

Rolling Bearing Anomaly Detection Based on Generative Adversarial Networks

Feng Hua

Huazhong University of Science and Technology, Wuhan Hubei
Email: 2248776173@qq.com

Received: Oct. 25th, 2019; accepted: Nov. 7th, 2019; published: Nov. 14th, 2019

Abstract

The imbalance of sample data is common in industrial system, and the number of normal samples is much larger than the number of abnormal samples. However, machine learning algorithms and deep learning methods, such as Naive Bayes (NB) and Convolutional Neural Networks (CNN), are difficult to obtain high recognition and classification accuracy when dealing with class imbalance problems, because they tend to guarantee the accuracy of majority classes. This paper proposes an anomaly detection method based on Generative Adversarial Networks (GAN). The generator structure in this method is a triple subnet of "encoder-decoder-encoder", and the training generator only needs to extract features from normal samples, so there is no need for abnormal samples in the training data set. The anomaly detection of the system is determined by the final score of the sample, in which abnormal scores are composed of apparent loss and potential loss. The highlight of the method in this paper is that it can detect abnormal data samples without training abnormal sample and diagnose faults by generating higher abnormal scores in the system. This project has verified the feasibility and effectiveness of this method on the rolling bearing data set obtained from Case Western Reserve University (CWRU). The accuracy of the method proposed in this paper in distinguishing abnormal samples from normal samples in the data set is 100%.

Keywords

Imbalanced Industrial Time Series, Anomaly Detection, Generated Antagonistic Network, Rolling Bearing Data

基于对抗生成网络的滚动轴承故障检测方法

华 丰

华中科技大学, 湖北 武汉
Email: 2248776173@qq.com

收稿日期: 2019年10月25日; 录用日期: 2019年11月7日; 发布日期: 2019年11月14日

摘要

在工业系统中普遍存在样本数据不平衡现象，正常样本数量远远大于异常样本数量。而传统的机器学习算法和深度学习方法，例如朴素贝叶斯和支持向量机，在处理类不平衡问题时，很难获得较高的识别分类准确率，因为它们往往会偏向保证多数类的准确率。为此，本文提出了一种基于生成对抗网络(GAN)的异常检测方法。这个方法中的生成器结构是“编码器-解码器-编码器”的三子网，并且训练该生成器只需要从正常样本中提取特征，所以训练数据集中就不需要异常样本。此系统的异常检测结果由样本的最终得分来判别，其中异常分数由表观损失和潜在损失组成。本文方法的亮点在于可以实现在无异常样本训练的情况下对异常数据样本做检测，通过系统生成更高的异常分数来诊断故障。本项目在凯斯西储大学(CWRU)获得的基准滚动轴承数据集上验证了该方法的可行性和有效性。本文提出的方法在数据集中区分异常样本与正常样本的准确率达到100%。

关键词

不平衡工业时间序列，异常检测，生成对抗网络，滚动轴承数据

Copyright © 2019 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 项目介绍

异常检测对于现代工业系统的可靠性和安全性至关重要。及时准确的异常检测有助于预防重大事故的发生，提高生产效率。然而，工业生产中数据类不平衡的情况比较严重，在正常条件下的样本比在异常条件下的样本普遍得多，为准确诊断工业设备故障造成了巨大的障碍。此外，工业系统总是具有非线性和不确定性，这对模型训练提出了很大的挑战[1]。

工业异常检测的数据一般是不同传感器在一定时间内记录的电流、温度等物理信号，也称为时间序列。对于工业异常检测领域，时间序列通常作为训练模型的输入数据[2]。一般以时间序列为输入，异常检测框架通常分为特征提取和故障识别两个阶段。通过特征提取算法，将时间序列预处理为低维特征向量，送入故障检测器进行故障检测[3]。机器学习算法作为一种强大的异常检测模式识别工具，已经成为关注的焦点，包括贝叶斯分类器、支持向量机方法等[4]。上述方法都是假设基于类平衡的情况，然而当数据分类不平衡时这些方法难以获得较高的精度[5]。除了数据集的类均衡假设外，标记数据也是训练阶段机器学习算法的关键。然而，在许多实际的工业系统中，来自异常情况下的样本数量往往很少。另外，当系统在正常状态下运行了很长一段时间后，突然出现异常，要准确定位异常的发生时间是极其困难的[6]。因此，不准确的异常标签也会对异常检测的准确性产生不利的影响。

当正常和异常的标签不平衡时，机器学习方法的分类器会牺牲少数类来保证多数类的准确性，这意味着分类结果会偏向于测试样本整体的正态性[5]。但是，对于工业系统中的异常检测，我们应该特别关注那些处于少数情况的类，如何能够准确捕捉和判别异常数据是当前工业系统中异常检测的重点。在2014年，由Goodfellow等人提出的生成对抗网络(GAN)为解决工业中类不平衡问题提供了一个新的思路。这个网络模型最先出现在“Generative Adversarial Networks”一文中，起初它被用于图像识别领域，并取得了卓越的成绩[7]。GAN的基本思想是通过一个具有随机数据点的生成器生成原型样本，这些随机数据点

