能极大减少训练时间同时提高准确率。

**关键词:**CNN 特征提取;Stacking 多模型融合;飞机液压系统;故障诊断

**中图分类号:**TP277; V233.91 **文献标志码:**B **文章编号:**1671-5276(2020)05-0193-04

## Fault Diagnosis of Aircraft Hydraulic System Based on CNN Feature Extraction and Model Fusion

LI Shiqi, ZHAO Dongbiao, SHEN Kenan, FENG Yingzheng

(College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

**Abstract:** To diagnose the fault of the aircraft hydraulic system effectively, CNN is used to extract the feature of the pressure signal of the aircraft hydraulic system. The extracted features are input into the linear model, decision tree, support vector machine, k-nearest neighbor algorithm for the fault diagnosis. Stacking model fusion technique is used to fuse models. The results show that the features extracted from CNN can be used to reduce the training time as well as improve the accuracy, compared with using CNN training directly for the fault diagnosis.

**Keywords:** CNN feature extracting; Stacking model fusion; aircraft hydraulic system; fault diagnosis

## 0 引言

智能故障诊断是未来飞机液压系统故障诊断发展的趋势之一<sup>[1-2]</sup>。国内通常采用小波包分解或者信息熵等方法对复杂的液压系统压力信号进行特征提取,然后输入 BP 神经网络或者支持向量机等学习器进行故障诊断分类<sup>[3-5]</sup>。这些特征提取的方法对人的经验要求比较高;不具备通用性;对于多传感器压力信号难以有效融合;单一模型有较大过拟合风险。

针对上述问题,提出了一种卷积神经网络(CNN)特征提取加多模型融合的故障诊断方法。使用一维多通道 CNN 可以直接接收多传感器一维压力时间序列,从而实现多传感器融合。选取少量数据对 CNN 进行有监督训练提取特征。再用提取出的特征训练如线性多分类、决策树、支持向量分类、k 邻近等学习器分别对其进行故障诊断。最后使用 Stacking 技术对这些模型进行融合。使用 Amesim 软件对典型飞机液压系统建模并进行故障仿真,验证了算法效果。结果表明,本文算法在准确率和训练用时上更优。

## 1 算法理论介绍

### 1.1 一维多通道 CNN

由于飞机液压系统的压力信号是一维时间序列曲线,而且不同位置的多个传感器可以采集多条曲线,为同时将多个传感器的一维时间序列输入 CNN,采用一维多通道 CNN。网络中的所有卷积层、池化层均使用一维结构。

设  $H_l$  为第  $l$  层的输入特征图,  $S$  为卷积操作的结果,  $H_{l+1}$  为卷积层的输出特征图,  $K$  为卷积核,  $i$  代表特征图上的索引,  $p$  代表卷积核上索引,  $c$  为通道的索引,  $H_l(:, c)$  为通道  $c$  的特征图,  $b$  为偏移量,  $f(x)$  为激活函数,则一维多通道卷积公式如下<sup>[6]</sup>:

$$S(i) = \sum_c \sum_p H_l(i+p, c) K(p) \quad (1)$$

$$H_{l+1} = f(S + b) \quad (2)$$

池化操作分为最大池化和均值池化。设第  $l$  层为池化层,  $k$  为池化窗口的大小,  $p$  代表池化窗口上的索引。最大池化和均值池化操作公式分别为式(3)和式(4):

$$H_{l+1}(i) = \frac{\max_{p \in K} H_l(i+p)}{p \in K} \quad (3)$$

$$H_{l+1}(i) = \frac{1}{k} \sum_p H_l(i+p) \quad (4)$$

## 1.2 线性多分类器

对数几率回归(logistic regression, LR)是一种用线性回归模型的预测结果去逼近真实标记的对数几率,虽然名字是“回归”,实际上是一种分类算法。设  $\mathbf{x}$  为输入向量,  $y$  为输出标记,  $w$  和  $b$  为待学习的权重和偏置。公式如下:

$$y = \frac{1}{1 + e^{-(w^T x + b)}} \quad (5)$$

当  $y > 0.5$  时预测为正类,  $y < 0.5$  时预测为反类,  $y = 0.5$  时可以随机预测。通过 OvO(One vs. One)或者 OvR(One vs. Rest)等拆分策略可以推广到多分类<sup>[7]</sup>。

