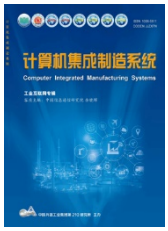


计算机集成制造系统
Computer Integrated Manufacturing Systems
ISSN 1006-5911,CN 11-5946/TP

《计算机集成制造系统》网络首发论文

题目：面向工业流程异常检测的均衡循环神经网络
作者：许荣斌，章宇，谢莹，刘志强，张以文，闻立杰
收稿日期：2022-08-23
网络首发日期：2023-03-23
引用格式：许荣斌，章宇，谢莹，刘志强，张以文，闻立杰．面向工业流程异常检测的均衡循环神经网络[J/OL]．计算机集成制造系统.
https://kns.cnki.net/kcms/detail/11.5946.tp.20230322.1713.012.html

网络首发时间：2023-03-23 15:50:05
网络首发地址：https://kns.cnki.net/kcms/detail/11.5946.tp.20230322.1713.012.html

面向工业流程异常检测的均衡循环神经网络

许荣斌[1]，章 宇[1]，谢 莹[1+]，刘志强[1]，张以文[2]，闻立杰[3]
(1. 莆田学院 机电与信息工程学院，福建 莆田 351100；2. 安徽大学 计算机科学与技术学院，安徽 合肥 230039；3.清华大学 软件学院，北京 100084)

摘 要：智能制造的迅速发展给网络安全防护带来巨大的机遇与挑战，各类安全威胁会造成严重的损失甚至灾难，已成为工业互联网亟待解决的问题。本文提出一种新的均衡循环神经网络，利用神经网络的适应性特点，采用长短期记忆网络（Long Short-Term Memory,LSTM）的门电路特性，针对工业互联网流数据随着时间推移异常检测准确性较低的问题，通过不同权重与当前输入数据重构得出遗忘门控、输入门控和输出门控。随后通过 sigmoid 激活函数求得预测结果，并将该结果作为门控循环单元网络（Gated Recurrent Unit, GRU）的网络层输入，由 GRU 网络层促使当前网络快速拟合，从而较快地获得较优的参数。本方法结合 LSTM 和 GRU 的优势，保留 LSTM 最后时刻的隐藏状态，作为下一层网络 GRU 的输入，使网络层的连接更加平滑，最大程度地保留 LSTM 所学习到的参数，获取隐藏特征，既可提高神经网络的精度，又可高效、快速地检测工业互联网络的异常。

关键词：循环神经网络；长短期记忆；门控循环单元；工业互联网；异常检测
中图分类号：TP311；　　**文献标识码**：A

Symmetric recurrent neural network for anomaly detection in industrial process

XU Rongbin[1]，ZHANG Yu[1]，XIE Ying[1+]，LIU Zhiqiang[1]，ZHANG Yiwen[2],WEN Lijie[3]
(1. School of Mechanical, Electrical and Information Engineering, Putian University, Putian 351100, China；2.School of Computer Science and Technology, Anhui University, Hefei, 230039, China；3.School of Software, Tsinghua University, Beijing 100084, China)

Abstract：The rapid development of intelligent manufacturing brings great opportunities and challenges to security protection. Various kinds of security threats may cause serious losses or even disasters, which have become an urgent problem to be solved in the Industrial Internet. In this paper, a novelsymmetric recurrent neural network is proposed, which utilizes the adaptability of neural network and the characteristic of gate circuit inLong Short-Term Memory (LSTM) network. This can solve the problem of low accuracy inanomaly detection for Industrial Internet streaming data over time. First, ForgetGate, Memory Gate and Output Gate are calculated by different weights and current input data.Then the prediction results are solved by sigmoidactivation function. And the results are used as the input of Gated Recurrent Unit (GRU) network layer, which promotes the rapid fitting of the current network. So that thebetterparameters could be obtained in a short time. The advantages of this method in combination with LSTM and GRU help to keep the last

收稿日期：2022-08-23；修订日期：2023-02-21。Received 23 Aug.2022;accepted 21 Feb. 2023.
基金项目：国家自然科学基金资助项目（62276146）；教育部人文社科交叉基金资助项目(20YJCZH197)；福建省自然科学基金资助项目(2020J01923,2021J011111)；莆田市科技计划资助项目(2020GP003).**Foundation items**：Project supported by the National Natural Science Foundation,China(No.62276146),the MOE Youth Project of Humanities and Social Sciences Foundation, China(No.20YJCZH197), the Natural Science Foundation of Fujian Province, China(No.2020J01923,2021J011111), and the Putian Technology Planning, China(No.2020GP003).

hidden stateof LSTM, which can be taken as the input of next layer for GRU.This connection keeps the neural network more smooth and maximum retention the parameters of LSTM. Finally, this novel method greatly improves the accuracy of neural network, which can both efficiently and quickly detect theanomaliesin Industrial Internet.

Keywords: recurrent neural network; long short-term memory; gated recurrent unit; industrial internet; anomaly detection

0 引言

随着计算机技术与工业互联网的迅速发展，国务院发布实施《中国制造 2025》规划，此规划将智能制造作为工业互联网研究的主攻方向，抢占制造业新一轮竞争制高点[1,2]。智能制造产业的快速发展将加快装备向智能化升级，高档数控机床、工业机器人、智能传感器和控制装备等新兴智能制造装备的研发与产业化速度进一步加快，智能装备制造企业、系统集成和设备服务企业进一步集聚协调发展[3]。从德国的工业 4.0、美国先进制造与工业互联网，以及中国制造 2025 规划，世界各国都在争相发力智能制造，这给制造业带来了巨大的市场和机遇[4]。