满足一定的分布(如高斯分布)。在图像的异常诊断领域,已经有一些基于 GAN 的具有竞争力的网络架构被提出,如 AnoGAN [8]、BiGAN [9]和 GANomaly [10]。这些基于 GAN 的方法只训练正常图像的模型,根据正常图像和异常图像的特征分布差异来区分异常图像。从这一点出发,基于 GAN 的模型对于学习和识别不平衡数据集是有效的,可以防止诊断结果偏向于多数类。然而,在工业应用中,基于 GAN 的异常检测方法非常少见。经过类似项目的调查研究,发现依次有基于 GAN 网络的机械故障检测方法[11]和基于 GAN 的不平衡数据故障诊断方法[12]被提出。这些研究启发可以进一步研究 GAN 在工业故障检测方面的有效性,特别是在没有异常数据的情况下。因此针对工业时间序列的特性,本文基于 GANomaly,改善了生成器整体的损失函数,以实现类不平衡场景的检测高精度。

在这项工作中,针对工业系统中异常数据的匮乏,本文提出了一种基于 GAN 的方法来解决智能异常检测问题。这个方法的系统由一个生成器和一个判别器组成,将生成器和判别器进行对抗训练得到训练后的模型来用于诊断。该系统的生成器基于卷积生成对抗网络,采用“Encoder-Decoder-Encoder”为结构的三子网络。为了提高诊断性能,在原始数据和 GAN 之间插入数据转换器对原始数据集进行预处理。通过在 CWRU 滚动轴承数据集的实验,验证了该方法的可行性和有效性。

本文的主要贡献如下:

a) 针对工业领域的不平衡时间序列问题,提出了一种新的基于 GAN 的异常检测方法。

b) 所提出的网络训练中只需要正常样本。这是一个比其他现有算法网络更贴合现实场景的网络。因为在真实的工业场景中,异常样本的数量往往非常少。实验基于 CWRU 的滚动轴承基准数据集进行了训练和调整,通过测试确保了该系统模型具有良好的诊断性能,并进一步验证了本文方法的有效性,使其可以作为解决工业数据类不平衡问题的新思路。

2. 项目背景

2.1. 深度学习用于智能故障诊断

随着各种工业系统中产生的时间序列数据流越来越动态、复杂和庞大,许多用于深度学习的异常检测技术得到了很好的发展。在调用这些深度学习模型的过程中,它们被赋予学习目的,但这并不意味着它们清楚最终的输出特性是什么。这些黑箱旨在为特定的数据集提取特定的模式,如长短时记忆网络(LSTM),递归神经网络(RNN),卷积神经网络(CNN)。虽然上述的这些深度学习模型在异常检测方面显现出了许多卓越的性能,但在面对不平衡的数据集时,其准确率仍然不能令人满意[13]。此外,由于数据表示的差异(如图像和时间序列),许多深度学习模型适用于图像领域,但在工业领域往往难以实施和应用。

2.2. 类不平衡问题

在基于不平衡时间序列的异常检测中,通常会考虑两种主要方法:数据级方法和算法级方法[14]。数据级方法通常利用采样策略改变不平衡的数据分布,其中广泛使用过采样和欠采样两种策略[15]。算法级的方法一般采用调整分类器以适应不平衡的数据,其中通常使用 bagging 和 boost ensemble-based 方法[16]。具体的有 Easy Ensemble [17]和 Balance Cascade [18]算法被提出用于处理类不平衡问题。Easy Ensemble [17]是通过多次从多数类样本有放回的随机抽取一部分样本生成多个子数据集,将每个子集与少数类数据联合起来进行训练生成多个模型,然后集合多个模型的结果进行判断。这种方法看起来和随机森林的原理很相似[14]。Balance Cascade [17]是通过一次随机欠采样产生训练集,训练一个分类器,对于那些分类正确的多数类样本不放回,然后对这个剩下的多数类样本再次进行欠采样产生第二个训练集,训练第二个分类器,同样把分类正确的样本不放回,以此类推,直到满足某个停止条件,最终的模型也是多个分类器的组合[17]。

2.3. GAN 和 GANOMALY

近年来，在类不平衡图像的异常检测中，对抗训练尤其是 GAN 占据了越来越重要的地位。GAN 最早由 Goodfellow 等人提出，被认为是一种无监督机器学习算法，在图像识别领域取得了突出的应用效果。GAN 的网络结构如图 1，使用原始数据对生成器进行训练，生成器(Generator)产生新的样本。生成器数据和真实训练数据输入到判别器(Discriminator)中，训练得到能正确区分数据类型的判别器。判别器和生成器相互对抗，生成器通过学习来产生更逼近真实数据的新样本，用于欺骗判别器，反之判别器也需要更好地区分生成数据与真实数据。自 2014 年以来，基于 GAN 的对抗性算法层出不穷，很多新的方法和框架被提出并取得了优秀的表现。

