

# 基于数据可视化的复杂系统信号时序识别方法

姜 婕<sup>1</sup>, 杨 威<sup>2</sup>, 冯俊涛<sup>3</sup>, 姜 帅<sup>1</sup>

(1. 北京交通大学 机械与电子控制工程学院, 北京 100044;

2. 北京航天自动控制研究所, 北京 100854; 3. 北京航空航天大学 软件学院, 北京 100191)

**摘要:** 针对复杂系统研发及运行过程中产生的大量信号可以表征系统运行的时序健康状态这一特性, 提出了一种基于数据可视化及卷积神经网络(convolutional neural networks, CNN)智能识别的时序特征识别方法; 该方法使用数据可视化技术将信号的时序特征映射至图像, 通过训练好的特征识别模型对信号可视化图像进行时序特征的识别, 可实现系统运行时的实时智能状态监测; 选取了三种典型信号的正常及异常特征, 通过模型构建及测试分析, 验证该方法对复杂系统信号的时序特征有良好的识别效果, 可应用于对时序要求较高的复杂系统进行状态监测及故障诊断。

**关键词:** 数据可视化; 卷积神经网络; Inception-v3 模型; 时序; 特征识别

## Time Series Feature Recognition of Complex System Signals Based on Data Visualization

JIANG Jie<sup>1</sup>, YANG Wei<sup>2</sup>, FENG Juntao<sup>3</sup>, JIANG Shuai<sup>1</sup>

(1. Mechanical, Electronic and Control Engineer, Beijing Jiao Tong University, Beijing 100044, China;

2. Beijing Aerospace Automatic Control Institute, Beijing 100854, China;

3. School of Software, Beihang University, Beijing 100191, China)

**Abstract:** In view of the fact that a large number of signals generated in the process of complex system development and operation can represent the time sequence health state of system operation, a time sequence feature recognition method based on data visualization and convolutional neural networks (CNN) intelligent recognition is proposed. This method is used by the data visualization technology to map the time sequence features of the signal to the image, and uses the trained feature recognition model to recognize the time sequence features of the signal visualization image, which can realize the real-time intelligent state monitoring of the system. The normal and abnormal characteristics of three typical signals are selected. Through model construction and test analysis, it is verified that the method has good recognition effect on the time sequence characteristics of complex system signals, it can be applied to condition monitoring and fault diagnosis of complex systems with high time sequence requirements.

**Keywords:** data visualization; convolutional neural network (CNN); inception-v3 model; time series; feature recognition

## 0 引言

复杂系统在运行过程中, 各传感器随时间及工况将产生不同的数字信号及模拟信号, 根据信号时序特征及阈值, 可以反映系统的运行状态, 在系统运行异常时及时对其进行干预。传统的信号处理及数据分析

方法往往基于规则及案例, 并不适用于复杂系统。机器学习领域的发展推动特征识别领域的发展, 使用机器学习解决此类问题成为可能<sup>[1]</sup>。

根据复杂系统(本文以飞行器为例)运行过程中产生的数据特点, 模型需要着重对数据的时序特征进

收稿日期: 2021-03-27; 修回日期: 2021-07-14。

作者简介: 姜 婕(1993-), 女, 山东省费县人, 硕士, 主要从事机电设备智能故障诊断方向的研究。

通信作者: 杨 威(1981-), 男, 北京市人, 硕士, 高级工程师, 主要从事精密仪器与机械方向的研究。

引用格式: 姜 婕, 杨 威, 冯俊涛, 等. 基于数据可视化的复杂系统信号时序识别方法[J]. 计算机测量与控制, 2022, 30(1): 252-257.

