

# 机器故障诊断技术文献综述

姓名：尹常攀

学号：M201870482

**摘要：****目的** 为了能更有效、准确地对复杂设备进行状态监测和故障诊断。**方法** 综述近年故障诊断技术中重要方法的基本原理、特点、局限性和研究现状。在大量文献的基础上，基于计算机技术、信号处理技术、人工智能技术和互联网技术讨论现代故障诊断技术的发展趋势。**结果** 故障诊断技术主要研究机器或机组运行状态的变化在诊断信息中的反映，分为基于系统模型、信号分析和人工智能等 3 类。**结论** 随着基础学科和前沿学科的不断发展和交叉渗透，故障诊断技术也在不断创新，未来的发展趋势主要集中于将不同人工智能技术以某种方式结合、集成或融合的混合集成诊断技术。

**关键词：**故障诊断技术；系统模型；信号分析；人工智能；混合集成诊断技术

## 一、前言

故障诊断技术起源于 20 世纪 60 年代的欧美国家,英国人主要是用它来研究摩擦、汽车、发动机领域;日本主要是在民用行业;法国和瑞士都开发了关于机械诊断的相关设备。我国对故障诊断技术的研究是在 20 世纪 80 年代初开始的,首先在石化企业进行了与故障诊断相关的工作,而后在钢铁、发电等行业能也开始广泛得到了运用,并呈现逐步发展的趋向。虽然起步比较晚,但经过 30 多年的发展,现在已经在航天、汽车等各个领域内取得了不错的成绩。

故障检测技术是在熟知了机械装备设备运行过程的各个状态之后,从而明确机械装备设备是否处于正常运行的状态,以便与及时的查找到并排除故障,查出原因,并熟知故障趋势的技术。通过故障诊断能够及时地、正确地对各种异常状态或故障状态做出诊断,预防或消除故障,对设备的运行进行必要的指导,提高设备运行的可靠性、安全性和有效性,以期把故障损失降低到最低水平;保证机械设备发挥最大的设计能力,制定合理的检查维修制度,以便在允许的条件下,充分挖掘设备潜力,延长服役期限和使用寿命,降低设备全寿命周期费用;通过检测监视、故障分析、性能评估等手段,为设备结构修改、优化设计、合理制造及生产过程提供数据和信息。

机器故障诊断是机械工程的领域之一,着重于找出机器中的故障的位置与类型,其中针对旋转机械的部分发展的相当完善,而旋转机械也是很常接触到的机械种类。目前根据故障诊断所采取处理方式可以将机器故障诊断的研究现状大致分为三个阶段:

- 1、通过直接对系统构建模型来分析故障类型或者发生的部位的系统模型阶段;
- 2、通过采集系统各个部位的信号对信号进行分析的信号分析阶段,其中的信号采集的方式包括振动监控、热影像、油粒子分析等。而信号的分析有谱密度分析、小波分析、小波转换、短时距傅立叶变换、Gabor 展开,维格纳准概率分布(WVD)、倒频谱、bispectrum、相关法、高分辨率频谱分析、波形分析(时域下的,因为频谱分析只关注频率分析,不考虑相位资讯)等;
- 3、采用如今的人工智能方法对从系统中获取到的数据进行分析的人工智能阶段。该阶段主要是将如今的智能算法应用在故障诊断,使特征的选取更多的依赖于数据与模型而非人为的主观选取。

## 二、机器故障诊断研究现状

### 1. 基于系统模型的方法

该方法可以深入系统本质的动态性质，与控制系统紧密结合，从而实现实时诊断、容错控制、系统修复与重构等故障，但系统模型难获得，易出现鲁棒性问题。

#### 1.1 状态估计诊断法

状态估计诊断法是通过被控过程的状态反映系统运行状态，再结合适当模型诊断故障的方法。当系统可观和部分可观时，重构被控过程的状态，将估计值与测量值进行比较，构成残差序列，以检测和分离系统故障。综合利用了系统的结构、功能、行为信息，算法简单，直接有效，但一些非线性系统难建立模型，使其受到了一定的限制。

刘春生等<sup>[1]</sup>提出了一种可以解决状态不可测时故障的状态估计器设计方法，可以用于报警和故障容错控制。随着故障对象复杂性的增加，对故障状态估计精确性需求会越来越高。

#### 1.2 参数估计法

参数估计法的基本思想是由模型参数序列计算过程参数序列并确定过程参数的变化量序列，基于此变化序列进行故障诊断。这种方法不需要计算残差序列，比状态估计诊断法更有利于故障的分离。目前较多的研究是将参数估计法与其他方法相结合。

BELLALIB 等<sup>[2]</sup>将参数估计法与神经模糊推理系统方法结合，提出了一种新的有效的自适应神经模糊推理系统方法，用于卫星姿态控制系统（ACS）故障诊断，这对解决复杂分布模型的参数估计问题有着重大意义。

### 2. 基于信号分析的方法

基于信号处理是一种传统的故障诊断方法，通过信号模型分析可测信号，提取方差、幅值、频率等特征值，从中获取与故障相关的征兆，达到对机械故障进行监测和诊断的目的。诊断过程如下图 1 所示。