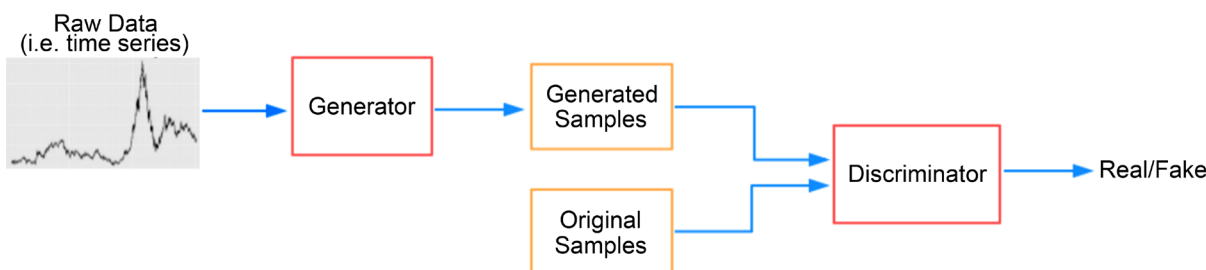


Figure 1. Structure of generate adversarial network
图 1. 生成对抗网络结构

在图像分类领域中，Akçay 等人提出了一种通用的异常检测体系结构 GANomaly。与以往的先进方法相比，该结构在多个基准图像数据集上具有显著的优越性和有效性[10]，给予我们在工业领域的异常检测一些启发。下面是对 GANomaly 的简要介绍。该模型中，其生成器由“编码器 - 解码器 - 编码器”三子网构成训练半监督网络，该结构使用深度卷积生成对抗网络(DCGAN) [18]，并在生成器中使用三个损失函数来捕获输入图像和潜在空间中的特征。该算法的特点之一是不考虑异常样本，并在图像数据集中实现了优异的异常检测性能。

论文的其余结构如下。第三部分提出了基于 GAN 的异常检测框架。第四节和第五节分别介绍了实验装置和实验结果。第六部分为结论和未来工作。

3. 项目方法

图 2 是本文提出的方法概述。图中可以看出在训练阶段，这个系统模型只考虑正常样本。本文使用数据转换器来对样本进行处理并提取有用的特征，这些特征经过转换可以更好的被模型学习。系统中生成器采用编码器 - 解码器 - 编码器三子网，判别器则是基于 DCGAN 网络结构。在检测阶段，异常样本可以通过比正常样本更高的异常分数来识别。

3.1. 方法介绍

给定一个时间序列数据集 $D = [X_1, X_2, X_3, \dots, X_n] \in R^{t \times n}$, $X_i = [x_{i1}, x_{i2}, x_{i3}, \dots, x_{iu}]^T \in R^t$ (for $i = 1, \dots, n$) 表示传感器在一段时间 t 内记录的样本。在数据集中划分出训练集和测试集，其中训练集用矩阵的形式可

表示为: $D_{train} = [X_1, X_2, X_3, \dots, X_b] = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1b} \\ \vdots & \vdots & \ddots & \vdots \\ x_{t1} & x_{t2} & \dots & x_{tb} \end{bmatrix}$ 。用训练集 $D_{train} \in R^{t \times b}$ (b 是训练样本的数量)的

一个子集训练本文所提出的模型。对应的测试样本是 $D_{test} = [D_{test}^v, D_{test}^u] \in R^{t \times (v+u)}$, v 和 u 分别是正常样本和异常样本的数量。