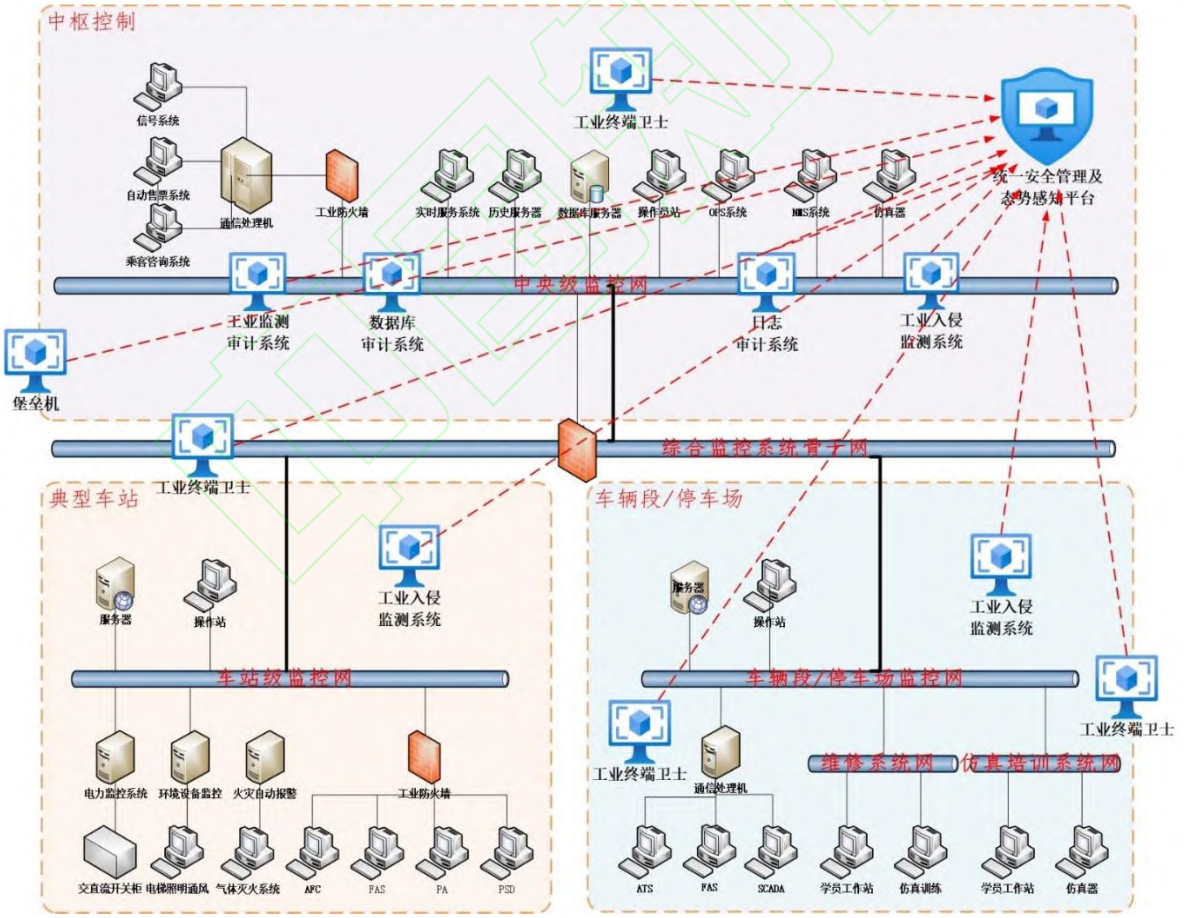


图 1 轨道交通系统协同工作流程

工业互联网的变革打破了传统工控系统（Industrial Control System, ICS）的封闭环境，融合了云计算、边缘计算、人工智能、大数据、物联网等技术，将生产制造环节与互联网信息系统连接起来，实现资源整合共

享、生产智能化与自动化，从而达到降低运营成本、缩短研制周期、提高生产效率等目标[5]。工业互联网包含网络、平台、数据、安全四大体系，既是工业数字化、网络化、智能化转型的基础设施，也是互联网、大数据、人工智能与实体经济深度融合的应用模式，同时也是一种新业态、新产业，将重塑企业形态、供应链和产业链。涉及电子设备生产、轨道交通、装备制造、纺织、钢铁、采矿、电力等诸多国家民生行业与实体经济重点产业，包含大量的业务系统和流程，实现工业互联网中安全、网络、IT 移动、工控等各种类型设备对接和控制[4,6]。图 1 为城市轨道交通系统协同工作流程，其中工业设备支持远程操控，设备间协同作业流程和调度逻辑清晰，通过工业互联网络与多接入边缘计算系统结合，部署轨道交通调度应用，满足轨道交通运行过程中对实时控制、数据集成与互操作、安全与隐私保护等方面的关键需求。

工业互联网安全防护研究工作已经成为世界研究热点，是全球制造业的主要发展方向和战略制高点[7]。当前，智能制造环境下的工控系统面临更多新的安全挑战，智能制造控制设备的高危漏洞、进口设备后门、高级持续性威胁、工业网络病毒以及无线技术应用等安全风险，会带来生产核心数据泄露、生产系统瘫痪等安全威胁，造成财产甚至生命的损失[8,9]。制造业成为全球攻击者的重要攻击目标，为智能制造业提供可靠的安全防护成为亟待解决的问题。攻击者的行为记录往往分散在不同类型的日志中，将会针对工业互联网众多业务流程进行漏洞探测、暴力破解等攻击，其异常模式占据较长的时间窗口[10]。因此需要研究优秀的计算模型[11]，来支持相关的安全分析功能，并利用机器学习方法进行网络入侵检测，从而在整体上把握工业互联网络空间安全状态。

针对上述问题，本文在已有的相关工作基础上对深度神经网络的激活函数与门控机制进行深入研究，并设计丰富的实验评估。主要贡献包括：1）改进传统的 LSTM 神经网络，更好地挖掘数据相关特征，缓解传统循环神经网络梯度饱和、梯度消失问题，提高运行效率；2）提出一种新的重置门控结构，在门控循环单元网络内部附加 sigmoid 激活函数，使得循环神经网络更加平滑；3）提出一种将 LSTM 与 GRU 结构相结合的新型神经网络，采用较为平缓的连接方式，有效解决单层网络运行效率低且精度不足的问题，加速识别工业互联网中的异常信号，保证工控网络的正常运行。

本文的章节安排有：第 1 节是典型案例，通过实际样例来阐述本文要解决的主要问题；第 2 节介绍相关循环神经网络，主要阐述本文涉及的循环神经网络概念和理论；第 3 节提出一种新的均衡循环神经网络，对 LSTM 和 GRU 分别做出改进，并结合二者设计出性能更高的神经网络模型；第 4 节介绍了测试和和训练数据，对比传统方法与新的均衡循环神经网络进行实验评估，并对各种方法的结果进行分析与比较；第 5 节对本文进行总结，给出未来的研究方向。

1 典型案例

1.1 工业互联网案例分析

工业控制系统很多是应用于国家关键领域的基础设施，一旦遭受网络攻击，会造成较为严重的损失。智能制造控制设备的高危漏洞、进口设备后门、高级持续性威胁、工业网络病毒以及无线技术应用等安全风险，会带来生产核心数据泄露、生产系统瘫痪等安全威胁，造成巨大财产甚至生命的损失。2018 年 8 月台积电工厂遭勒索病毒攻击，停摆 3 天损失高达 1.7 亿美元；2021 年 7 月由于勒索软件攻击导致服务器离线，英国北方铁路耗资 1700 万英镑采购的自助售票系统陷入瘫痪，超过 420 个车站受影响；2022 年 3 月黑客通过暴力破解入侵美国征信巨头 TransUnion 一台存有大量消费者数据的 SFTP 服务器，导致 TransUnion 将为受影响的消费者免费提供身份保护年度订阅服务，预计成本将超过 114 亿元。

图 1 所示的城市轨道交通系统协同工作流程包含 3 层业务流程，每层都包含有一个或多个子系统，由若干设备连接而组成业务子流程，流程与流程之间再通过交换机或无线连接，通信设备与网或外网直接连接。开放的端口、未修复的漏洞、未认证的接口等问题都会成为黑客便捷入侵的攻击点。加大关键信息基础设施防护力度，对工业互联网各业务系统和流程实行统一动态监测，确保工业互联网安全平稳可靠运行，提升整体安全认知能力，对我国工业互联网的发展有着至关重要的意义。

1.2 问题分析

工业互联网业务流程异常动态检测逐渐在智能制造、电力、能源、轨道交通等行业得到很多应用。工控业务流程异常检测综合了安全防护多方面技术，从建立起来到良好运行，需要结合具体的使用环境和行业特点，建立不同场景的算法模型。近些年随着大数据、云计算、边缘计算和人工智能等技术的发展，逐渐在工业互联网安全领域深入应用，工控业务异常检测的发展趋势是：

（1）深度融合大数据和人工智能技术。通过在融合使用深度学习、知识图谱等大数据分析算法和人工智能模型，从整体上把握网络空间安全状态，对针对关键信息基础设施和重要信息系统的网络攻击和重大网络安全威胁实现可知、可管、可控、可溯、可预警，及时发现并精确预警及处置。

（2）系统可以动态扩展和云化。随着云计算基础设施的大量使用，要求对安全威胁和攻击的处置能力也是可以随着云计算平台扩展而可动态扩展的，实现网络安全态势感知系统的基础平台云化，使其态势感知能力可以随着保护对象的规模变化而动态变化。

（3）提供精准预测和防御处置。在大数据挖掘与分析技术持续发展的基础上对网络安全态势进行深度感知和整体把握，对网络空间的整体安全态势发展给出更精准的预测和积极防御处置建议。

2 循环神经网络

2.1 循环神经网络基本结构

循环神经网络（Recurrent Neural Network, RNN）是一类以序列数据为输入，在序列的演进方向进行递归，且所有节点按链式连接的递归神经网络[12]。循环神经网络的显著特点是具有记忆性，可以处理与时间有关的数据，也能处理与顺序有关的数据。