行识别。RNN 及其衍生模型虽然可以对时序数据进行专业的处理, 但是仍然存在计算量大处理效率低且数据量大时存在梯度弥散或梯度爆炸的可能, 对于复杂系统实时性要求高且运行时间长、数据量大的情况存在难以弥补的缺陷。

本方法通过将时序特征映射至图像, 使用 CNN 对时序特征图进行分类识别实现对时序数据的处理, 处理速度快, 且在复杂系统长时间运行产生的大量数据进行时序处理时效果依然稳定。

本文第一章对目前常用的识别模型进行了概述与对比, 结合需求特点最终确定复杂系统在时序特征识别时所用模型的类型。

第二章使用 ECharts 将飞行器运行时的二进制流数据进行可视化处理实现了时序特征的映射。

第三章结合前两章的飞行器数据流的时序特征提取方法, 补充了图像预处理方法及 TFRecord 数据集的制作方法, 将大量的飞行器历史测试数据制作成飞行器运行过程中各信号时序变化的时序特征数据集。

第四章构建 Sample 模型、修改 Inception 模型的同时, 使用第三章制作的数据集在 Tensorflow 环境下进行模型的训练及调参, 完成 CNN 模型的构建。

第五章使用第三章制作的数据集, 对第四章构建的 CNN 模型进行了验证及分析, 最终在第六章通过总结与概括得出结论。本文的研究路线如图 1 所示。

## 1 模型方案选择

在对飞行器各系统数据进行处理及分析时, 主要考虑其时序变化特征及故障数据分类的准确性。现在常用的时序识别方法主要集中在语义识别、语音识别及预测等方面。与此相关的神经网络主要有具有记忆功能的循环神经网络 (RNN, recurrent neural network), 解决 RNN 长期依赖问题的长短期记忆网络 (LSTM, long short-term memory) 及其它衍生网络<sup>[2]</sup>。CNN 对于一对一映射关系规则严谨, 尤其对于图像的识别及分类有良好的效果, 因此从 RNN、LSTM 及 CNN 三种神经网络进行对比研究。

文献 [3] 优化了一种 RNN 方法, 用来预测短时间内特定海洋的区域碰撞风险, 该方法的预测精度较高且效果良好。

文献 [4] 使用 LSTM 作为准确估算锂离子电池 SOC 的方法, 其估计误差小且均方根误差低于 1%, 验证方法可行。

文献 [5] 融合了两个二维的 CNN 神经网络, 对超声波心动图进行了时域特征及空域特征, 实现了对超声波心动图的视点分类, 在心脏病诊断方面辅助优势明显。

文献 [6] 使用改进后的 Inception-v3 模型进行乳腺病理图像的二分类及多分类, 取得了良好的识别精度, 提高了识别效率。

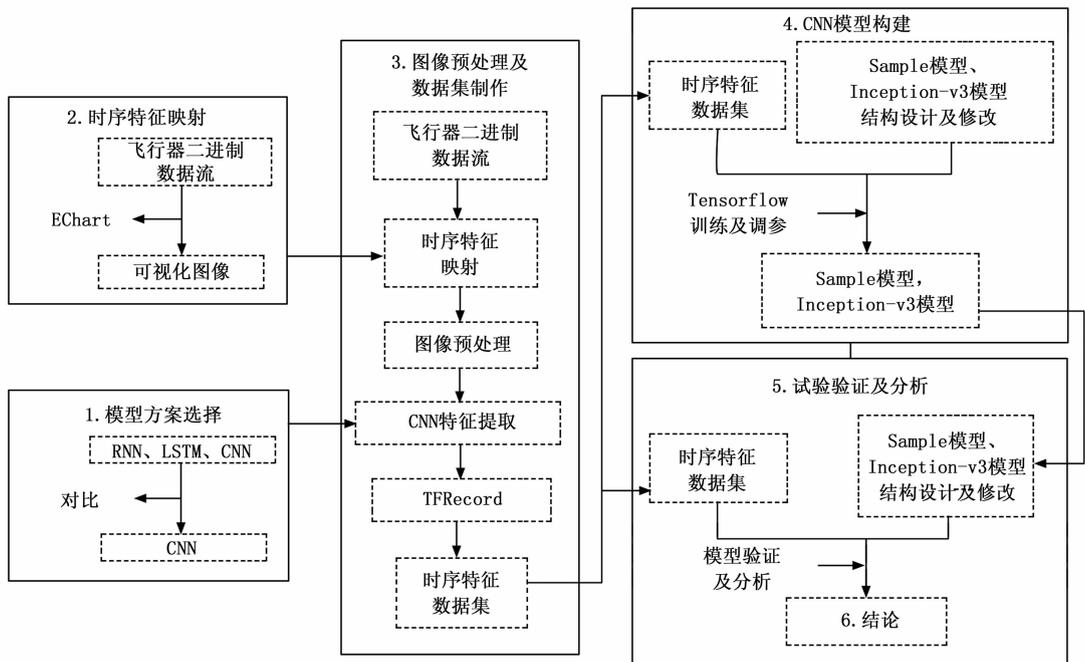


图 1 研究路线图

根据 RNN、LSTM 及 CNN 的特点, 结合复杂系统运行过程中关键信号的识别需求分析如下:

1) RNN 对于时序处理效果较好, 但是在模型训练时易出现由于共享参数导致的梯度弥散或梯度爆炸。

2) LSTM 解决了 RNN 的梯度问题, 但是计算费时, 且对于量级较大的长序列表现不佳。由于飞行器数据变化幅度一般较小, 且飞行时间长、数据量大, 使用 LSTM 存在明显的限制约束。

3) CNN 的输入输出长度固定, 便于实现确定的映射到确定的输出, 且识别性能普遍良好。

基于上述考虑, 采用 CNN 作为复杂系统信号时序特征的识别的基本模型。选择使用可视化技术将信号的时序特征映射至图像, 通过 CNN 对时序特征图像的认识实现复杂系统信号时序特征的识别的方法进行时序特征识别。

## 2 时序特征映射

在系统运行过程中, 采集不同时序状态下的特征信号 (电压、电流、角度变化、速度变化及位移等), 按照产生的时间顺序存储至数据库中。使用数据可视化工具, 将信号数据按照时序 (以时间为横坐标) 进行二维图像化处理, 实现信号时序特征的图像映射, 可以更直观地反映其时序变化趋势及特征。

如图 2 所示, 表示转台转动角度随时间变化的部分可视化图像。因为本文的目的为对信号图像时序特征的处理, 所以在可视化时对其进行去单位化的处理, 统一横、纵坐标尺度, 只保留其时序特征曲线形状的有效性。

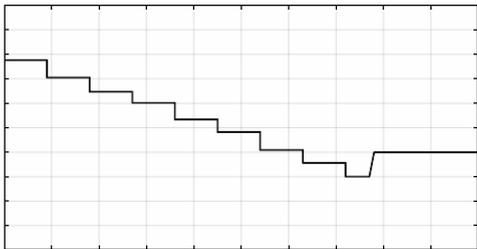


图 2 部分转动角度信号的可视化图像

使用 echarts 可以将时序数据可视化为二维图表, 通过对图表进行分析归纳其时序特征。由于神经网络构建使用 Python 语言, 因此使用 Pyecharts 将时序特征映射至图像<sup>[7]</sup>。

时序数据图表绘制的代码逻辑如表 1 所示, 以时间 (timestamp) 为横坐标, 信号数据 (data) 为纵

坐标进行曲线绘制。

表 1 Pyecharts 曲线绘制代码

行号	代码
1	From pyecharts import Line
2	attr = timestamp
3	l = data
4	line = Line()
5	line.add("line", attr, l)
6	overlap = Overlap()
7	overlap.add(line)
8	Overlap

## 3 图像预处理及数据集

可视化处理后的数据图像按照故障源进行分类, 为了加快处理速度, 提高模型的识别及训练效率, 对图像进行以下操作:

1) 统一图像尺寸及分辨率。由于模型识别主要针对曲线的时序趋势特征进行研究与识别, 并不在意从图像的度量单位及时间单位, 因此在图像预处理时对曲线坐标进行模糊处理。且为加快图像处理速度, 降低图像分辨率, 只保留曲线特征像素。

2) 按照类别进行编号。按照图像的分类对图像文件进行分类编号。

以如图 3 所示的六种时序特征为例进行时序特征识别说明, 预处理前信号时序特征图像及其分类如 3 所示, 预处理后信号时序特征图像及其分类如 4 所示。因本文的目的为对信号图像时序特征的处理, 因此在可视化时对其进行去单位化的处理, 统一横、纵坐标尺度, 只保留其时序特征曲线形状的有效性。