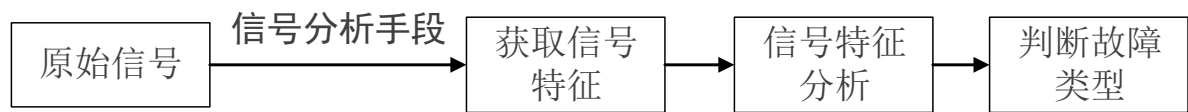


图 1 基于信号分析的故障诊断结构

## 2.1 EMD 时频分析方法

机械设备发生故障时,会出现非平稳特性的振动信号,经验模态分解(EDM)是一种优秀的非平稳信号处理方法,可以应用于机械设备动态分析与故障诊断。该方法无需预先设定基函数,就可将信号自适应地分解成不同尺度的本征模式函数(IMF),具有多分辨率和自适应性的特点,能有效地提取原信号的特征信息。在实际应用中,EMD 存在模态混叠、频率分辨率低的问题。

1、Norden E. Huang 等人<sup>[3]</sup>于 1998 年首先提出了 EMD 时频分析方法,原始信号经 EMD 分解获得一系列 IMF 分量,提取包络谱中不同特征频率的振幅比作为特征向量,将特征向量输入到 SVM 中,自动判别故障模式。HHT 处理非平稳信号的基本过程是:首先利用 EMD 方法将给定的信号分解为若干固有模态函数(以 Intrinsic Mode Function 或 IMF 表示,也称作本征模态函数),这些 IMF 是满足一定条件的分量;然后,对每一个 IMF 进行 Hilbert 变换,得到相应的 Hilbert 谱,即将每个 IMF 表示在联合的时频域中;最后,汇总所有 IMF 的 Hilbert 谱就会得到原始信号的 Hilbert 谱。

2、V.K. Ra 等<sup>[4]</sup>对 HHT 变换产生的 IMF 分量进行 FFT 变换,通过比较说明了该方法能够提高 HHT 算法的效率。

3、曹冲锋等<sup>[5]</sup>针对现有降噪方法的不足,提出一种集成的 EMD 分解降噪方法,能够有效消除高斯噪声及降低脉冲干扰,抑制模态混叠现象。

4、戴桂平和刘彬<sup>[6]</sup>研究了提取信号瞬时参数的 EMDPHS 法,用小波去噪对信号进行预处理,再进行 EMD 分解,能够有效地去除噪声,提高 EMD 的分解质量。

5、Yu Yang 等人<sup>[7]</sup>提出一种基于 IMF 包络谱的故障特征提取方法。

## 2.2 小波变换法

小波变换(wavelet transform, WT)是一种新的变换分析方法,它继承和发展了短时傅立叶变换局部化的思想,同时又克服了窗口大小不随频率变化等缺点,能够提供一个随频率改变的“时间-频率”窗口,是进行信号时频分析和处理的理想工具。它的主要特点是通过变换能够充分突出问题某些方面的特征,能对时间(空间)频率的局部化分析,通过伸缩平移运算对信号(函数)逐步进行多尺度细化,最终达到高频处时间细分,低频处频率细分,能自动适应时频信号分析的要求,从而可聚焦到信号的任意细节,解决了 Fourier 变换的困难问题,成为继 Fourier 变换以来在科学方法上的重大突

破。

1、Nanjing<sup>[8]</sup>采用了小波变换滤波分析轴承的低频信号，可以把轴承的低频信号准确地提取出来，但由于低频段信号能量较低，还受其他中低频干扰源影响，精确诊断故障也不容易，并且小波变换不能对信号的高频进一步分解，不能准确有效地提取共振的高频带信息。

2、NikolaouNG 等<sup>[9]</sup>提出了使用小波包变换(WPT)分析系统振动信号来诊断轴承的局部缺陷。

3、V. Purushotham 等人<sup>[10]</sup>将离散小波变换(DWT)用于处理滚动轴承多故障情况，提出了用隐马尔科夫模型(HMMs)来实现故障模式的识别。

4、J.Pineyro 等<sup>[11]</sup>采用 Haar 小波提取滚动轴承早期故障信号。

5、MoriK.等<sup>[12]</sup>对轴承的振动信号采用离散小波变换分解,基于小波变换系数的变化能成功在线诊断轴承的早期疲劳剥落故障。

6、Abderrazek Djebala 等人<sup>[13]</sup>基于优化的小波多分辨率分析技术，提出一种信号降噪方法。利用峭度值作为优化和评估准则，选择小波变换中的相应参数。

7、Khalid F. Al-Raheem 等人<sup>[14]</sup>提出一种基于拉普拉斯小波包络的功率谱的方法。利用峭度最大值准则优化拉普拉斯小波的形态参数。模拟和实验的结果都表明该方法在谐振频带内提取缺陷特征参数的有效性。

8、A. Tabrizi 等人<sup>[15]</sup>将小波包分解(WPD)、总体经验模态分解(EEMD)和支持向量机(SVM)相结合，用于轴承的早期故障诊断。WPD 主要用于降噪处理，EEMD 提取信号的特征向量，通过 SVM 算法实现的超平面构造的样本用于检测轴承的状态。文中提出的三种方法的结合可以有效的实现轴承状态的分类。