循环神经网络 RNN 具备环状箭头，用以表示网络循环，展开后如图 2 所示。t 时刻的输出 O_t 不仅受到 t 时刻的输入数据影响，还受到上一时刻 t-1 输入数据的影响，而 t-1 时刻又受到 t-2 时刻的影响，以此类推，可以把 t-n 时刻的输入数据的影响都加在当前的时刻中。通过此结构，序列数据之间的信息尽可能地包含在网络中，也就把语义信息中前面表达的意思传递到网络后部。因此 RNN 广泛应用于序列数据的处理，可以根据数据来理解具体信息，处理语言、文字、信息等数据的分析与预测。

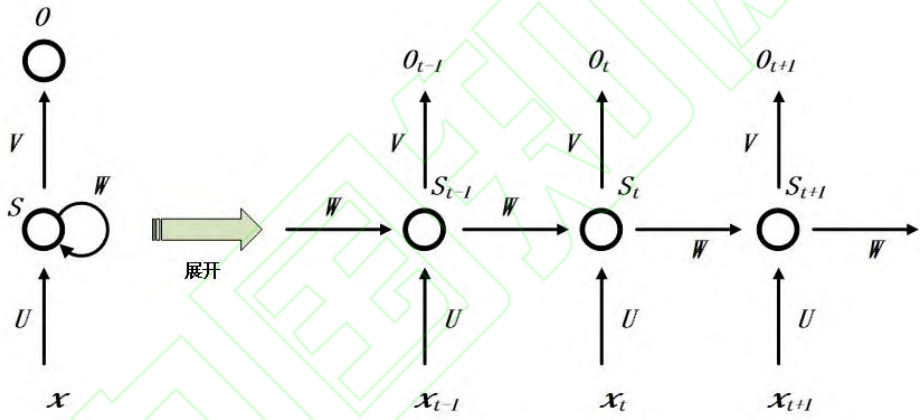


图 2 循环神经网络基本结构图

工业互联网中的通信数据正是与时间有关且按一定顺序组建起来，如果数据里的词的顺序或者频率被篡改，则不能表达出正确的指令。工业互联网通信数据在不同的时间点，数据会有所不同，数据的大小随时间的变化而变化，序列数据的每一个数据点与其左右的数据都有关联，数据不是孤立存在的。因此，采用循环神经网络不仅能获取到自身神经元的信息，也可以获取到其他神经元所传递的信息。这比原有的前馈神经网络能学习到更多时序信息，更加符合生物神经网络的结构。

2.2 门控循环神经网络

如图 2 所示，RNN 中的循环重复结构实现了参数共享，解决了历史数据的利用问题，很适合应用于工业互联网中的序列数据。然而传统 RNN 在其梯度下降过程中，由于链式法则出现累乘项，当t很大时，累乘项的值趋向于 0。因此 RNN 容易出现梯度消失现象，使得参数更新缓慢，甚至是停止更新。

研究者在循环神经网络的基础上衍生设计出了一些优化方法，最为经典的优化网络是通过加入门控机制

（Gating Mechanism）进行优化，也就是基于门控的循环神经网络（Gated RNN）[13]。加入门控机制的网络需要选择是否遗忘曾经积累数据的遗忘门控，以及是否需要保留现有数据的记忆门控，使循环神经网络的学习能力增强，精度提高，很大程度上解决了梯度爆炸、消失的问题。Gated RNN 中较为出名的网络是影响深度学习多年的长短期记忆网络（Long Short-Term Memory Network,LSTM）[14]，门控循环单元网络（Gated Recurrent Unit,GRU）[15]则是近些年提出的较为新颖的方法，两者各有其优缺点，本质都是门控机制。

3 均衡循环神经网络

3.1 快速长短期记忆神经网络

长短期记忆神经网络是为了改善神经网络对过去数据的长期依赖问题，在加入门控机制的基础上，选择遗忘之前积累的部分信息，并且有选择地加入当前数据组成新的信息，从而得到较高精度的预测结果[16]。LSTM 神经网络结构如图 3 所示。

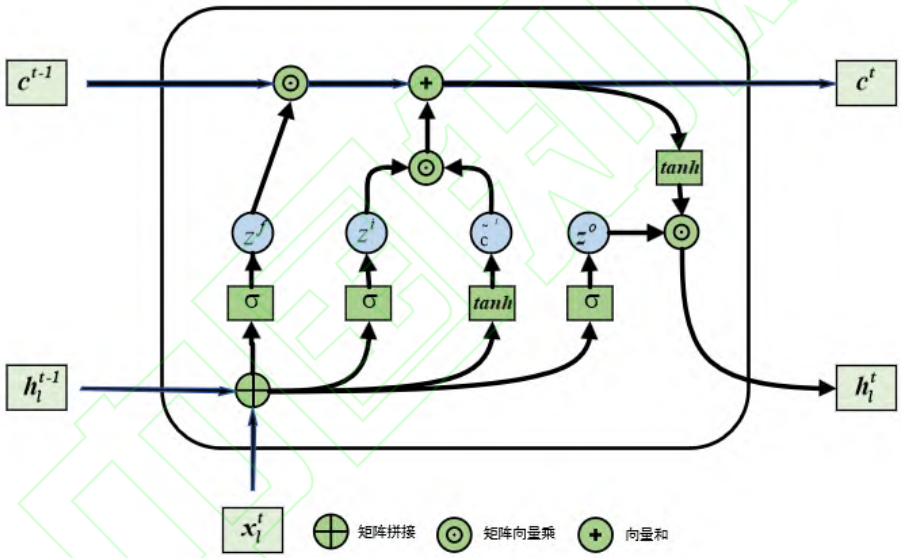


图 3 长短期记忆神经网络结构

（1）首先将当前输入与上一时刻隐藏状态结合得到的候选状态:

$$\tilde{c}^t = tanh(W^c * x_l^t + U^c * h_l^{t-1} + b^c) \quad (1)$$

式中x_l^t表示当前输入，h_l^{t-1}表示上一时刻的隐藏状态（起始状态时将h_l^{t-1}设为 0），$\tilde{c}^t$表示当前候选状态。W^c和U^c分别为x_l^t和h_l^{t-1}的权重，b^c为候选状态的偏移量，该激活函数一般使用 tanh 函数。

（2）通过不同权重与当前输入计算得出遗忘门控、输入门控和输出门控：

$$z^f = relu\left(W^f * x_l^t + U^f * h_l^{t-1} + b^f\right) \quad (2)$$

$$z^i = relu\left(W^i * x_l^t + U^i * h_l^{t-1} + b^i\right) \quad (3)$$

$$z^o = relu(W^o * x_l^t + U^o * h_l^{t-1} + b^o) \quad (4)$$

式中z^f、z^i、z^o分别为遗忘门控、输入门控和输出门控。b^f、b^i、b^o为各自的偏移量，该激活函数一般使用 Logistic sigmoid 函数。然而，传统的 LSTM 方法拟合效率相对较低，为了提高循环神经网络拟合效率，本文采用 Relu 作为神经激活函数，利用其所具有的单侧抑制特点，更好地挖掘相关特征，在一定程度上解决梯度饱和、梯度消失的问题。

（3）通过公式 2 和 3，计算出遗忘门控和输入门控，并分别与上一时刻的内部状态及公式 1 求出候选状态，通过矩阵相乘计算出新的内部状态：

$$c^t = z^f \odot c^{t-1} + z^i \odot \widetilde{c^t} \tag{5}$$

此处c^{t-1}为上一时刻的内部状态（起始状态时将c^{t-1}设为 0），c^t为新的内部状态，$\odot$为矩阵运算。

（4）通过公式 4 得到的输出门控与公式 5 求出的内部状态矩阵相乘，求出当前的隐藏状态：

$$h_l^t = z^o \odot tanh(c^t) \tag{6}$$

式中h_l^t为当前时刻的隐藏状态，本激活函数一般使用 Tanh 函数。

（5）通过所得出的隐藏状态计算出结果：

$$y_l^t = sigmoid\left(W_l^y * h_l^t\right) \tag{7}$$

利用上一时刻的隐藏状态（h_l^{t-1}）以及上一时刻的内部状态（c^{t-1}），通过当前输入得到当前时刻的隐藏状态、内部状态以及当前的输出，从而不断学习优化权重和偏移量，以得到较优的预测率，该激活函数一般使用 Logistic sigmoid 函数。