图像预处理后, 需要将其制作成数据集, 以便后续进行模型训练测试及验证。在此采用 TFRecord 作为数据集文件生成的方法及格式。

## 4 CNN 模型构建

在构建模型时, 为了体现模型对时序特征的识别效果, 使用了两种模型: 构建了 Sample 模型作为基础识别模型, 运用迁移学习思想, 使用 Inception-v3 模型及其预训练参数。

### 4.1 Sample 模型构建

由于示例的时序特征及数据集较简单, 构建一个八层的 CNN 网络的以利用数据集进行训练。CNN 网络结构包含两个卷积层, 两个池化层, 两个全连接

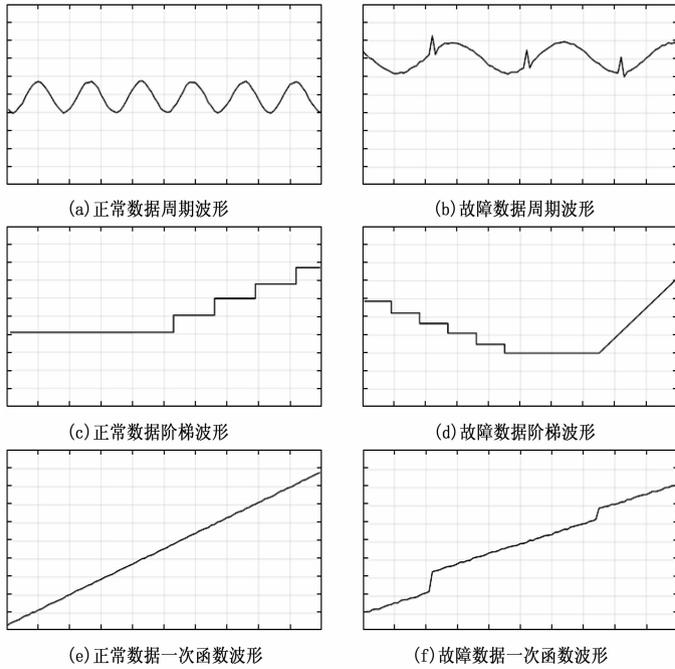


图 3 可视化后信号时序特征图

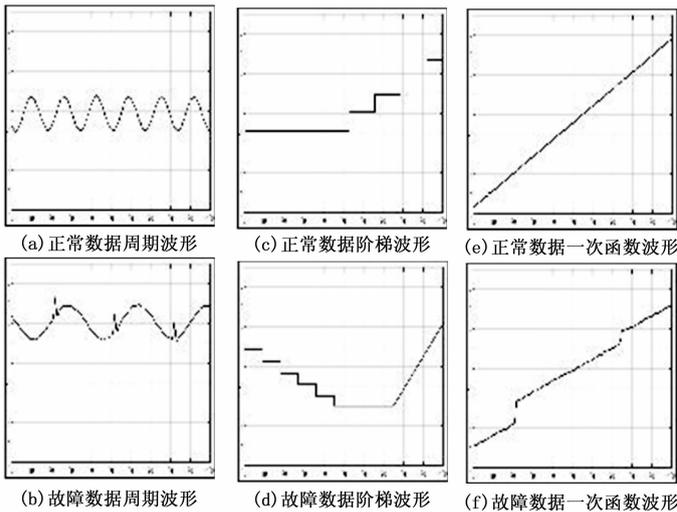


图 4 预处理后的信号时序特征图

层, 一个分类层, 使用 ReLU 作为激活函数。大致过程为: “初始化权值参数→计算损失值→调制权值参数” 过程的循环。

在对 Sample 模型完成结构设计及修改后, 在 tensorflow 环境下使用根据历史试验数据制作好的飞行器时序特征数据集进行模型训练及调参, 完成 Sample 模型的训练。其训练的最终参数如表 2 所示。

由于 Sample 模型简单, 层级较浅且参数较少, 数据集的数据量不大且时序特征少, 因此设定训练次数为 500 此即可获得收敛度较高的模型, 其训练过程中的准确率曲线如图 5 所示, 损失曲线如图 6 所示。