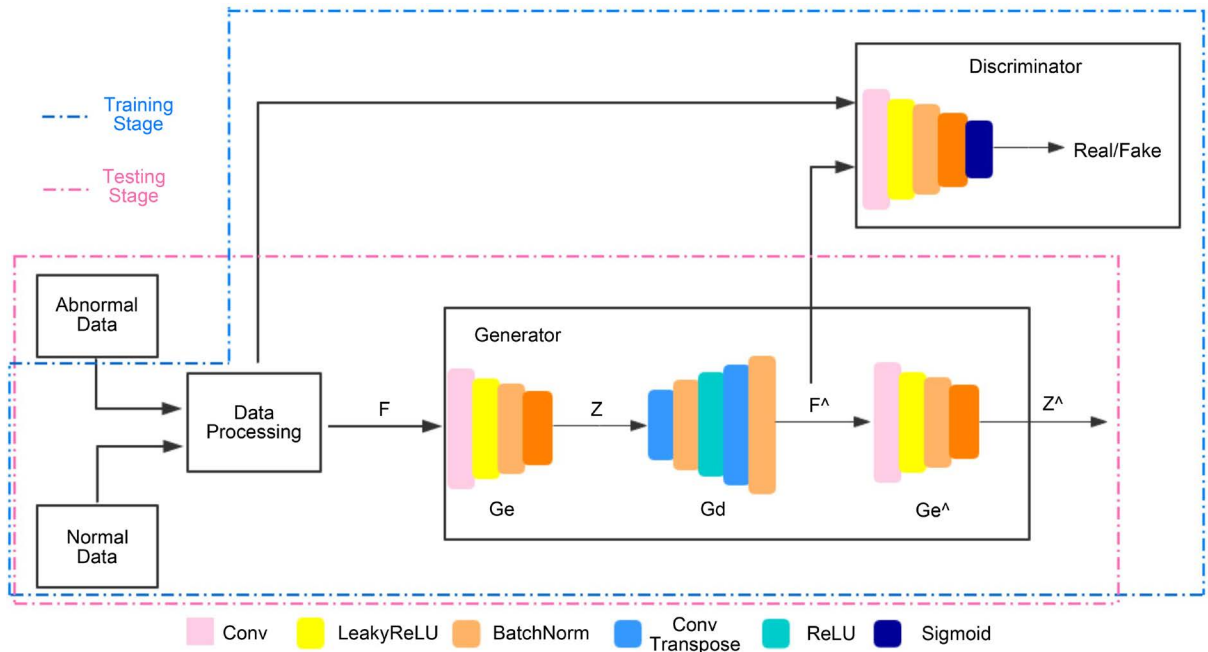


Figure 2. The proposed method framework
图 2. 本文方法框架

在训练的步骤中，用 D_{train} 训练一个 GAN-Based 模型 M。训练过程的目标是最小化 D_{train} 中每个子数据集的输出。在训练过后，用于测试的数据集 D_{test} 将会放入模型 M。训练后的生成器将对故障样本和正常样本进行相应的编码和解码。由于训练后的网络只学习了正常数据的可能表示模式，如果用异常数据样本 D_{test}^u 作为输入，那么模型 M 的输出与正常输入 D_{test}^n 的输出相比，会有很大的偏差。这个偏差值最终帮助确定异常样本的存在。

3.2. 训练的过程

本文提出的方法中其训练阶段的主要目的是正常条件下生成尽可能小的样本分数的模型。本文方法的网络结构由数据转换器、生成器和判别器三部分组成。基本网络架构用 DCGAN 表示。在生成器的设计中，开发了一种由“Encoder-Decoder-Encoder”组成的三子网。在将数据输入生成器之前，需要设计一个数据转换器将一维的时间序列数据转为二维的图像数据，可以帮助 DCGAN 更好的提取和学习样本特征。基于普适性原则，不对原始数据进行其他处理，因此任何时间序列都可以封装到这些特征中。在训练过程中，首先在正常条件下提取时间序列的特征，然后利用本文设计的异常检测器获得这些特征的数据分布和可能的代表模式。在测试阶段，将异常样本输入训练好的异常检测器，异常数据的特征分布会得到比正常样本更高的分数，以此为差异来识别和诊断。

3.2.1. 生成器和判别器

在生成器中，第一个编码器用于学习样本原始特征 F 的表征，第二个编码器用于生成再生特征 F^{\wedge} 。同时，解码器 G_d 用于试图重建 F^{\wedge} 。整个过程如下：

a) G_e 由卷积层，Batch-Norm 层和 Leaky ReLU 激活层组成。 G_e 将特征 F 降维到潜在表示 Z (通常 $h < q$)，其中 h 是潜在向量 Z 的维数， q 是训练集中每个子集的特征数量。

$$Z = G_e(F), Z \in R^{h \times b}, \tag{1}$$

b) G_d 采用卷积转置层(反卷积)、ReLU 激活层和 Batch-Norm 层构建。 G_d 用 Z 来重新创建 F^{\wedge} , F^{\wedge} 的维度和 F 是一样的。

$$F^{\wedge} = G_d(Z), \tag{2}$$

c) 编码器 G_e^{\wedge} 的结构和 G_e 是一样的但是参数不同, 输出 Z^{\wedge} 的维度和 Z 是相同的。

$$Z^{\wedge} = G_e^{\wedge}(F^{\wedge}), \tag{3}$$

该生成器保证了原始特征集 F 的特征不变的同时可以得到潜在向量 Z 的模式。判别器(Discriminator)采用 DCGAN 中引入的标准判别器网络, 用于判断输入数据是真实的还是生成的。根据判别器的反馈再对生成器进行调整和训练。在定义了整个网络架构之后, 本文将称述如何定义训练的损失函数。

3.2.2. 目标函数

在训练阶段, 因为只训练正常数据集 D_{train} , 所以模型 M 只会获得正常数据集下的模式。但是在测试阶段, M 需要通过输出更高的异常分数来确定异常样本。那就意味着 G_d 和 G_e 将会解码潜在表示 Z 和重新编码 F^{\wedge} , 这与在训练阶段获得的模式类似。之后 F^{\wedge} 和 Z^{\wedge} 将不可避免地原来的 F 和 Z 产生差异, 这样有助于我们识别异常故障。