3.2 门控循环单元网络

门控循环单元网络 GRU 可以在减少问题求解计算量同时，获得较好的计算精度。GRU 的核心是拥有两个门控，重置门控与更新门控，这两个门控能很好防止梯度爆炸或梯度消失。尤其对于大数据量的计算问题，使用 GRU 能够高效、快速地实现网络拟合。

（1）通过不同的权重计算得到重置门控和更新门控：

$$z^r = sigmoid\left(W^r * x_g^t + U^r * h_g^{t-1} + b^r\right) \tag{8}$$

$$z^z = sigmoid\left(W^z * x_g^t + U^z * h_g^{t-1} + b^z\right) \tag{9}$$

式中z^r、z^z分别为重置门控和更新门控，其中x_g^t为当前输入，h_g^{t-1}为上一时的隐藏状态，W^r、W^z、U^r、U^z分别为重置门控和更新门控各自输入及隐藏状态的权重，b^r、b^z为各自的偏移量。该激活函数一般使用 Logistic sigmoid 函数。

（2）通过当前时刻的输入，上一时刻的隐藏状态以及公式 8 求出的重置门控，求出当前的候选状态：

$$h_g = sigmoid\left(W^h * x_g^t + U^h * h_g^{t-1} + b^h\right) \quad (10)$$

$$\tilde{h}^t = \tanh\left(h_g * z^r\right) \quad (11)$$

式中h_g为当前的候选预备状态，W^h为当前输入的权重，U^h为上一时刻隐藏状态的权重，b^h为当前的偏移量，$\tilde{h}^t$为当前的候选状态。为了使得 GRU 函数能更加平滑获取到数据信息，本文通过先通过 sigmoid 激活函数取得当前候选预备状态，在与重置门控相乘求得当前的候选状态。

（3）当前时刻的隐藏状态可通过下式计算得出：

$$h_g^t = z^z \odot h_{t-1} + (1 - z^z) \odot \tilde{h}^t \quad (12)$$

其中h_g^t为当前时刻的隐藏状态，更新门控由公式 9 式求出。z^z的范围为 0-1，用z^z和 1-z^z分别乘以上一时刻的隐藏状态和当前候选状态。若z^z数据较大，表示会较多的保留上一时刻的隐藏状态信息，而选择放弃多数的当前候选状态信息；若z^z数据较小，则结果相反。

该组公式仅通过两个门控即可实现神经网络的拟合，由于其网络相对简化，加快了问题的求解过程，从而使其能达到较好的效果。

GRU 神经网络的结构如图 4 所示。

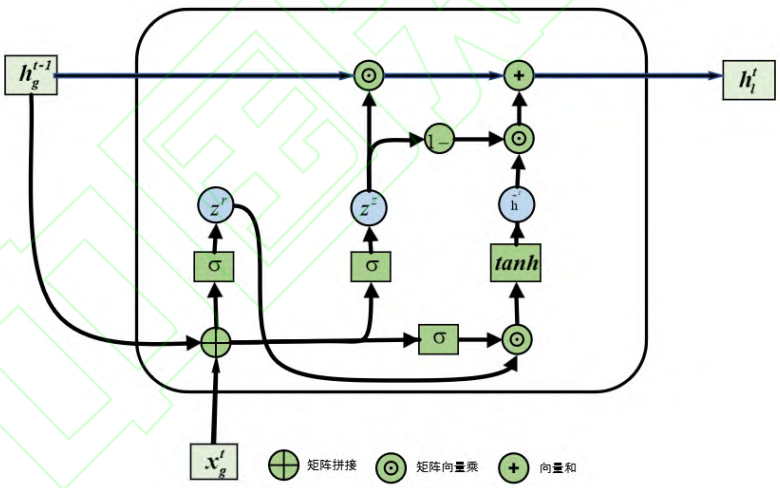


图 4 GRU 门控循环单元网络结构

（4）求出最后结果：

$$y_g^t = h^t \odot W_g^y + b^y \quad (13)$$

其中y_g^t为当前输出，W_g^y与b^y为权重与偏移。

3.3 均衡 LSTM-GRU 循环神经网络

通过前述两个神经网络结构的改进，可以发现 LSTM 由于其网络的复杂性，能学习到的参数较多，获得的精度较高。但是由于 LSTM 的复杂网络导致计算量较大，需要通过多次学习才能得到较高的精度。GRU 则

由于其门控的减少，较大幅度地降低了计算量，使得该网络能很快地达到较好的效果。然而由于 GRU 门控较少，会导致其在精度上始终不够理想。

基于上述 LSTM 与 GRU 的基本理论和改进，本文提出一种新的神经网络结构——均衡 LSTM-GRU 神经网络（SymmetricLSTM-GRU, SYM-LSTM-GRU），在保留各自优势的同时，解决两个网络所存在的问题。本方法能在较短的时间内达到比 LSTM 精度更高的效果，又能在很大程度上减少时间开销。均衡 LSTM-GRU 神经网络通过对输入数据赋予不同权重，计算出遗忘、输入和输出门控，交叉使用 sigmod 与 tanh 激活函数求解预测结果，将改进的 LSTM 网络层计算得出的结果作为 GRU 网络层的输入，促使 GRU 网络层快速拟合。新方法既可保证问题求解的高精度，又可实现高效、快速的网络拟合，能够均衡地满足不同问题的需求。

均衡 LSTM-GRU 循环神经网络的具体实现如下：

（1）由公式(1)-(7)计算得出的结果作为 GRU 网络的输入。由于 LSTM 网络较为复杂，为了防止过拟合，使用 dropout 层随机丢弃掉一些数据；同时，为了保持一致性，也丢弃掉部分隐藏状态的数据。

$$x_g^t = dropout(y_l^t). \tag{14}$$

$$h_l^t = dropout(h_l^t) \tag{15}$$

式中y_l^t表示 LSTM 层的输出，x_g^t表示 GRU 层的输入，h_l^t表示 LSTM 层的隐藏状态。

（2）为了最大程度保证网络的连续性，同时考虑 GRU 神经网络内部同样存在的隐藏状态，本文提出一种新的数据连接方法，提取出 LSTM 神经网络训练得到的最后一个时刻的隐藏状态，用于 GRU 神经网络初始的隐藏状态，而不是全部初始化为全零作为隐藏状态。由于 LSTM 拥有更多的门控，因此，该隐藏状态中会包含更多的有效信息。然而，如果使用全零作为 GRU 隐藏状态的初始化，可能需要花费更多的时间来重新学习有效信息，导致时间和空间开销增大，并且很有可能没有单网络的效果好。于是选择将 LSTM 最后时刻的隐藏状态作为 GRU 的隐藏状态，能保证高效、快速的同时，有效提高网络的学习效率，达到更高的精度。

$$h_g^t = h_l^t \tag{16}$$

（3）将从 LSTM 得到的隐藏状态和输出数据作为 GRU 神经网络的输入，通过公式(8)-(13)计算得到输出结果。均衡 LSTM-GRU 神经网络的具体架构如图 5 所示。