表 2 Sample 模型训练参数表

名称	Name	数值
学习速率	learning-rate	0.05
样本批次容量	Batchsize	100
权重衰减	weightdecay	0.1
训练次数	epoch-number	500

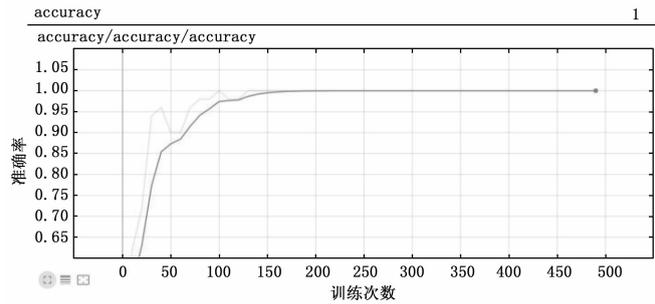


图 5 Sample 模型准确率曲线

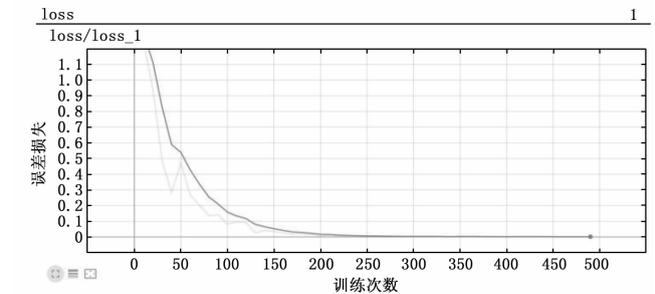


图 6 Sample 模型损失曲线

分析训练过程中的准确率曲线, 在迭代 150 次以后准确率在 99.5% 左右收敛稳定。

分析训练过程中的误差损失曲线, 在 150 次左右时误差损失稳定在 0.1 以下, 且仍具有收敛趋势, 在迭代 200 次左右时基本收敛稳定。

#### 4.2 Inception-v3 模型构建

Google Inception Net 在大量数据的处理, 尤其是内存或计算资源有限的场合具有计算效率高且分类准确的特点, 采用全局平均池化层取代全连接层, 一般称之为 Inception-v1。后续优化的 Inception-v2 中加快了收敛速度<sup>[10-11]</sup>。在 Inception-v3 中为了降低参数量, 减轻过拟合, 将二维卷积层拆分成两个一维卷积层<sup>[12]</sup>。

为充分利用训练好的 Inception-v3 模型结构及预训练参数, 仅对 Inception-v3 模型瓶颈层的输出来训练一个新的全连接层, 以处理时序特征图像的分类问题, 完成模型的适配。

在对 Inception-v3 Sample 模型完成结构设计及

修改后，在 tensorflow 环境下使用根据历史试验数据制作好的飞行器时序特征数据集进行模型训练及调参，完成 Inception-v3 模型的训练其训练的最终参数如表 3 所示。

表 3 Inception-v3 模型训练参数表

名称	Name	数值
学习速率	learning-rate	0.05
样本批次容量	Batchsize	100
权重衰减	Weightdecay	0.1
训练次数	epoch-number	1000
输入图片尺寸	IMG_W * IMG_H	299 * 299
分类数	N_CLASSES	6

由于 Inception-v3 模型较为复杂，层级较深参数较多，因此设定训练次数为 1 000 次可获得收敛度较高的模型，其训练过程中的准确率曲线如图 7 所示，损失曲线如图 8 所示。