由于生成器的结构是采用“Encoder-Decoder-Encoder”组成的三个子网络, 因此生成器的最终损失函数将由欺骗损失(Fraud Loss)、表面损失(Apparent Loss)和潜在损失(Latent Loss)三部分组成, 接下来将会分别对三个损失函数进行说明。

欺骗损失: 在这里引入 fraud loss 的目的是为了诱导判别器将生成器生成的样本误判为真实的工业样本。将生成的样本输入判别器, 通过判别器输出计算的 fraud loss, 公式如下:

$$L_f(F) = \sum_{i=1}^N \sigma(C(F^{\wedge}), \alpha) \tag{4}$$

公式中 σ 是二进制交叉熵损失函数, $C(F^{\wedge})$ 是样本 i 被预测为实数的概率。为了骗过判别器和促进生成器来生成尽可能真实的假样本, 我们定义 $\alpha = 1$ 。

表面损失: 生成器仅学习正常条件下的电位模式并尽可能真实地重构生成的样本是不够的, 因此我们通过测量真实样本与生成样本之间的 L_1 差异来定义表现损失 L_a , 公式如下:

$$L_a(F) = \sum_{i=1}^N \|F - F^{\wedge}\|, \tag{5}$$

潜在损失: 除了欺骗损失和表面损失外, 我们还定义了潜在损失, 以最小化真实样本的潜在表示与生成样本的编码瓶颈特征之间的距离。这种损失可以帮助学习真实样本和生成样本中的潜在表征。另外, 隐性损失与表现损失在测试阶段产生最终的异常分数。

$$L_l(F) = \sum_{i=1}^N \|G_e(F) - G(F)\|_2. \tag{6}$$

上述三个子损失函数构成了生成器的主体损失函数, 公式可以表示为:

$$L_g(F) = \omega_f * L_f(F) + \omega_a * L_a(F) + \omega_l * L_l(F), \tag{7}$$

在公式中 ω_f , ω_a , ω_l 被用来调整 L_f , L_a 和 L_l 在生成器损失函数中的权重。

在判别器中, Salimans 等人[19]提出采用特征匹配损失 \mathcal{L} 进行对抗学习来减少 GAN 训练的不稳定性。用公式可以表示为:

$$L_d(F) = \sum_{i=1}^N \|\mathcal{L}(F) - \mathcal{L}(G_d(G_e(F)))\|. \tag{8}$$

我们使用 Adam 优化器[20]来反向传播并更新方程(8)和(9)。

3.3. 测试步骤

在测试阶段，我们的模型使用潜在损失和表面损失对检测到的测试数据样本进行评分。判定样本分数的公式可以定义为：

$$T(F) = \lambda * L_a(F) + (1 - \lambda) * L_l(F) \tag{9}$$

在这一部分中，我们使用 ω_a ， ω_l 的比率作为合适的加权值超参数 λ ，意思是此处使用最好的训练结果，即最小的生成器损失和判别器损失。因为本文提出的方法只在正常数据样本上训练异常检测器，所以实验中的异常检测器将只学习和识别正常样本的潜在模式和数据分布。在这个判定公式中，正常样本得到的 $T(F)$ 将接近于 0，而对于异常样本，公式得到的值将相对大很多。在测试中，通过公式 $T(F)$ 的值的波动可以很容易地找到异常。

4. 实验设置

为了评估本文方法的可行性和有效性，实验和测试将在凯斯西储大学(CWRU)获得的滚动轴承数据上进行，实验台如图 3 所示。

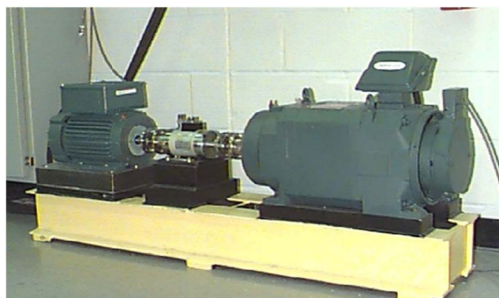


Figure 3. Testing bed in CWRU
图 3. CWRU 实验台

4.1. 数据集描述

这是一个基准轴承异常检测数据集，通过在电机上使用加速度计测量轴承的振动信号得到。采用电火花加工方法对电机轴承进行了故障样本创建。分别在内滚道(IR)，滚动部件(*i.e.* Ball)和外滚道(OR)创建了直径范围在 0.007 英寸到 0.028 英寸的故障点。故障轴承被重新安装到测试电机中，并记录电机负载从 0 到 3 马力(电机速度从 1720 到 1797 rpm)。数据集的采样频率为 48 kHz，每个数据集文件由三种类型的信号组成，即驱动端加速度计信号、风扇端加速度计信号和基础加速度计信号。表 1 总结了实验中数据集的详细信息。