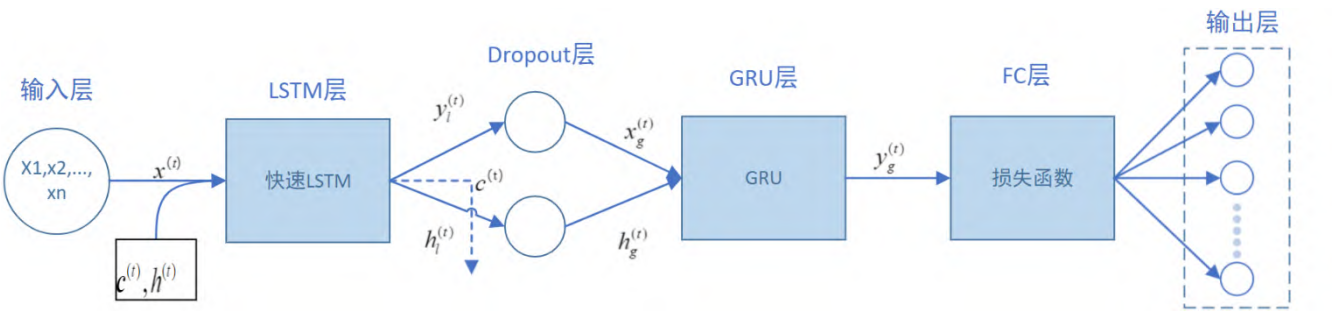


图 5 均衡 LSTM-GRU 结构

（4）通过各类损失函数，计算出均衡 LSTM-GRU 的求解精度。

均衡 LSTM-GRU 的伪代码如算法 1 所示。

算法 1: 均衡$LSTM-GRU$神经网络

输入：数据集$\hat{X}$: $x^{(t)}$
隐藏状态和内部状态$c^{(t)}$，$h^{(t)}$
输出：预测结果集$\hat{Y}$:$y^{(t)}$

```
nn.Sequential:
c^(t)，h^(t)=[0,0,...,0]
y^(t)，c^(t)，h^(t) =nn.LSTM(x^(t)，c^(t)，h^(t))
x^(t) = nn.dropout(y^(t))
h^(t) = nn.dropout(h^(t))
    y^(t)，h^(t) =nn.GRU(x^(t)，h^(t))
        Return y^(t)
```

4 实验评估

本节将给出均衡 LSTM-GRU 神经网络的实验评估。对不同类型的工业互联网数据进行评估分类，从网络中获取到各类信息并进行标签化处理，提供给均衡 LSTM-GRU 神经网络进行学习。

4.1 数据处理与分析

本实验中的数据集来源于 Kaggle[1]，由 20000 条数据组成，包含一系列长短不同的字符串。由于神经网络结构不能直接对字符串数据进行检测，需要将其转换成数值数据。在将字符串数值化过程中，采用字典映射法将原始数据集中的分词映射成数字，有效地解决字符串数据无法被神经网络结构检测的问题。由于工业互联网络数据输入是固定的，而每个日志的长度并非是相等的，这会导致网络无法正常使用。考虑到该因素，在字符数值化基础上，本文采用嵌入（Embedding）方法将原有特征映射到新的特征空间，使每个特征的输入格式相同，以保证接下来的实验输入格式相同。

工业互联网异常数据集主要包含的标签有 DoS 攻击、恶意爬取、数据探测、漏洞探测、暴力破解和正常

[1] https://www.kaggle.com/datasets/praveengovi/emotions-dataset-for-nlp?datasetId=605165&sortBy=voteCount

信息。在监控网络安全的过程中，一旦有信息泄露、恶意攻击等一系列事件发生，对智能制造行业来说，会造成巨大的损失。因此，快速准确地对异常情况进行分类非常重要。

工业互联网数据集标签的设置如表 1 所示，类别 1 表示正常状态，其他数值均表示异常状态。其中，最重要的就是异常状态的识别问题，即检测通信数据是否为异常状态。若检测出异常信息，则需要将异常信息分析结果返回给系统做出相对应的异常处理，并将异常信息增加到异常日志中。

表 1 工业互联网通信数据基本分类

标签	正常信息	DoS 攻击	恶意爬取	数据探测	漏洞探测	暴力破解
序号	1	2	3	4	5	6

4.2 参数设置

神经网络的性能受到超参数影响很大，尤其是门控循环神经网络参数更多。因此神经网络结构的设计对工业网络安全的检测非常重要，同时如何选取合适的超参数对于高效的神经网络也有着非常大的影响。对于本工作主要有以下几方面参数设置。

（1）隐藏层的层数：由于工业互联网数据集并不是非常复杂，只需将隐藏层数设置为 1 至 3 层。网络层数越多，神经网络的能力就越强，但也可能导致网络发生过拟合。因此，并不考虑过多的神经网络层数。

（2）学习率：该参数是循环神经网络中较为重要的超参数，本实验分别选择 0.01、0.005、0.001 和 0.0005 四个有代表性的学习率进行学习。

（3）隐藏层的神经元个数：由于嵌入（Embedding）层将输入数据映射成长度为 128 的向量，为了能精确的计算训练数据量，本实验将其神经元数量设置为 256。

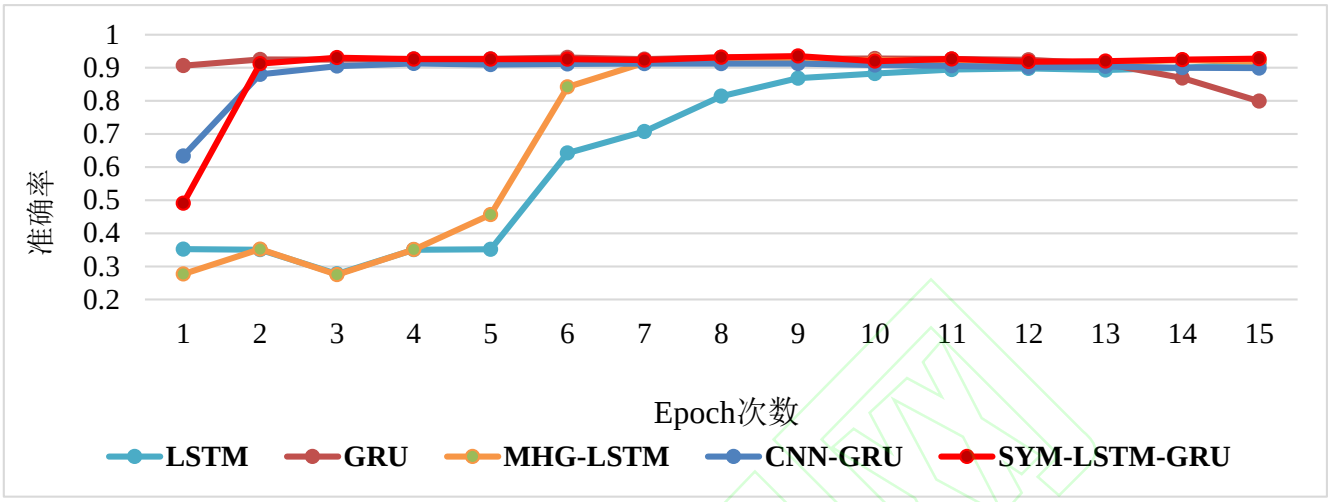
以上超参数对网络性能都有交叉影响，不能单独进行考虑，并且与迭代次数密切相关。本实验在尽可能保证较高精度的基础上，尽量减少训练时间，防止发生过拟合，并选择合适的迭代次数。通过上述超参数的设置，寻找具有较高性能的神经网络模型，可快速、准确地做出工业流程异常检测与识别。