## 2.3 共振解调方法

轴承局部损伤故障会引起冲击尖脉冲信号，由于这些冲击尖脉冲具有很宽的频带，经传感器接收后产生共振，输出 1 频率与传感器固有频率接近的减幅振荡信号，振荡信号跟原来的尖脉冲信号相比，从幅值上被放大，从时间上被展宽，因而故障信息突出。再利用高中心频率的带通滤波器有效地使冲击故障信息与低频干扰信号隔开，提高信噪比。最后对滤波器输出的信号进行解调处理，如求绝对值、取包络线等，得到低频的共振解调信号。

1、Rubini R.和 Meneghetti U.<sup>[16]</sup>利用包络解调的方法提取时域信号中所包含的轴承

内环、外环及滚动体的早期微弱故障信息,求取故障特征频率,并研究了载荷对测量结果的影响。

2、J.Antonino-Daviu<sup>[17]</sup>将希尔伯特变换用于感应电动机的偏心相关组件的瞬态检测,将小波的近似系数和的分量相互转换,可靠地实现了动态偏心故障的诊断。

3、Min-Chun Pan 和 Wen-Chang Tsao<sup>[18]</sup>利用 EMD 的带通滤波性质,选择一个或多个适当的 IMF 分量,用于后续的包络分析,该方法能够有效的诊断出轴承滚动体多故障的情况。

4、Wei Sun 等人<sup>[19]</sup>将离散小波变换与包络分析相结合用于提取轴承振动数据的特征谱,谱的互相关系数被用于识别轴承的不同工作状态。通过比较,该方法要优于传统的特征损伤频率提取技术。

5、Junsheng Cheng 等人<sup>[20]</sup>将广义解调时频分析与包络阶次谱技术相结合,用来诊断齿轮的故障。将齿轮箱中产生的多分量信号用广义解调时频分析方法分解成单分量信号,对各单分量信号进行包络分析,角域重采样各包络信号,再对重采样信号进行包络分析获得相应的包络阶次谱,从包络阶次谱中可以得到齿轮的工作情况。

6、Myeongsu Kang 等人<sup>[21]</sup>提出将时变多分辨率包络谱分析和特征判别分析用于轴承故障诊断。将利用声发射技术采集到的信号在时域上进行分区,获得轴承失效的原始信息。对选择后的信号用多级有限脉冲响应滤波器组进行处理以获取异常症状信息最丰富的子带信号,对于带信号进行谱分析。

### 3. 基于数据驱动的人工智能方法

该方法主要是将如今的智能算法应用在故障诊断,对原始信号进行预处理后从中选取若干个特征,然后以特征值作为输入,故障类型作为输出,进行有监督式的学习,学习后所得的模型用于故障诊断。并随着智能的提升使特征的选取越来越依赖于数据与模型而减少人为的主观选取。过程如图 2 所示。

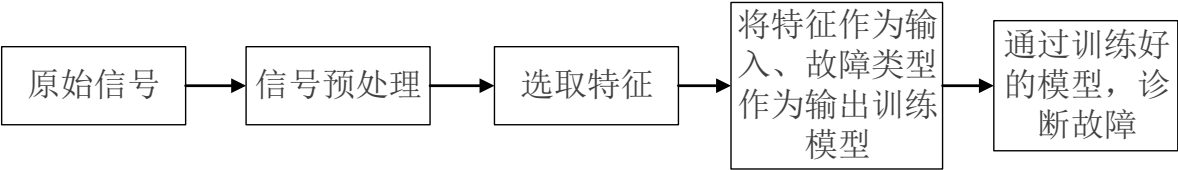


图 2 基于数据驱动的人工智能故障诊断结构

### 3.1 机器学习法

在故障检测和诊断中，有些数学分类模型其实是属于监督式学习法，会利用有标签的资料集组成的训练集来分析，目的是要准确的识别冗余、故障以及异常的样本。在过去数十年来，基于机器学习方法在研究领域中已发展了许多的分类法预处理模型<sup>[13]</sup>。

1、Zexian Wei 等人<sup>[22]</sup>提出一种基于 AP 聚类算法和自适应特征选择技术的智能方法用于轴承故障诊断。利用总体经验模态分解（EEMD）和小波包变换（WPT）精确的提取故障特征信息。为了解决聚类算法效率问题并避免维数问题，提出自适应特征选择技术。

2、Myeongsu Kang 等人<sup>[23]</sup>提出一种基于二进制蝙蝠算法故障特征分析的早期低速轴承可靠故障诊断方法。首先将小波包变换应用于声发射获得的信号以计算基于小波的故障特征，然后利用基于二进制蝙蝠算法（BBA）的判别故障特征分析技术从原始产生的特征向量中过滤出最具鉴别性的故障特征。最后将提出的方法应用于一对多支持向量机中来识别低转速轴承的多损伤情况。

3、袁海满等<sup>[24]</sup>利用粒子群优化算法(PSO)的全局寻优搜索策略对相关向量机(RVM)的核函数的参数进行优化，解决了 RVM 自身无法取得最佳核函数参数的固有不足，并已经应用于电力变压器的故障诊断。