Table 1. Data parameters collected by CWRU

表 1. CWRU 数据集的参数

	CWRU 数据
信号类型	振动信号
信噪比	高
采样频率(kHz)	48
运转速度(rmp)	1730, 1750, 1772, 1797
故障直径	0.007, 0.014, 0.021, 0.028
故障类型	IR, OR, B

4.1. 数据集处理

我们将正常数据样本分为训练集和测试集两个部分。对于来源于 CWRU 的滚动轴承数据集，得到的初始训练数据集 D_{train} 都是传感器在一段时间 t 内记录的样本。这些数据可以写成矩阵型式，

$$D_{train} = [X_1, X_2, X_3, \dots, X_b] = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1b} \\ \vdots & \vdots & \ddots & \vdots \\ x_{t1} & x_{t2} & \dots & x_{tb} \end{bmatrix}. \quad (10)$$

在项目的实际训练中发现样本的数量以及样本的多样性需要得到进一步扩充，因此采用了对初始数据进行分割的方法来处理数据集。将时间 t 内记录的样本进行分割，依次取连续不间断的 3136 个数据点作为新样本。这个新数据样本是一个 1×3136 的矩阵 $D_i = [X_{i1}, X_{i2}, X_{i3}, \dots, X_{i3136}]$ 。然后将数据转换成 $4 \times 28 \times 28$ 的矩阵，此处可以理解为将一维的原始数据转变成了 4 张尺寸为 $1 \times 28 \times 28$ 的图像数据作为输入模型的训练数据，经过处理得到训练集 $D_{train} \in R^{t \times b}$ (b 是训练样本的数量)。众所周知 GAN 这一网络模型最初就是在图像识别领域展现了优异的性能，因此对数据进行该项处理可以优化模型对样本特征的学习和提取。

对应的测试样本是 $D_{test} = [D_{test}^v, D_{test}^u] \in R^{t \times (v+u)}$ ， v 和 u 分别是正常样本和异常样本的数量。所以这个项目的样本总数 n 可以表示为 $n = b + v + u$ 。在 CWRU 的数据集中，实验使用正常状态条件下 $b = 400$ 的样本来训练异常检测器，同时设置了 $v = 541$ 和 $u = 383$ 的没有标签的样本用于测试集。

4.3. 实施细节

实验的具体实施放在 PyTorch 中，并使用了 Adam 优化网络来实现本文提出的方法。首先设定训练模型的参数，初始化学习率设为 $lr = 0.0001$ ，动量 β_1, β_2 的值分别设为 0.5 和 0.999。对于两个数据集，每个模型在训练时 epochs 都设为 50，batch-size 即一次训练样本的数量设为 32。本文提出的方法的具体网络结构在表 2 中进行了详细的说明。

Table 2. Model network structure
表 2. 模型网络结构

Layer	Filters	Ks/S1	Output Size
Encoder			
Input	-	-	$4 \times 28 \times 28$
Conv1 (LeakyReLU1/Batchnorm1)	32	$3 \times 3/1 \times 1$	$32 \times 28 \times 28$
Conv2 (LeakyReLU2/Batchnorm2)	32	$4 \times 4/2 \times 2$	$32 \times 13 \times 13$
Conv3 (LeakyReLU3/Batchnorm3)	64	$3 \times 3/1 \times 1$	$64 \times 13 \times 13$
Conv4 (LeakyReLU4/Batchnorm4)	32	$3 \times 3/2 \times 2$	$32 \times 6 \times 6$
Conv5	16	$3 \times 3/1 \times 1$	$16 \times 6 \times 6$
Decoder			
Input	-	-	$16 \times 6 \times 6$
ConvTrans1 (Batchnorm1/ReLU1)	32	$3 \times 3/1 \times 1$	$32 \times 6 \times 6$
ConvTrans2 (Batchnorm2/ReLU2)	64	$3 \times 3/2 \times 2$	$64 \times 13 \times 13$
ConvTrans3 (Batchnorm3/ReLU3)	32	$3 \times 3/1 \times 1$	$32 \times 13 \times 13$
ConvTrans4 (Batchnorm4/ReLU4)	32	$4 \times 4/2 \times 2$	$32 \times 28 \times 28$
ConvTrans5 (Tanh)	4	$3 \times 3/1 \times 1$	$4 \times 28 \times 28$

Continued

	Discriminator		
Input	-	-	$4 \times 28 \times 28$
Conv1 (Batchnorm1/LeakyReLU1)	32	$3 \times 3/1 \times 1$	$32 \times 26 \times 26$
Conv2 (Batchnorm2/LeakyReLU2)	32	$3 \times 3/2 \times 2$	$32 \times 12 \times 12$
Conv3 (Batchnorm3/LeakyReLU3)	64	$3 \times 3/1 \times 1$	$64 \times 10 \times 10$
Conv4 (Batchnorm4/LeakyReLU4)	64	$3 \times 3/2 \times 2$	$64 \times 4 \times 4$
Conv5	128	$3 \times 3/1 \times 1$	$128 \times 2 \times 2$
Flatten	-	-	1×512

Ks/S^{-1} = Kernel Size/Stride.