4.3 基于均衡循环神经网络的实验评估

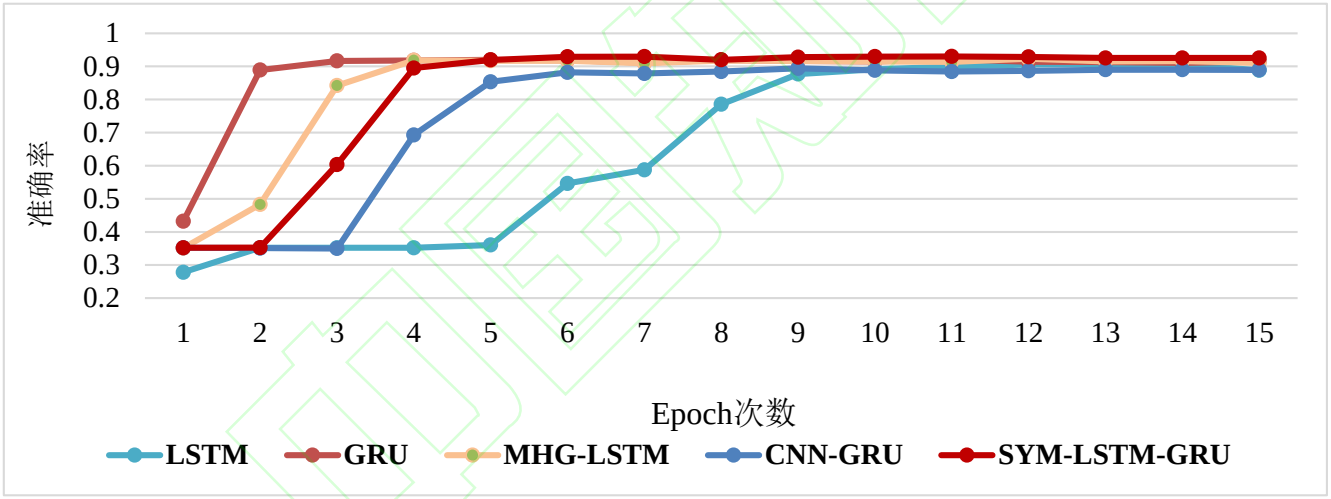
均衡 LSTM-GRU 的网络层比普通循环神经网络更加复杂，对超参数的影响更大。在多分类问题中，即使一个极小的改动，也可能会导致对异常分类的结果产生较大影响。不同参数的组合有多种形式，需要从中选取较为优化的组合来获取较优的结果。

在神经网络模型诸多的评价指标中，准确率是非常有说服力的指标。因此首先将准确率作为评价指标来进行实验分析。为了寻找到较优的参数组合，分别采用 4.2 节中所设置的参数组合分别对各网络进行训练，从中选取 1 至 3 层网络，将准确率作为评判标准以训练出精度效果较高的方法。通过对比传统方法和改进循

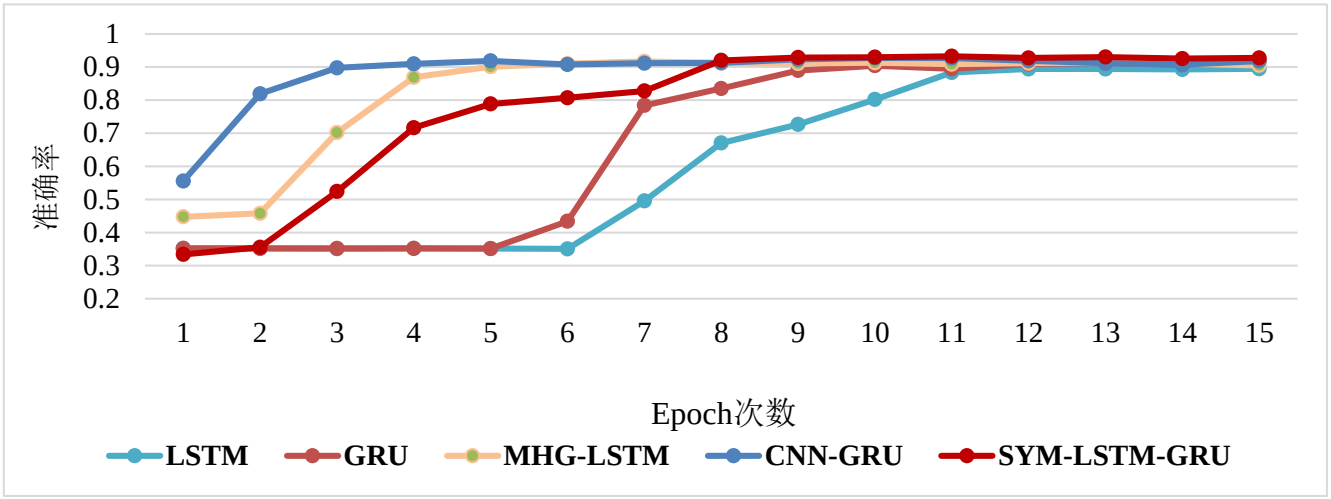
环神经网络方法[17, 18]，如图 6 所示，可以看出与 LSTM 网络有关的方法拟合速度都比较慢。本文提出的 SYM-LSTM-GRU 网络结合了 LSTM 与 GRU 网络的优势，相比于其他方法，SYM-LSTM-GRU 在准确率和拟合速度上均有大幅提升。