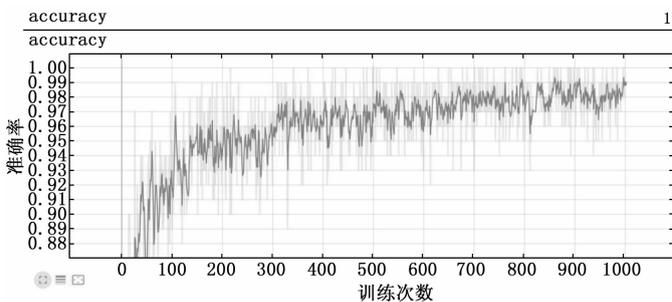


图 7 Inception-v3 模型准确率曲线

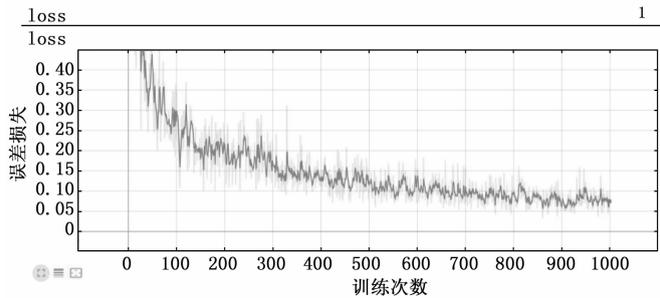


图 8 Inception-v3 模型损失曲线

分析训练过程中的准确率曲线，在迭代 100 次以内时虽然波动较大但是收敛迅速，900 次以后基本稳定在 97% 以上，收敛效果较好。

分析训练过程中的误差损失曲线，在 100 次之前收敛速度较快，100 次以后收敛速度变缓但收敛趋势良好，在迭代 900 次以后基本收敛至 0.1 以下。

### 5 试验验证与分析

模型训练结束后，需要对模型的识别率、准确率

进行验证分析。使用训练好的神经网络在验证集中随机读取一张图像进行识别，将读取图像及其识别结果进行打印输出。步骤如下所述。

Step1: 读取一张图片，对图片进行预处理操作。

Step2: 应用 API 中的结构构建新的网络结构。

Step3: 载入模型，对模型结构的输出结果进行 softmax，即为预测值。

Step4: 图像读取与识别，打印准确率及识别的图像。

识别结果示例如图 9 所示；对六种时序特征波形，使用 Sample 模型及 Inception-v3 模型分别进行识别测试 100 次，统计测试结果如表 4 所示。

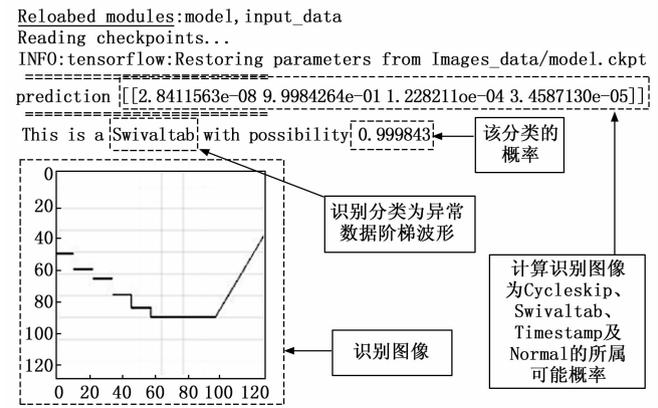


图 9 识别结果打印示意图

表 4 Sample 模型及 Inception-v3 模型识别结果测试统计表

波形名称	Sample 模型/次	Inception-v3 模型/次	概率统计 %
正常数据周期波形	99	99	99
异常数据周期波形	100	100	100
正常数据阶梯波形	99	98	98.5
异常数据阶梯波形	100	100	100
正常数据一次函数波形	100	96	98
异常数据一次函数波形	99	98	98.5
概率统计/%	99.667	98.167	—

根据以上测试识别结果，Sample 模型的识别准确率为 99.667%，识别效果要优于识别准确率为 98.167% 的 Inception-v3 模型，但是经过实际的训练与识别，Inception-v3 的识别速度普遍比 Sample 模型快。

从识别分类来看，正常数据阶梯波形的识别概率为 100%，但是异常数据阶梯波形的识别概率仅为 98.5%，由于正常数据及异常数据的阶梯波形相似、Sample 模型及 Inception-v3 模型识别结果的统一性说明识别差距的原因可定位为正常数据及异常数据阶梯波形数据集的差距不大，两类阶梯波形数据采集的

样本问题导致 Sample 模型及 Inception-v3 模型对阶梯波形正常、异常的识别结果存在误差。

## 6 结束语

本文提供了一种基于数据可视化的复杂系统信号时序特征识别的方法。不需要深入了解复杂系统的运行及控制原理, 即可对时序特征进行识别。通过采集信号数据进行可视化处理, 将信号时序特征映射至图像生成数据集; 构建、训练、测试和验证模型, 形成最终的时序特征识别模型, 然后通过此模型以模型预测的形式对信号数据的可视化时序特征进行识别。