4、XiaoLi Zhang 等人<sup>[25]</sup>提出一种新的基于多变量集成的增量式支持向量机智能故障诊断方法。该方法可以应用于复杂复合故障及故障严重程度的诊断。

5、Md. Mamunur Rashid 等人<sup>[26]</sup>提出一种数据挖掘方法用于故障识别和异常检测。首先利用 FFT 对振动信号进行预处理提取频率信号，然后创建滑动窗口关联频率模式树（SAFP），利用 SAFP 算法得到相关频率模式，完成故障识别。

### 3.2 深度学习法

随着人工神经网络的研究进展，以及更复杂层次的深度学习算法的发明，已经发展了新颖的分类模型来处理故障检测及诊断任务。大部分浅学习模式会从信号中提取一些特征值，因此和原来信号相比，会有维度缩减的情形。连续小波转换的小波量值图再配合卷积神经网络，可以直接分类成正常的案例及异常的案例。这类的技术避免了忽略重要故障讯息的问题，在故障检测及诊断上的性能也有提升。此外，透过将信号转换为影像的建构过程，可以用卷积神经网络，透过振动影像的特征来识别故障的信号。

1、Jafar Zarei 等人<sup>[27]</sup>提出一种基于人工神经网络（ANNs）的智能方法用于检测轴承

损伤。将振动信号通过由神经网络设计的去除非轴承故障分量（RNFC）滤波器，利用模式识别技术进行故障分类。

2、Haidong Shao 等人<sup>[28]</sup>提出一种增强的深度特征融合方法用于旋转机械故障诊断。利用自消噪编码器（DAE）和收缩自编码器（CAE）构造了一种新的深度自动编码器来增强特征学习能力，采用局部保持投影（LPP）融合深度特征来改善学习特征的质量，融合的深层特征送入 SoftMax 训练智能诊断模型。

3、Krishna K R 等<sup>[29]</sup>提出了一种新的卷积深度置信网络（CDBN）用于轴承故障诊断。用自动编码器用于压缩数据并减小尺寸，用高斯可见单元构造新颖的 CDBN 以学习代表性特征，并通过指数移动平均值用于改善构建的深层模型的性能。

4、Wen Long 等人<sup>[30]</sup>提出了一种基于 LeNet-5 的新 CNN 用于故障诊断，通过将信号转换成二维图像的转换方法，该方法可以提取转换后的二维图像的特征，能够消除人工提取的影响。



### 三、发展趋势预测

传统的单一人工智能方法对现代大型复杂设备进行故障时，具有精度低，通用性差等缺点。针对上述困难，个人认为可以将不同的人工智能技术以某种方式结合、集成或融合，如 Bagging、Boosting 等集成方式，形成混合智能故障诊断技术，提高诊断与预示系统的敏感性、鲁棒性和精确性，具有解决复杂问题的能力。如集成方式如下图 3 所示。