(a) 1-Layer 训练精度

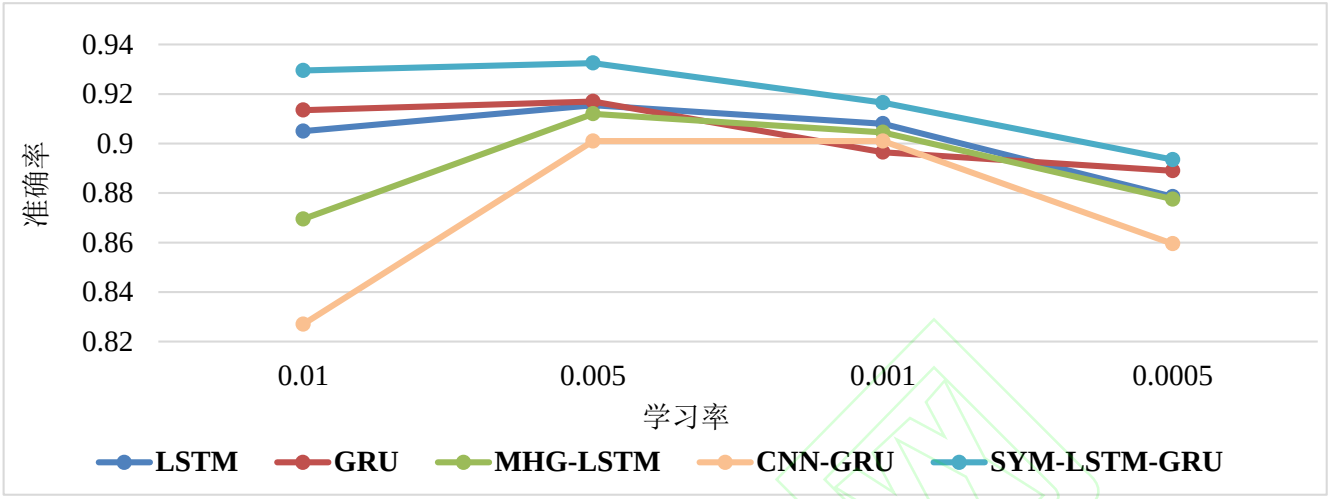


(b) 2-Layer 训练精度

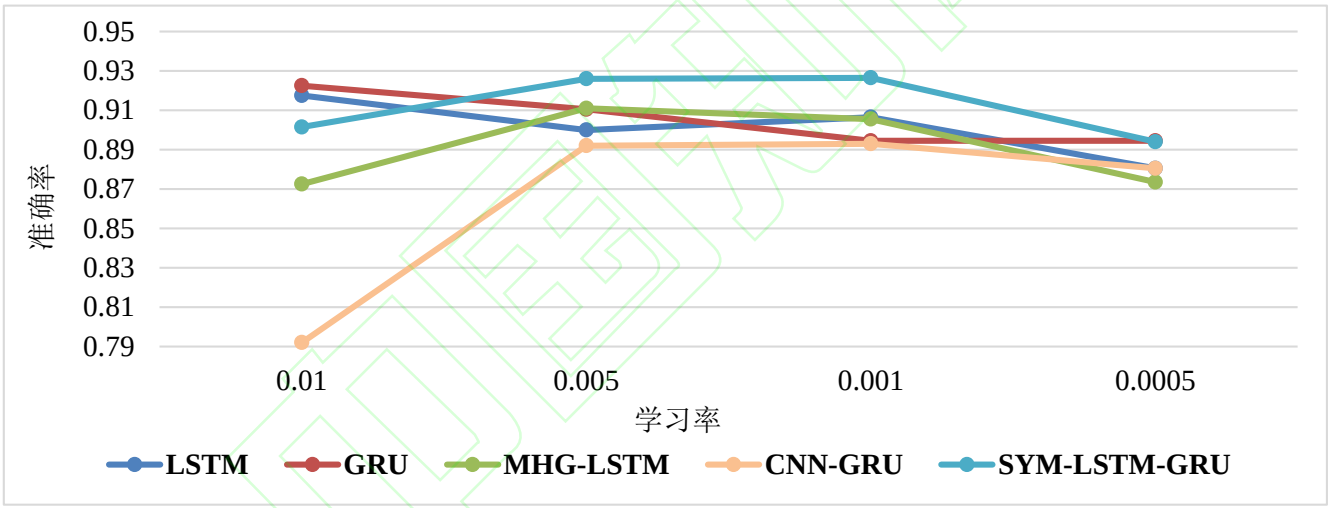


(c) 3-Layer 训练精度

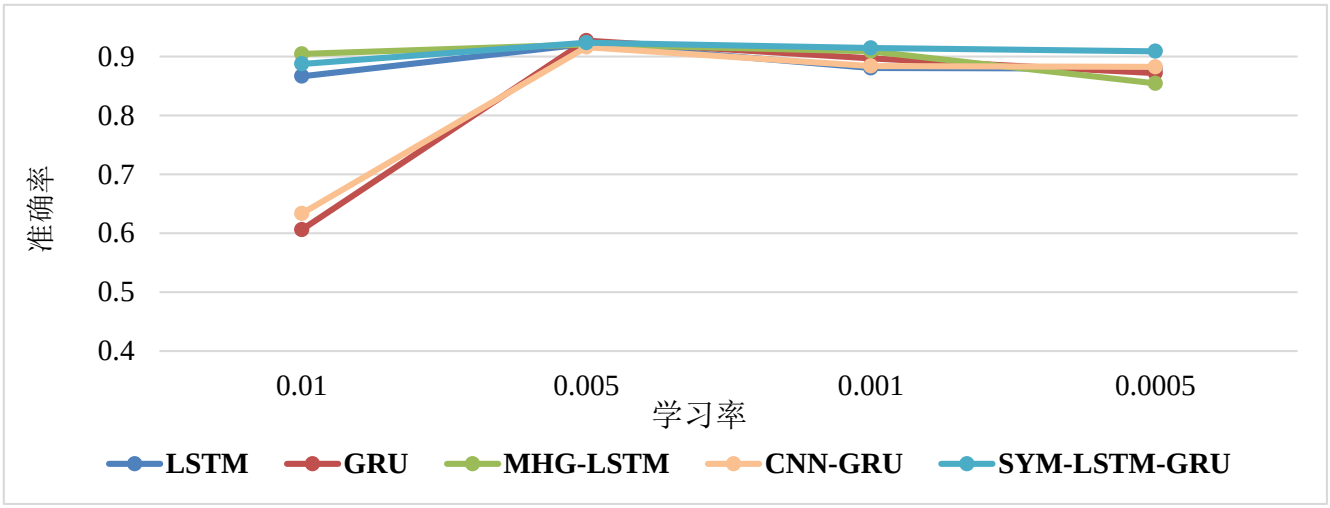
图 6 不同层数神经网络的训练精度



(a) 1-Layer 神经网络



(b) 2-Layer 神经网络



(C) 3-Layer 神经网络

图 7 不同层数神经网络的测试精度

其次，在各类超参数组合的情况下，选取 1 至 3 层网络，通过学习率与 4 种方法进行对比。如图 7 所示，与 GRU 网络有关的方法在不同学习率与网络层数所获得的准确率结果变化较大，稳定性不高。本文所提出的 SYM-LSTM-GRU 神经网络相比其他方法，仅在学习率为 0.01 时与其他方法所获得的准确率有高有低，其他情况下均优于其他方法且比较稳定，都获得了较高的准确率。

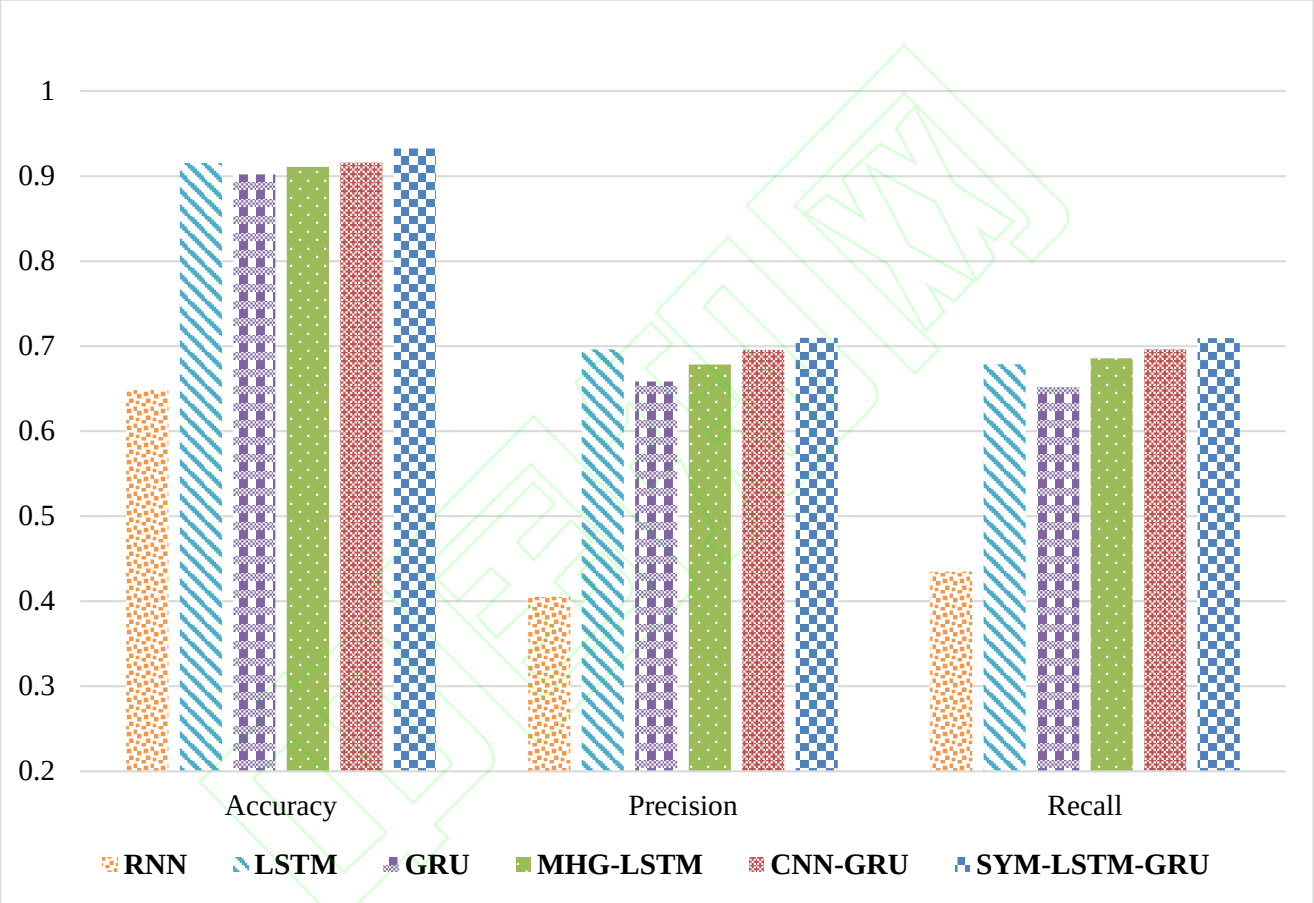


图 8 不同方法异常检测指标的平均值

再者，为了验证 SYM-LSTM-GRU 方法的有效性，采取了三种精度测试方法来进行实验对比，包括准确率（Accuracy）、查全率（Recall）和精确率（Precision）。如图 8 所示，SYM-LSTM-GRU 在所有指标上均优于其他模型。其中，高准确率说明新神经网络结构在处理较为平衡的数据集上有较好的表现。针对网络数据异常检测，查全率则更具有较高的参考意义。因为如果意外错失一些关键的异常信息，很可能会导致出现严重的后果。