## 1.3 决策树分类器

决策树通过树型结构对样本进行学习。树的节点代表某个属性,该节点下的路径代表属性的不同取值。每个叶节点代表某个类。随机森林(random forest, RF)<sup>[8]</sup>是多个决策树通过 bagging 策略集成的,RF 能引入随机性,有效减少过拟合。

## 1.4 支持向量分类器

支持向量机(support vector machine, SVM)通过寻找一个超平面划分两类样本,离超平面最近的几个训练样本称为“支持向量”,两个不同类别的支持向量到超平面距离和称为“间隔”,支持向量机算法试图找到间隔最大的超平面来划分不同类<sup>[9]</sup>。SVM 同样可以使用拆分策略推广到多分类。

## 1.5 $k$ 邻域分类器

$k$  邻近算法( $k$ -nearest neighbor, KNN)的训练过程就是保存训练集所有数据。当新的样本输入时,计算训练集中距离它最近的  $k$  个样本,根据这  $k$  个样本的类别来决定新样本的类别<sup>[10]</sup>。

# 2 算法流程

图 1 为算法流程示意图。首先对典型飞机液压系统进行 Amesim 建模,通过改变某些部件参数来模拟飞机液压系统的常见故障。采集不同故障状态以及正常状态的信号,进行归一化处理并通过划窗重采样进行数据集增强。将训练集分为 A 和 B 两个子集。用 A 集数据对 CNN 进行有监督训练。再将数据集 B 输入训练好的 CNN,提取中间层的特征,将中间提取出的特征输入其他学习器进行训练。之后将多个学习器的结果用 Stacking 融合技术融合。最后将测试集输入训练好的模型,验证算法效果。

## 2.1 CNN 特征提取

CNN 各个层的结构如表 1 所示。CNN 可以接受一维多通道数据,可将不同飞机液压系统中不同传感器获取到的一维压力时间序列输入 CNN。表中 filter 指卷积核(又

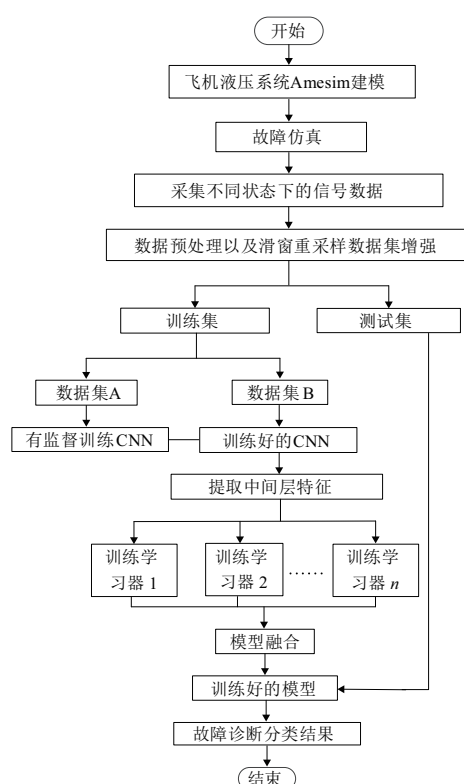


图 1 算法流程示意图

叫滤波器)或者池化滤波器。filter 大小  $1 \times 100$  指窗口大小为 100 的卷积核在特征图上滑动,步长指的是 filter 在特征图上滑动步长。

表 1 CNN 各层参数

| 网络层   | 特征尺寸           | 特征通道数 | filter 大小      | filter 步长    |
|-------|----------------|-------|----------------|--------------|
| 输入层   | $1 \times 600$ | 3     | —              | —            |
| 卷积层 1 | $1 \times 120$ | 64    | $1 \times 100$ | $1 \times 5$ |
| 池化层 1 | $1 \times 60$  | 64    | $1 \times 2$   | $1 \times 2$ |
| 卷积层 2 | $1 \times 30$  | 64    | $1 \times 20$  | $1 \times 2$ |
| 池化层 2 | $1 \times 15$  | 64    | $1 \times 2$   | $1 \times 2$ |
| 全连接 1 | $1 \times 256$ | 1     | —              | —            |
| 全连接 2 | $1 \times 128$ | 1     | —              | —            |
| 输出层   | $1 \times 6$   | 1     | —              | —            |