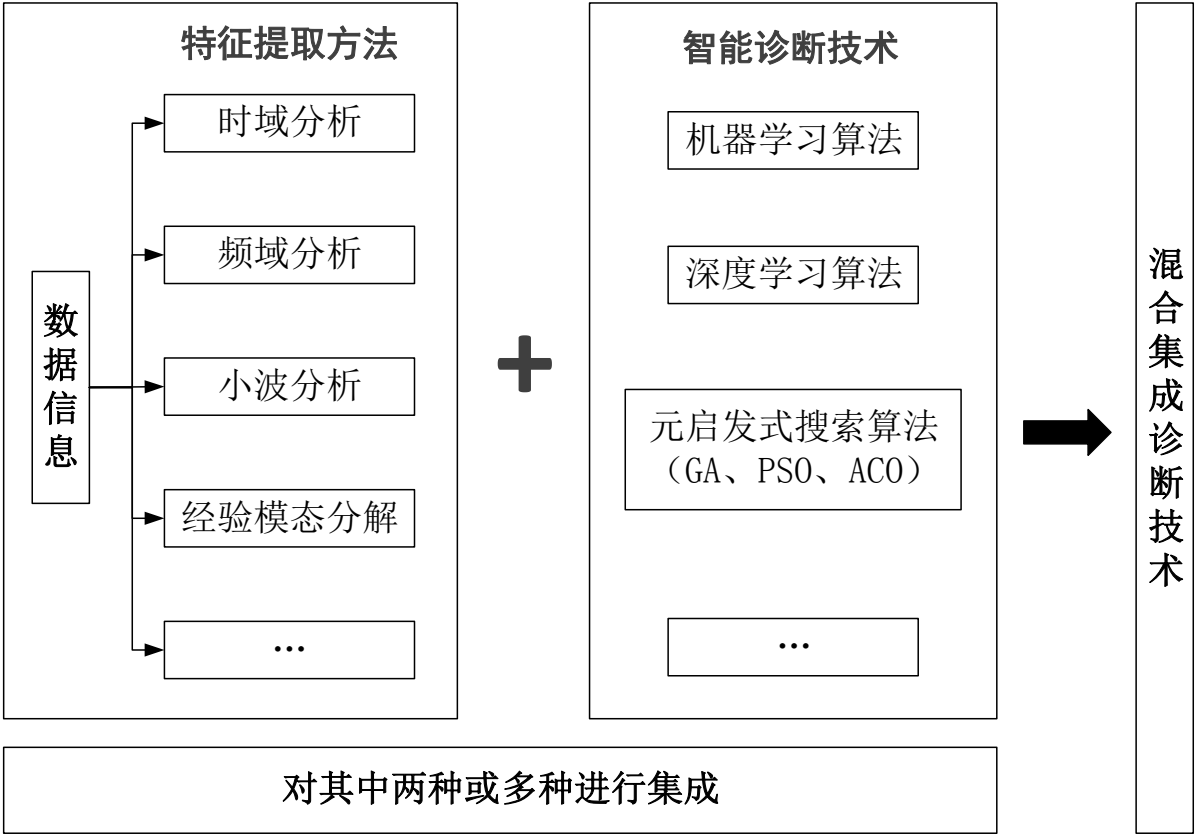


图 3 故障诊断集成技术结构

在如今工业 4.0、物联网的发展洪流下，机械设备的自动化程度会越来越高，多种物力场的耦合程度会不断增加，未来发展要重视多种信息的融合和多种不同方面特征的提取，对不同智能技术进行更深层次的融合与集成，这样才能充分发挥混合智能故障诊断技术的优势，有效解决机械早期故障的预示与诊断。

## 四、参考文献

- [1] LIU Chun-sheng, HU Shou-song. A Kind of Intelligent Fault Diagnosis of Nonlinear Systems Based on State Estimation[J]. Journal of Control and Decision, 2005,20(5): 557—561.
- [2] BELLALI B, HAZZAB A, BOUSSERHANE I K, et al. Parameter Estimation for Fault Diagnosis in Nonlinear Systems by ANFIS[J]. Procedia Engineering, 2012,29(4): 2016—2021.
- [3] Huang N E, Shen Z, Long S R, et al. The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis[J]. Proceedings Mathematical Physical & Engineering Sciences, 1998, 454(1971):903-995.
- [4] Rai V K, Mohanty A R. Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert–Huang transform[J]. Mechanical Systems & Signal Processing, 2007, 21(6):2607-2615.
- [5] 曹冲锋, 杨世锡, 杨将新. 大型旋转机械非平稳振动信号的 EEMD 降噪方法[J]. 振动与冲击, 2009, 28(9):33-38.
- [6] 戴桂平, 刘彬. 基于小波去噪和 EMD 的信号瞬时参数提取[J]. 计量学报, 2007, 28(2):158-162.
- [7] Yang Y, Yu D, Cheng J. A fault diagnosis approach for roller bearing based on IMF envelope spectrum and SVM[J]. Measurement, 2007, 40(9–10):943-950.
- [8] Nanjing. Application of Wavelet Envelope Analysis to Rolling Bearing Diagnosis[J]. China Mechanical Engineering, 2000.
- [9] Nikolaou N G, Antoniadis I A. Rolling element bearing fault diagnosis using wavelet packets[J]. Coal Mine Machinery, 2009, 35(3):197-205.
- [10] Purushotham V, Narayanan S, Prasad S A N. Multi-fault diagnosis of rolling bearing elements using wavelet analysis and hidden Markov model based fault recognition[J]. Ndt & E International, 2005, 38(8):654-664.
- [11] Piñeyro J, Klempnow A, Lescano V. Effectiveness of new spectral tools in the anomaly detection of rolling element bearings[J]. Journal of Alloys & Compounds, 2000, 310(1):276-279.
- [12] Mori K, Kasashima N, Yoshioka T, et al. Prediction of spalling on a ball bearing by applying the discrete wavelet transform to vibration signals[J]. Wear, 1996, 195(1–2):162-168.
- [13] Djebala A, Ouelaa N, Hamzaoui N. Detection of rolling bearing defects using discrete wavelet analysis[J]. Meccanica, 2009, 43(3):339-348.

[14] Al-Raheem K F, Roy A, Ramachandran K P, et al. Rolling Element Bearing Fault Diagnosis Using Laplace-Wavelet Envelope Power Spectrum[J]. Eurasip Journal on Applied Signal Processing, 2007, 2007(1):70-70.

[15] Tabrizi A, Garibaldi L, Fasana A, et al. Early damage detection of roller bearings using wavelet packet decomposition, ensemble empirical mode decomposition and support vector machine[J]. Meccanica, 2015, 50(3):865-874.

[16] Rubini R, Meneghetti U. APPLICATION OF THE ENVELOPE AND WAVELET TRANSFORM ANALYSES FOR THE DIAGNOSIS OF INCIPIENT FAULTS IN BALL BEARINGS[J]. Mechanical Systems & Signal Processing, 2001, 15(2):287-302.

[17] Antonino-Daviu J, Rodriguez P J, Riera-Guasp M, et al. Transient detection of eccentricity-related components in induction motors through the Hilbert–Huang Transform[J]. Energy Conversion & Management, 2009, 50(7):1810-1820.

[18] Pan M C, Tsao W C. Using appropriate IMFs for envelope analysis in multiple fault diagnosis of ball bearings[J]. International Journal of Mechanical Sciences, 2013, 69(4):114-124.