5. 实验结果

在实验时，将凯斯西储大学(CWRU)收集的滚动轴承数据输入本文方法构建的系统中，生成器经过训练集数据的训练生成初始的网络模型，接着将生成器生成的样本输入判别器，根据判别器反馈的结果，对生成器进行新的训练，直到判别器能成功被生成器生成的样本所欺骗，此时生成器的训练收敛。接着调用生成器训练后的网络模型对测试样本进行测试，得到并输出该样本数据的得分情况。所提出方法的CWRU数据集得分如图4所示。

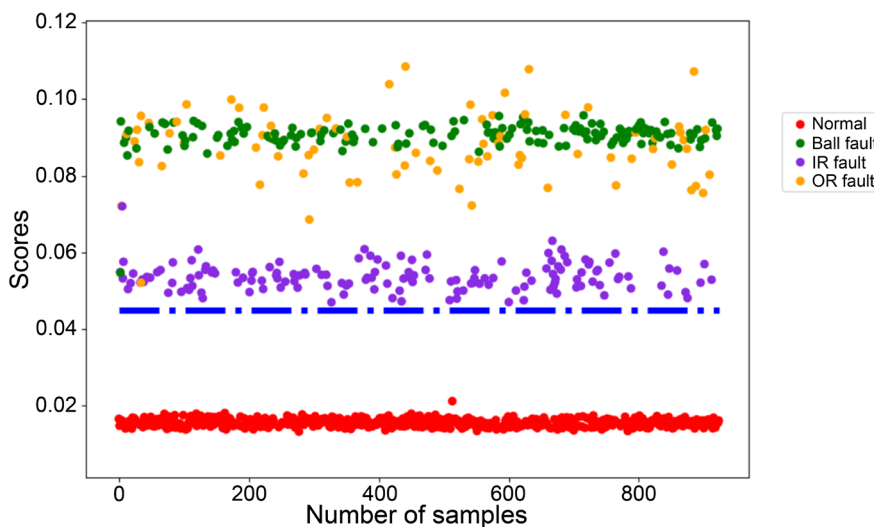


Figure 4. CWRU data experimental results

图 4. CWRU 数据实验结果

从这张输出结果图中可以看出因为滚动轴承的故障情况不同，所以每种情况所得分值也明显不同。同时表示阈值的蓝色虚线清晰的分隔了正常和异常样本。红色的数值点是正常样本在生成器中的得分情况，可以清楚的看出正常样本的 Scores 值区间小于 0.02 并总体呈均匀分布，此处可以进一步验证该生成器的训练收敛了。图中紫色的点代表内滚道故障样本的得分情况，该类型样本的整体分布较为混乱，得分区间在[0.04, 0.06]之间。图中黄色的点代表外滚道故障样本的得分情况，该类型的整体得分情况最为分散。此外图中绿色的点代表滚动部件故障样本的得分情况，此故障的得分情况相较于其余两种较为稳定，得分区间在[0.08, 0.10]之间。该得分结果还表明，滚动故障(图中的 Ball fault)数据的模式与正常样本数据的模式最接近，而外滚道故障(图中的 OR fault)与正常数据模式的差异最大。

6. 项目结论

本文提出了一种基于生成对抗网络(GAN)的针对工业时间序列不平衡的异常检测体系结构。与当前大多数的深度学习方法相比,这种体系结构的特点是只需要使用正常的的数据样本进行训练,不用考虑工业系统中样本分类不均衡的问题。在异常检测之前,我们根据本项目工业数据的特征和网络模型的特点设计了一个数据转换器。然后利用“Encoder-Decoder-Encoder”构建的生成器输出较大的异常分数来检测异常样本的存在。该网络结构在 CWRU 滚动轴承数据集上实现了 100% 的测试精度。

随着工业领域技术的进步,以及设备的计算能力和存储能力的提高,工业中的传感器现在可以收集比以往更多更复杂的多元时间序列数据。在未来的工作中,针对多元的时间序列数据,如何将不同维度的信息进行融合或者如何对这些工业数据进行多进程的学习和识别,会有更多的想法和尝试。