文中选取了复杂系统中特征参数随时间变化的周期波形、阶梯波形、一次函数波形三种典型波形在正常及异常情况下的时序特征进行了研究及验证, 证实本文所述方法对单一特征波形的故障有良好的识别效果。下一步可以针对复杂特征波形及多特征融合波形的识别进行研究及改进, 增强本文所述模型识别方法的适用性及实用性。

本文研究基于数据可视化的复杂系统信号时序特征识别的方法对大型复杂系统的信号时序特征识别具有一定的适用性, 不仅可以应用在飞行器在飞行过程中的故障诊断, 还可以应用于态势感知与态势评估、智能指挥决策等等方向。

### 参考文献:

[1] 姚宏洋, 林晓烘, 宗思光. 基于时序的 D-S 证据理论雷达辐射源识别 [J]. 舰船电子工程, 2019, 39 (1): 47-49, 117.

[2] 夏瑜璐. 循环神经网络的发展综述 [J]. 电脑知识与技术, 2019, 15 (21): 182-184.

[3] LIU D P, WANG X, CAI Y, et al. A Novel Framework of Real-Time Regional Collision Risk Prediction Based on the RNN Approach [J]. Journal of Marine Science and Engineering, 2020, 8 (3): 224.

[4] GENG P, XU X Y, TARASIUK Tomasz. State of charge estimation method for lithium-Ion batteries in all-electric ships based on lstm neural network [J]. Polish Maritime research, 2020, 27 (3): 100-108.

[5] GAO X H, LI W, LOOMES M, WANG L Y. A fused deep learning architecture for viewpoint classification of echocardiography [J]. Information Fusion, 2017, 36: 103-113.

[6] 詹翔, 张婷, 林聪, 等. 基于深度学习的乳腺癌

理图像分类实验方法 [J]. 计算机应用, 2019, 39 (S2): 118-121.

[7] 齐超, 周艳尼, 万上英, 等. 电力大数据的可视化展现技术 [J]. 电子技术与软件工程, 2018 (4): 183-185.

[8] 陈南真. 复杂场景低分辨率人脸识别及其在身份识别系统的应用 [D]. 成都: 电子科技大学, 2018.

[9] 龙满生, 欧阳春娟, 刘欢, 等. 基于卷积神经网络与迁移学习的油茶病害图像识别 [J]. 农业工程学报, 2018, 34 (18): 194-201.

[10] HU J, SHEN L, SUN G. Squeeze-and-Excitation Networks [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, 18-23.

[11] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C] // International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012: 1097-1105.

[12] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2818-2826.

[13] 刘燕军. 神经网络的自学习及其应用 [D]. 合肥: 中国科学技术大学, 2018.

[14] 黄一龙, 李培峰, 朱巧明. 事件因果与时序关系识别的联合推理模型 [J]. 计算机科学, 2018, 45 (6): 204-207, 234.

[15] 李西平, 谷立臣, 寇雪芹. 基于超声时序神经网络目标识别的塔机安全预警 [J]. 中国机械工程, 2016, 27 (16): 2190-2195.

[16] 何巍, 胡久辉, 赵婷, 等. 基于模型的运载火箭总体设计方法初探 [J]. 导弹与航天运载技术, 2021 (1): 12-17, 32.

[17] 张浩博, 付冬梅, 周珂. 时序增强的视频动作识别方法 [J]. 模式识别与人工智能, 2020, 33 (10): 951-958.

[18] 高明柯, 赵卓, 逢涛, 等. 基于注意力机制和特征融合的手势识别方法 [J]. 计算机应用与软件, 2020, 37 (6): 199-203.

[19] 罗盛炜. 健康监护时序信号解析与应用算法研究 [D]. 杭州: 浙江大学, 2020.

[20] 郭耀华. 基于深度学习的车辆驾驶状态识别算法研究 [D]. 北京: 北京邮电大学, 2019.