[19] Sun W, Yang G A, Chen Q, et al. Fault diagnosis of rolling bearing based on wavelet transform and envelope spectrum correlation[J]. Journal of Vibration & Control, 2013, 19(6):924-941.

[20] Cheng J, Yang Y, Yu D. The envelope order spectrum based on generalized demodulation time–frequency analysis and its application to gear fault diagnosis[J]. Mechanical Systems & Signal Processing, 2010, 24(2):508-521.

[21] Kang M, Kim J, Wills L M, et al. Time-Varying and Multiresolution Envelope Analysis and Discriminative Feature Analysis for Bearing Fault Diagnosis[J]. IEEE Transactions on Industrial Electronics, 2015, 62(12):7749-7761.

[22] Wei Z, Wang Y, He S, et al. A novel intelligent method for bearing fault diagnosis based on affinity propagation clustering and adaptive feature selection[J]. Knowledge-Based Systems, 2016, 116(C):1-12.

[23] Myeongsu Kang, Jaeyoung Kim, Jong-Myon Kim. Reliable fault diagnosis for incipient low-speed bearings using fault feature analysis based on a binary bat algorithm[J]. Information Sciences, 2015, 294(C):423-438.

[24] YUAN Hai-man, WU Guang-ning, GAO Bo. DGA Based Fault Diagnosis of

Electromechanical Power Transformer Based on Limit Learning of Particle Swarm Optimization[J]. High Voltage Apparatus, 2016, 52(11): 176—180.

[25] Zhang X L, Wang B J, Chen X F. Intelligent fault diagnosis of roller bearings with multivariable ensemble-based incremental support vector machine[J]. Knowledge-Based Systems, 2017, 89(Switzerland):56-85.

[26] Rashid M M, Amar M, Gondal I, et al. A data mining approach for machine fault diagnosis based on associated frequency patterns[J]. Applied Intelligence, 2016, 45(3):1-14.

[27] Zarei J, Tajeddini M A, Karimi H R. Vibration analysis for bearing fault detection and classification using an intelligent filter[J]. Mechatronics, 2017, 24(2):151-157.

[28] Shao H, Jiang H, Zhao H, et al. An enhancement deep feature fusion method for rotating machinery fault diagnosis[J]. Knowledge-Based Systems, 2016.

[29] Wen L, Li X, Gao L, et al. A New Convolutional Neural Network-Based Data-Driven Fault Diagnosis Method[J]. IEEE Transactions on Industrial Electronics, 2018, 65(7): 5990-5998.

[30] Krishna K R, Ramachandran K I. Machinery Bearing Fault Diagnosis Using Variational Mode Decomposition and Support Vector Machine as a Classifier[C]//IOP Conference Series: Materials Science and Engineering. IOP Publishing, 2018, 310(1): 012076.

## 附录（最近一篇文献阅读笔记）

	Journal Title: IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS IF: 7.168
ARTICLE INFO	A New Convolutional Neural Network-Based Data-Driven Fault Diagnosis Method Available 2018 Author: Long Wen, Xinyu Li, Liang Gao, Member, IEEE, and Yuyan Zhang
<p>（阅读日期：2018-11-13）一种基于数据驱动的卷积神经网络故障诊断方法</p> <p><b>摘要：</b>由于对新出现问题的早期检测可以节省宝贵的时间和成本，这在制造系统中至关重要。随着智能制造的发展，数据驱动故障诊断成为一个热门话题。然而，传统数据驱动的故障诊断方法依赖于专家提取的特征。特征提取过程是一项繁重的工作，且对最终结果有很大影响。深度学习（DL）提供了一种自动提取原始数据特征的有效方法。卷积神经网络（CNN）是一种有效的 DL 方法。在此研究中，采用的是一种基于 LeNet-5 的新 CNN 用于故障诊断。通过将信号转换成二维图像的转换方法，该方法可以提取转换后的二维图像的特征，能够消除人工提取的影响。所提出的方法在三个著名数据集上测试（电机轴承数据集、自吸式离心泵数据集和轴向柱塞液压泵数据集），分别达到 99.79%，99.481% 和 100% 的预测精度。结果与其他 DL 及传统方法进行了比较，包括自适应深度 CNN、稀疏滤波、深度置信网络和支持向量机。比较表明，提出的基于 CNN 的数据驱动故障诊断方法取得了显著的改进。</p> <p><b>关键词：</b>故障诊断 卷积神经网络 图像分类 数据驱动</p> <p><b>方法&amp;原理</b></p> <ol style="list-style-type: none"><li>1、开发了一种新的数据预处理方法“信号-图像转换方法”来提取原始数据的二维特征，而不需要预定义的参数。</li><li>2、提出了一种改进的 CNN，采用更深的结构和零填充方法来增加特征的非线性。</li><li>3、基于 CNN 的故障诊断方法与其他 DL 和传统方法相比有显著的改进。</li></ol> <p><b>核心思想简介：</b></p> <ol style="list-style-type: none"><li>1、信号-图像转换方法</li></ol> <p>如图 1 所示，在“信号-图像转换方法”中，时域原始信号按时域顺序转换成图像的像素点。为了获得 <math>M \times M</math> 尺寸的图像，先从原始信号中随机获得长度为 <math>M^2</math> 的</p>	