采用 4.2 节中的参数设置，学习率为 0.005，隐藏层数目为 2，神经元个数为 256。如图 8 所示，SYM-LSTM-GRU 比传统的 RNN 方法在准确率上高出 25%以上，比其他改进循环神经网络方法在准确率上也提高了 2%到 5%，可以验证本文对循环神经网络内部的改进能有效地提高网络的准确率。除了准确率，查全率和

精确率也十分重要。通过对比图 8 中各方法的查全率和准确率，可以明显发现 RNN 性能相对较差。本文所提出的 SYM-LSTM-GRU 结合了 LSTM 和 GRU 优势，并改进了二者的结构，使 SYM-LSTM-GRU 网络更为有效和稳定。实验证明了 SYM-LSTM-GRU 可以高效、快速地适应工业互联网中各业务流程的异常检测，体现出本方法所具备的均衡能力。因此，通过十次实验结果计算出的平均值，在上述三种常用的多分类指标比较结果中，验证了本文提出的新神经网络结构的高效性。

最后，为了进一步总结不同方法的训练过程与获得的精度，本文给出了各循环神经网络的优缺点。如表 2 所示，RNN 是一种简单快速的循环神经网络，由于其没有门控机制进行梯度删减，容易造成梯度爆炸的问题，导致其性能较低。而使用带有门控的神经网络，性能都可以达到较高的水平。LSTM 网络的门控较多，使其具有记忆强、精度高等特性。然而，由于 LSTM 的复杂性，导致其需要花费较多的迭代才能达到拟合效果，所以 LSTM 网络运行相对较慢。GRU 网络减少了门控数量，提高了收敛和运行速度，却降低了网络的精度，是精度最低的网络结构。MHG-LSTM 和 CNN-GRU 分别对 LSTM 和 GRU 做出了改进，提高了原有方法的准确率。本文提出 SYM-LSTM-GRU 则同时解决了传统循环神经网络的缺陷，构成一种拟合速度快、精度高的新型网络结构。相比于 MHG-LSTM 和 CNN-GRU，SYM-LSTM-GRU 在获得较高准确率的同时，大幅提高了循环神经网络的性能，能非常好的应用于工业互联网中各业务流程的异常检测与识别。

表 2 循环神经网络算法比较

网络	特征	复杂度	性能	计算时间
RNN	梯度爆炸、消失	正常	较低	快
LSTM	门控多，记忆强	较复杂	较高	较慢
GRU	门控少，收敛快	一般复杂	较高	快
MHG-LSTM	门控多，记忆强	较复杂	较高	正常
CNN-GRU	门控少，收敛较快	一般复杂	一般	正常
SYM-LSTM-GRU	有效记忆	较复杂	非常高	正常

5 结束语

工业互联网各业务流程通信数据是一种流数据，包含着大量对工业设备实时控制、远程操作与智能调度等信息，当前智能制造快速发展给工业互联网安全防护带来了巨大的挑战。本文提出一种新的均衡循环神经网络 SYM-LSTM-GRU，结合 LSTM 和 GRU 的优势，并对门控单元网络进行改进，使得该网络能更加高效、精准地从大量数据中判断异常信息，发现工业互联网络中的潜在威胁。为了保证循环神经网络连接的平滑性，本文提出改进 LSTM 层的激活函数，并在 GRU 输入层进行参数重构使得层级之间的连接更加紧密，从而实现循环神经网络的快速拟合。实验结果表明，新的均衡循环神经网络无论是在精度还是效率上都可获得较好的效果。

未来的研究中，将进一步优化当前均衡网络，从而更加提高循环神经网络的性能。此外，将对更多的异

质数据集设计其迁移工作，使其对于各类异常的检测都具有更优的表现。

参考文献

[1]王建民,刘建勋.面向智能制造的业务过程管理与服务技术专题前言[J].软件学报,2018,29(11):3239-3240.

[2] 杨婷, 张嘉元, 黄在起等. 工业控制系统安全综述[J]. 计算机研究与发展, 2022, 59(5): 1035-1053.

[3] 江平宇, 史皓良, 杨茂林等. 面向工业互联网的社群化制造模式及 3D 打印测试床研发[J]. 中国科学:技术科学, 2022, 52(1): 88-103.

[4] Xu L D, Xu E L, Li L. Industry 4.0: state of the art and future trends[J]. International journal of production research, 2018, 56(8): 2941-2962.

[5] Sisinni E, Saifullah A, Han S, et al. Industrial internet of things: Challenges, opportunities, and directions[J]. IEEE transactions on industrial informatics, 2018, 14(11): 4724-4734.

[6]胡致远, 胡文前, 李香等. 面向业务可达性的广域工业互联网调度算法研究[J]. 电子与信息学报, 2021, 43(9): 2608-2616.

[7] 李印, 陈勇, 赵景欣等. 泛在计算安全综述[J]. 计算机研究与发展, 2022, 59(5): 1054-1081.

[8] 杨安, 孙利民, 王小山, 石志强. 工业控制系统入侵检测技术综述[J]. 计算机研究与发展, 2016, 53(9): 2039-2054.

[9]Sukiasyan A,Badikyan H,Pedrosa T,Leitao P. Secure data exchange in Industrial Internet of Things[J]. Neurocomputing,2022,484: 183-195.

[10] 徐雪松, 金泳, 曾智等. 应用于工业互联网数据安全的分层轻量级高通量区块链方法[J]. 计算机集成制造系统, 2019, 25(12): 3258-3266.

[11] 胡致远, 胡文前, 李香等. 面向业务可达性的广域工业互联网调度算法研究[J].电子与信息学报, 2021, 43(9): 2608-2616.

[12] Mikolov T, Karafiát M, Burget L, et al. Recurrent neural network based language model[C]. Interspeech. 2010, 2(3): 1045-1048.

[13] Tang D, Qin B, Liu T. Document modeling with gated recurrent neural network for sentiment classification[C]. Proceedings of the 2015 conference on empirical methods in natural language processing. 2015: 1422-1432.

[14] Fischer T, Krauss C. Deep learning with long short-term memory networks for financial market predictions[J]. European journal of operational research, 2018, 270(2): 654-669.

[15] Zhao R, Wang D, Yan R, et al. Machine health monitoring using local feature-based gated recurrent unit networks[J]. IEEE Transactions on Industrial Electronics, 2017, 65(2): 1539-1548.

[16] Xu R, Cheng Y, Liu Z, et al. Improved Long Short-Term Memory based anomaly detection with concept drift adaptive method for supporting IoT services[J]. Future Generation Computer Systems, 2020, 112: 228-242.

[17] 赵雅雪, 王旭, 蒋传文,等. 基于最大信息系数相关性分析和改进多层级门控 LSTM 的短期电价预测方法[J]. 中国电机工程学报, 2021, 41(1): 135-146.

[18] Liao G, Gao W, Yang G, et al. Hydroelectric generating unit fault diagnosis using 1-D convolutional neural network and gated recurrent unit in small hydro[J]. IEEE Sensors Journal, 2019, 20: 9352-9363.

作者简介：

许荣斌（1981-），男，安徽黄山人，副教授，博士，研究方向：智能制造、工作流技术，E-mail:xuscholar@126.com；

章 宇（2000-），男，福建福州人，本科，研究方向：智能制造、业务流程管理；

+谢 莹（1981-），女，安徽黄山人，教授，博士，研究方向：机器学习、社交网络，通讯作者，E-mail:xieyingahu@126.com；

刘志强（1982-），男，山西大同人，讲师，博士，研究方向：服务计算、信息安全；
张以文（1976-），男，安徽马鞍山人，教授，博士，研究方向：服务计算、云计算和大数据分析；
闻立杰（1977-），男，河北唐山人，副教授，博士，研究方向：工作流技术、流程数据管理；