用 A 集监督训练 CNN,目的是让 CNN 的每一层获得不错的权重从而使网络具备提取抽象特征能力。将 B 集输入到训练好的网络,提取全连接层 2 尺寸为 128 的特征。此时向量的总数据维度从 1 800 降到了 128,此时的特征已经过滤掉一些无效信息并提取到有用的抽象特征,能极大地加快训练速度,简化训练难度,提高训练准确率。

## 2.2 Stacking 模型融合技术

Stacking 模型融合技术<sup>[11]</sup>是一种两级的模型融合技术,将初级学习器学习到的分类结果作为次级学习器的输入,再进行一次学习。

Stacking 模型融合技术示意图如图 2、图 3 所示。首



### 3.2 CNN 特征提取效果分析

若用在 A 集上训练的 CNN 直接预测 B 集的故障诊断结果,准确率仅为 84.3%,这是因为 A 集只占总数据集的很少一部分,但是足够提取出有效的特征用于下一步训练。

图 6 为 A 集大小对 CNN 训练时间以及最终准确率的影响。图 7 为从 CNN 不同中间层提取的特征对后续训练时间以及最终准确率的影响。由图可知,为使准确率高的同时训练时间尽可能少, A 集大小取 20%,提取 CNN 中间层的卷连接层 2 作为特征进行下一步训练。

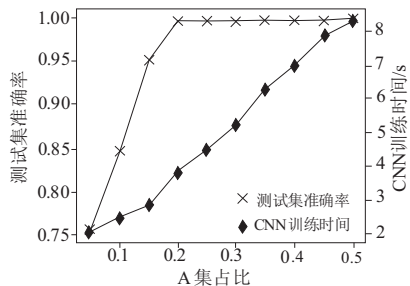


图 6 A 集占比对结果的影响

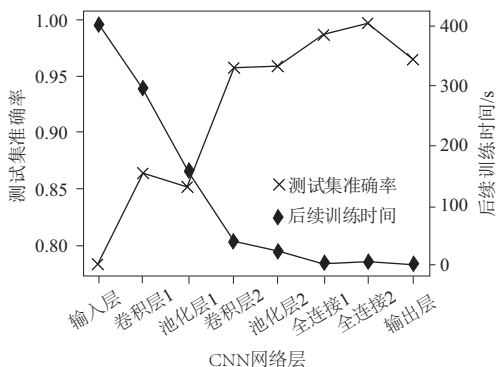


图 7 CNN 不同层特征训练结果

### 3.3 故障诊断结果分析

表 4 为各学习器用 CNN 提取到的特征进行训练(表中 i)与直接用原始输入压力曲线训练(表中 ii)的对比。由表可知,经过 CNN 提取到的特征能极大加快训练速度,提高故障诊断的准确率,降低模型大小。

表 4 用特征与用原始数据训练对比

| 模型  | 准确率/% |      | 训练耗时/s |        | 模型大小/KB |        |
|-----|-------|------|--------|--------|---------|--------|
|     | i     | ii   | i      | ii     | i       | ii     |
| KNN | 98.8  | 90.3 | 0.82   | 9.40   | 3 743   | 52 205 |
| LR  | 99.2  | 68.2 | 1.03   | 189.00 | 0.37    | 86.00  |
| SVM | 99.3  | 64.2 | 0.24   | 178.00 | 0.32    | 86.00  |
| RF  | 98.6  | 92.7 | 0.31   | 4.97   | 0.70    | 0.70   |

表 5 为用不同次级学习器进行 Stacking 融合的比较

以及和直接用简单投票法融合的准确率比较。可以看到,使用 LR 作为次级学习器准确率最高。

表 5 不同模型融合方法比较

| 融合方法       | 融合后准确率 |
|------------|--------|
| KNN 作次级学习器 | 0.994  |
| LR 作次级学习器  | 0.997  |
| SVM 作次级学习器 | 0.832  |
| RF 作次级学习器  | 0.992  |
| 简单投票法      | 0.990  |