分段信号。令  $L(i)$  代表分段信号的值，其中  $i=1..M^2$ ； $P(j,k)$  表示图像的像素强度，其中  $j=1..M$ ， $k=1..M$ ，如等式(1)所示。

$$P(j,k) = \text{round} \left\{ \frac{L((j-1) \times M + k) - \text{Min}(L)}{\text{Max}(L) - \text{Min}(L)} \times 255 \right\} \quad (1)$$

(PS: 对公式的理解， $L((j-1) \times M + k)$  代表是第  $j$  行第  $k$  个像素点的信号值，后面是归一化的计算方式，把信号值投射到 0-255 之间)

函数  $\text{round}(\cdot)$  是舍入函数，整个像素值已经从 0 到 255 进行归一化，这只是灰度图像的像素强度。本文通常使用  $2 \times 2$  滤波器，每层图像特征的大小将减少一半，因此  $M$  的推荐值为  $2^n$ ，如 16、32、64、128 等。在本文中， $64 \times 64$  和  $16 \times 16$  的选择取决于信号数据量。

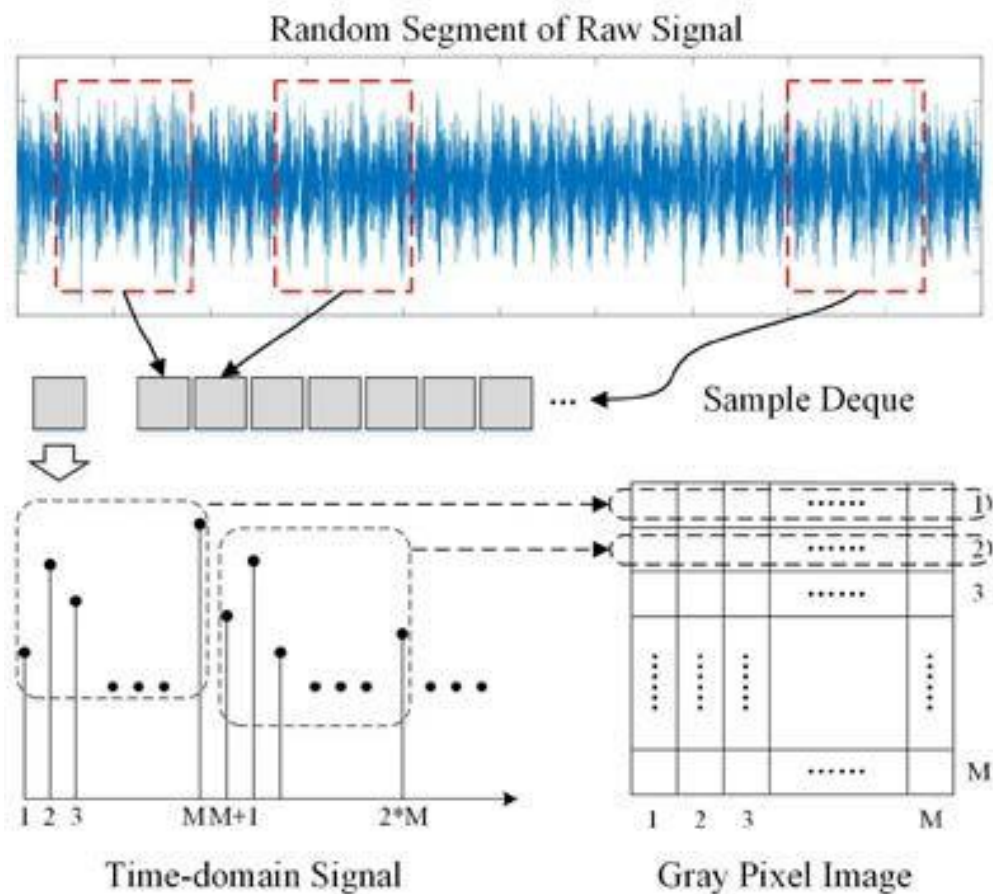


图 1 信号-图像的转换方法

这种数据处理方法的优点是它提供了一种探索原始信号的二维特征的方法。更重要的是，这种数据预处理方法可以在没有任何预定义参数的情况下进行计算，并且可以尽可能减少专家的经验。

## 2、改进的 CNN 结构

LeNet-5 中的图像尺寸为  $32 \times 32$ 。但为了改善分类结果，图像大小根据本研究中信号量的大小而变化。在第一和第二种情况下，信号数据的体积较大，而较大的

图像大小可以促进分类结果。所以这些情况下的图像尺寸是  $64 \times 64$ 。相反，由于信号数据量在第三种情况下相对较小，因此图像的大小为  $16 \times 16$ 。

在所提出的 CNN 模型中，它们包含四个交替的卷积层和池层，其中一个或两个完全连接的层用于  $64 \times 64$  图像。虽然所提出的 CNN 模型仅包含两个交替的卷积层和共用层，其中一个完全连接层用于  $16 \times 16$  图像。填充方法也与原来的 LeNet-5 不同，在我们的研究中使用了零填充。图 2 显示了所提出的用于  $64 \times 64$  图像的 CNN 模型的基本结构。