参考文献

- [1] Kantz, H. and Schreiber, T. (2004) *Nonlinear Time Series Analysis*. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511755798>
- [2] Mao, W., Feng, W., Liang, X. and Zhang, X. (2019) A Novel Deep Output Kernel Learning Method for Bearing Fault Structural Diagnosis. *Mechanical Systems and Signal Processing*, **117**, 293-318. <https://doi.org/10.1016/j.ymssp.2018.07.034>
- [3] Yang, H., Mathew, J. and Ma, L. (2005) Fault Diagnosis of Rolling Element Bearings Using Basis Pursuit. *Mechanical Systems and Signal Processing*, **19**, 341-356. <https://doi.org/10.1016/j.ymssp.2004.03.008>
- [4] Lei, Y., Jia, F., Lin, J., Xing, S. and Ding, S.X. (2016) An Intelligent Fault Diagnosis Method Using Unsupervised Feature Learning towards Mechanical Big Data. *IEEE Transactions on Industrial Electronics*, **63**, 3137-3147. <https://doi.org/10.1109/TIE.2016.2519325>
- [5] Yin, H. and Gai, K. (2015) An Empirical Study on Preprocessing High-Dimensional Class-Imbalanced Data for Classification. 2015 *IEEE 17th International Conference on High Performance Computing and Communications*, 2015 *IEEE 7th International Symposium on Cyberspace Safety and Security*, and 2015 *IEEE 12th International Conference on Embedded Software and Systems*, New York, 24-26 August 2015, 1314-1319. <https://doi.org/10.1109/HPCC-CSS-ICSS.2015.205>
- [6] 张仁斌, 许辅昊, 刘飞, 等. 基于 K-均值聚类的工业异常数据检测[J]. 计算机应用研究, 2018, 35(7): 266-270.
- [7] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014) Generative Adversarial Nets. *Advances in Neural Information Processing Systems (NIPS 2014)*, 2672-2680.
- [8] Schlegl, T., Seeböck, P., Waldstein, S.M., Schmidt-Erfurth, U. and Langs, G. (2017) Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery. In: Niethammer, M., et al., Eds., *Information Processing in Medical Imaging. IPMI 2017. Lecture Notes in Computer Science*, Springer, Cham, 146-157. https://doi.org/10.1007/978-3-319-59050-9_12
- [9] Donahue, J., Krähenbühl, P. and Darrell, T. (2016) Adversarial Feature Learning. arXiv:1605.09782. <https://arxiv.org/abs/1605.09782>
- [10] Akcay, S., Atapour-Abarghouei, A. and Breckon, T.P. (2018) GANomaly: Semi-Supervised Anomaly Detection via Adversarial Training. arXiv:1805.06725. <https://arxiv.org/abs/1805.06725>
- [11] Liu, J., Qu, F., Hong, X. and Zhang, H. (2019) A Small-Sample Wind Turbine Fault Detection Method with Synthetic Fault Data Using Generative Adversarial Nets. *IEEE Transactions on Industrial Informatics*, **15**, 3877-3888. <https://doi.org/10.1109/TII.2018.2885365>
- [12] Han, T., Liu, C., Yang, W. and Jiang, D. (2019) A Novel Adversarial Learning Framework in Deep Convolutional Neural Network for Intelligent Diagnosis of Mechanical Faults. *Knowledge-Based Systems*, **165**, 474-487. <https://doi.org/10.1016/j.knosys.2018.12.019>
- [13] 邵俊杰, 董伟, 冯志. 基于机器学习的工业控制网络异常检测方法[J]. 信息技术与网络安全, 2019, 38(6): 17-20+25.
- [14] Wu, Z., Lin, W. and Ji, Y. (2018) An Integrated Ensemble Learning Model for Imbalanced Fault Diagnostics and Prognostics. *IEEE Access*, **6**, 8394-8402. <https://doi.org/10.1109/ACCESS.2018.2807121>
- [15] Japkowicz, N. (2000) The Class Imbalance Problem: Significance and Strategies. In: *Proceedings of the 2000 International Conference on Artificial Intelligence*, 111-117.

- [16] Galar, M., Fernandez, A., Barrenechea, E., Bustince, H. and Herrera, F. (2012) A Review on Ensembles for the Class Imbalance Problem: Bagging-, Boosting-, and Hybrid-Based Approaches. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, **42**, 463-484. <https://doi.org/10.1109/TSMCC.2011.2161285>
- [17] Nanni, L., Fantozzi, C. and Lazzarani, N. (2015) Coupling Different Methods for Overcoming the Class Imbalance Problem. *Neurocomputing*, **158**, 48-61. <https://doi.org/10.1016/j.neucom.2015.01.068>
- [18] Radford, A., Metz, L. and Chintala, S. (2015) Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. arXiv:1511.06434. <https://arxiv.org/abs/1511.06434>
- [19] Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A. and Chen, X. (2016) Improved Techniques for Training GANs. *30th Conference on Neural Information Processing Systems (NIPS 2016)*, Barcelona, Spain, 2234-2242.
- [20] Kingma, D.P. and Ba, J. (2014) Adam: A Method for Stochastic Optimization. arXiv:1412.6980. <https://arxiv.org/abs/1412.6980>