图 8 为 Stacking 模型融合后与各个模型的对比。由图可知,融合后的模型比所有原模型的准确率都要高。

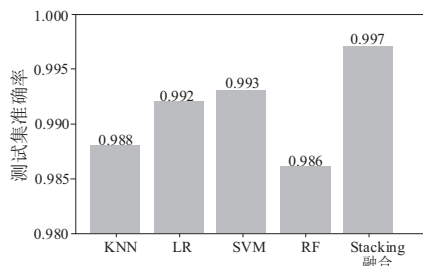


图 8 Stacking 模型融合与各个模型对比图

表 6 为 CNN 特征提取加 Stacking 模型融合与直接用 CNN 训练全部数据集进行故障诊断的对比。由表可知,在准确率方面多模型融合略高,在训练耗时上多模型融合更少,但是所占的空间更大。

表 6 两种方法效果对比

| 对比项目    | CNN 特征提取+多模型融合 | CNN 直接故障诊断 |
|---------|----------------|------------|
| 测试集准确率  | 0.997          | 0.982      |
| 训练时间/s  | 7.8            | 16.0       |
| 模型大小/MB | 8.3            | 4.6        |

## 4 结语

本文采用 CNN 特征提取加 Stacking 多模型融合技术对飞机液压系统进行了故障诊断。使用 Amesim 软件对典型飞机液压系统建模并进行故障仿真来验证算法效果,结果如下:

- 1) 使用少量数据监督训练 CNN 也能提取出不错的中间特征,用提取出的特征训练比用原始输入信号训练在各方面都有较大提升。
- 2) Stacking 多模型融合准确率高于各个单一模型。
- 3) 相比于直接用 CNN 训练全部数据集进行故障诊断, CNN 特征提取加多模型融合训练速度明显更快,准确率也有少量提升。

(下转第 199 页)

的产生是由于导波本身的多模态传播特性造成的,因地脚螺栓腐蚀检测最关心的是腐蚀缺陷的截面积缺失率,一定程度的定位误差对于检测来讲是可接受的。

#### 4) 灵敏度确定

对于深度分别为 4 mm、5 mm、7 mm 的不同腐蚀缺失率的模拟缺陷进行检测,将系列检测数据通过线性拟合公式( $y=ax+b$ )进行拟合后得到增益值随距离远近的变化曲线,如图 7 所示。这些曲线可以近似描述缺陷位置、回波增益和缺陷当量尺寸之间的关系,可以依此建立地脚螺栓的腐蚀损伤评价体系。

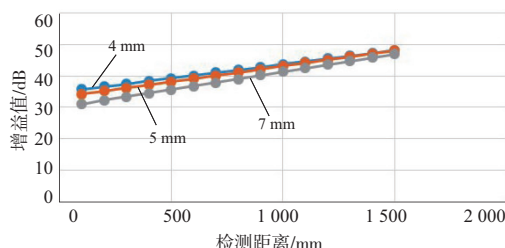


图 7 M42 mm 螺栓增益值随着距离变化曲线图

由图 7 可以看出每个不同腐蚀缺失率模拟缺陷的距离增益曲线均可近似看成是线性的。在此基础上,可通过在地脚螺栓模拟试样上所需检测最远处设置最小允许缺陷,调整仪器增益使该缺陷回波高度处于 80%,并以此增益作为检测灵敏度。在实际检测中,以该方法设置检测灵敏度对未知缺陷进行检测时,可检测出整根地脚螺栓上大于最小允许缺陷当量的缺陷,并根据上述距离增益曲线得到相应腐蚀缺陷当量,进而估算腐蚀缺失率。

## 4 结语

本文从理论上分析了柱面超声导波的频散特性,制作了不同深度缺陷的地脚螺栓模拟试样,采用相控阵超声柱面导波技术对模拟试样进行检测,分析了混凝土包覆和探头频率对检测结果的影响,并建立地脚螺栓的腐蚀损伤评价体系,得出以下结论:

1) 可以选用衰减极小值频率对应的导波模态来检测地脚螺栓。

2) 混凝土包裹对地脚螺栓检测效果没有明显影响,表明该检测方法可适用于埋地地脚螺栓不开挖条件下的现场检测。

3) 实际检测中,可根据距离-增益变化曲线确定缺陷当量,用于评估地脚螺栓腐蚀损伤量。

相控阵超声柱面导波技术检测螺栓具有操作简单、结果直观、可重复性好、现场检测人员要求低等优点,该技术还可应用于其他行业类似螺栓的检测。

#### 参考文献:

- [1] 张成涛,刘宗林,唐小辉,等. 输电铁塔的腐蚀与防护[J]. 全面腐蚀控制,1998,12(3):11-15.
- [2] 王秀玉,朱德补,程学启. 对运行输电线路铁塔防腐问题的探讨[J]. 山东电力技术,2006(6):55-57.
- [3] 默增禄,程志云. 输电线路杆塔的腐蚀与防治对策[J]. 电力建设,2004,25(1):22-23,36.
- [4] 周正干,冯海伟. 超声导波检测技术的研究进展[J]. 无损检测,2006(2):57-63.
- [5] NA W B, KUNDU T. Inspection of interfaces between corroded steel bars and concrete using the combination of a piezoelectric zirconate-titanate transducer and an electromagnetic acoustic transducer[J]. Experimental Mechanics, 2003, 43(1):24-31.
- [6] BEARD M D, LOWE M J S, CAWLEY P. Development of a guided wave inspection technique for rock bolts[J]. Aip Conference Proceedings, 2002,615(1):1318.
- [7] VIENS M. Leaky torsional acoustic modes in infinite clad rods[J]. Journal of the Acoustical Society of America, 1994, 95(2):701-707.
- [8] 何存富,吴斌. 超声柱面导波技术及其应用研究进展[J]. 力学进展,2001,31(2):203-214.
- [9] 吴斌,张瑞芳,刘秀成,等. 基于纵向模式超声导波陷频特性的钢绞线拉力测量新方法[J]. 机械工程学报,2016(12):9-15.
- [10] 滕德金,王东方,滕德巧. 基于超声相控阵检测技术的连杆检测方法研究[J]. 机械制造与自动化,2018,47(6):198-202.

收稿日期:2020-07-15

(上接第 195 页)

#### 参考文献:

- [1] 周汝胜,焦宗夏,王少萍. 液压系统故障诊断技术的研究现状与发展趋势[J]. 机械工程学报,2006(9):6-14.
- [2] 郑永梅,叶宏飞,汪小飞,等. 飞机液压系统故障诊断的现状与发展趋势[C]. 航空试验测试技术学术交流会论文集,北京:中国航空学会,2016:4.
- [3] 窦丹丹,姜洪开,何毅娜. 基于信息熵和 SVM 多分类的飞机液压系统故障诊断[J]. 西北工业大学学报,2012,30(4):529-534.
- [4] 刘泽华,李振水. 基于小波包特征熵的飞机液压系统故障诊断方法研究[J]. 计算机测量与制,2009,17(9):1705-1707,1710.
- [5] 李耀华,王星州. 飞机液压系统故障诊断[J]. 计算机工程与应用,2019,55(5):232-236,264.
- [6] LECUN Y, BOTTOU L, BENGIO Y, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural

Computation, 1989, 11(4):541-551.

- [7] 孙志远,鲁成祥,史忠植,等. 深度学习研究与进展[J]. 计算机科学,2016,43(2):1-8.
- [8] 周志华. 机器学习[M]. 北京:清华大学出版社,2016:33-34.
- [9] 方匡南,吴见彬,朱建平,等. 随机森林方法研究综述[J]. 统计与信息论坛,2011,26(3):32-38.
- [10] BURGESS J C. A tutorial on support vector machines for pattern recognition[J]. Data Mining and Knowledge discovery, 1998, 2(2):121-167.
- [11] 徐慧丽. Stacking 算法的研究及改进[D]. 广州:华南理工大学,2018.
- [12] 王海涛. 飞机液压元件与系统[M]. 北京:国防工业出版社,2012:190-196.

收稿日期:2019-07-02