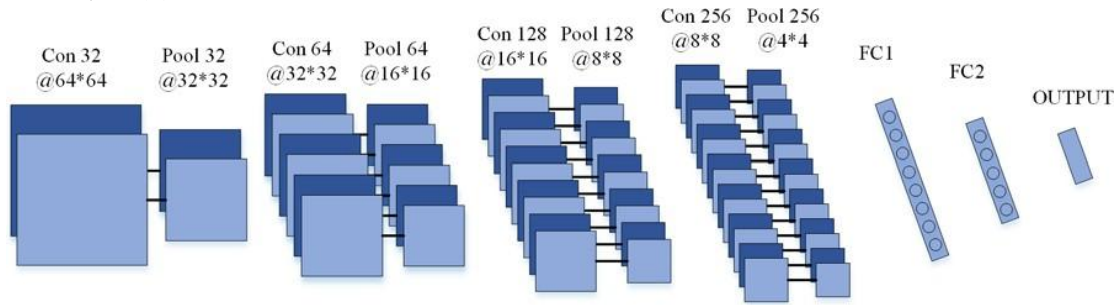


图 2 提出的  $64 \times 64$  图像的 CNN 结构

### 3、零填充

填充方法是控制特征尺寸大小的重要技术。将零填充方法应用于 CNN 模型以防止尺寸损失，如图 3 所示。设  $M$  为输入尺寸， $N$  为输出尺寸， $F$  为滤波器宽度， $S$  为步幅。左侧  $PL$  和右侧  $PR$  上的填充数量可以用公式(2-5)计算。 $\text{ceil}(\cdot)$ 和  $\text{floor}(\cdot)$ 是  $\text{ceil}$ 和  $\text{floor}$  函数，(PS 分别是向上取整和向下取整)。

$$N = \text{ceil}\left(\frac{M}{S}\right) \quad (2)$$

$$PT = (N - 1) \times S + F - M \quad (3)$$

$$PL = \text{floor}\left(\frac{PT}{2}\right) \quad (4)$$

$$PR = PT - PL \quad (5)$$

在零填充方法中，零将自动添加以填充卷积过程。图 3 显示了一维的零填充方法的一个例子。参数为  $M=5$ ， $S=1$ ， $F=3$ ，则填充结果为  $PL=1$ ， $PR=1$ ， $N=5$ 。应该指出的是，本研究中卷积层和汇集层的步幅设置为 1。

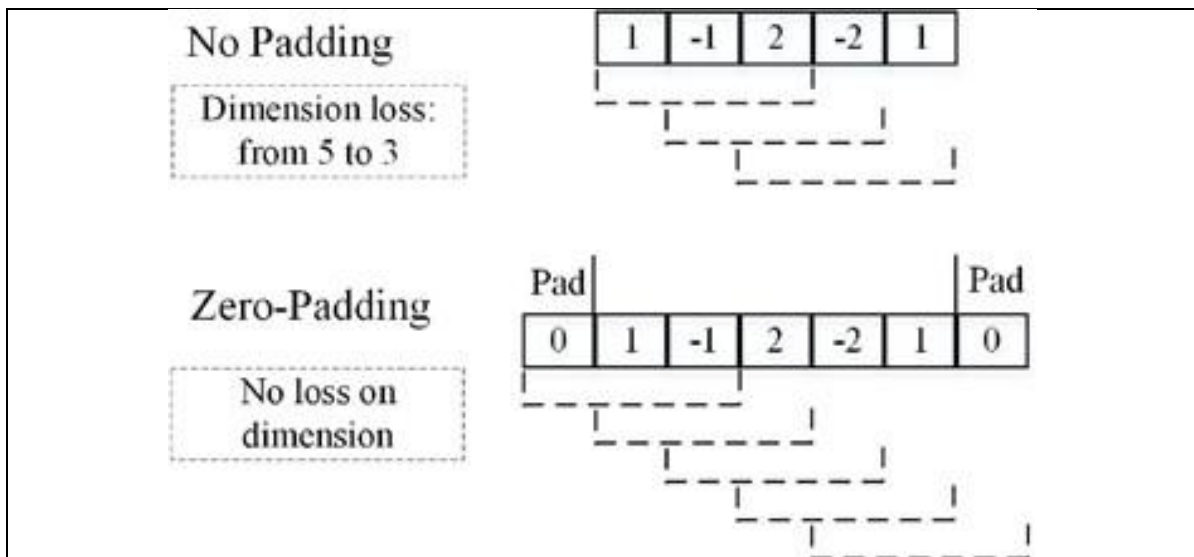


图 3 CNN 中使用的零填充方法

### 文献阅读小结

该文献阅读后，个人觉得可以从两方面进行思考：

第一查看关于主轴、丝杠等部件论文是否用过这种将原始信号算法转换成图像的算法。

第二在接下来的文献阅读中思考，是不是可以将原始信号通过另一种方式转换成图像，然后在应用这种算法，比如直接对原始信号与时间的曲线图作为图像源输入，或者对原始信号先进行某种变换（小波变换、HHT 变换），然后在应用这种 CNN 算法进行处理。