

DOI:10.19344/j.cnki.issn1671-5276.2024.05.040

# 基于数据关联表征的工业零件检测

黎洸安<sup>1,3</sup>, 杨冬平<sup>2,3</sup>

- (1. 福州大学 先进制造学院, 福建 晋江 362251;  
2. 之江实验室 人工智能研究院混合增强智能研究中心, 浙江 杭州 311101;  
3. 中国科学院 海西研究院泉州装备制造研究中心, 福建 晋江 362216)

**摘要:**为实现自动化工业生产中零件的自动识别,对深度残差网络的残差结构进行改进。将储备池模块应用到残差网络的残差连接结构中,使得输入数据的各个区域互相关联后重新进行表征。将提出的模型在工业零件数据集以及公开数据集上与其他深度学习模型进行比较。结果表明:在工业零件数据集上提出的具有数据关联表征的残差网络 ResNet18-RC 比 ResNet18 提高了 0.17%,且均比其他模型的识别率高。在 CIFAR-10、CIFAR-100、Tiny-ImageNet 等公开数据集上,具有数据关联表征的残差网络 ResNet50-RC 分别比 ResNet50 提高了 0.35、0.62、0.54、1.31 个百分点的精度,具有很好的图像检测性能。

**关键词:**图像识别;残差神经网络;储备池计算;数据关联表征;工业零件

**中图分类号:**TP391.41 **文献标志码:**A **文章编号:**1671-5276(2024)05-0191-04

## Industrial Parts Detection Based on Data Correlation Representation

LI Weian<sup>1,3</sup>, YANG Dongping<sup>2,3</sup>

- (1. School of Advanced Manufacturing, Fuzhou University, Jinjiang 362251, China; 2. Zhijiang Lab, Human-Machine Augmented Intelligence Research Institute of Artificial Intelligence Reserch Center, Hangzhou 311101, China;  
3. Quanzhou Equipment Manufacturing Reserch Center of Haixi Institute, Chinese Academy of Science, Jinjiang 362216, China)

**Abstract:** For the realization of automatic identification of industrial parts in automated industrial production, the residual structure of deep residual network is upgraded. The reservoir module is applied to the residual connection structure of the residual network so that each area of the input data can be represented after being correlated with each other. The proposed model is compared with other deep learning models on industrial parts dataset and public dataset. The experimental results show that the proposed residual network with data correlation representation Resnet18-RC is 0.17%, better than ResNet18 on the industrial parts dataset, and the recognition accuracy is higher than other models. The public dataset like CIFAR-10, CIFAR-100 and Tiny-Imagined indicates that the residual network Resnet50-RC is respectively 0.35, 0.62, 0.54, 1.31 per cent, higher than ResNet50 in terms of accuracy, and has good image recognition performanc.

**Keywords:** image recognition; residual neural network; reservoir computing; data correlation representation; industrial parts

## 0 引言

近年来,随着人工智能技术的发展,自动化的生产车间已经越来越普遍。因此,对于如螺丝、螺母等零件的正确识别非常重要。近年来有人通过 GAPS0-SVM 方法进行钣金零件图像识别<sup>[1]</sup>。机器错误地识别和抓取零件可能会造成难以估量的损失。为了解决这个问题,通过先进的图像识别模型算法来完成对零件的正确选择尤为重要。

图片分类算法的研究一直是最近几年来的热点,期间许多性能强悍的模型不断涌现。比如在 2012 年的 ILSVRC<sup>[2]</sup> 分类挑战大赛中,基于卷积

结构的深度网络 AlexNet<sup>[3-4]</sup> 取得了很好的成绩。除此之外,VGG<sup>[5]</sup> 网络模型的 top-5 准确率也达到了 93.2%。而 ResNet<sup>[6]</sup> 通过残差学习的方式,解决了 CNN 深度很深时性能退化的问题,提高了网络的可扩展深度,并在 ILSVRC 中达到了很高的分类准确率。

在图像识别的算法中,卷积模块、残差结构以及注意力机制<sup>[7]</sup> 等模块互相组合,有效提取图像特征。除了这些经典的深度神经网络模块,其他的一些机制或许可以被利用。近年来,一种新的机器学习范式—储备池计算(reservoir computing, RC)<sup>[8]</sup>,开始被广泛地使用和研究,储备池也在语

**基金项目:**国家自然科学基金项目(12175242)

**第一作者简介:**黎洸安(1998—),男,广东惠州人,硕士研究生,研究方向为深度学习和机器学习,bbigcat111@163.com。

音识别等一些工程领域有很好的效果。

本文对残差神经网络的残差结构进行改进。通过将处于混沌边缘状态的储备池模块<sup>[9-11]</sup>应用在残差神经网络的残差层中,提出了具有数据关联表征残差结构的残差神经网络模型。在工业零件数据集上,改进 ResNet18 模型,并对相关的图像识别模型进行训练并对比验证。除此之外,为验证提出的方法在图像识别模型的有效性,本文在 ResNet50 的残差网络基础上进行改进,在 CIFAR-10, CIFAR-100<sup>[12]</sup>, Tiny-ImageNet<sup>[13]</sup> 等图像数据集上与其他网络模型进行对比实验,充分验证了本文所提出的模型对于正确物体图像识别的可能性。

### 1 模型与方法

#### 1.1 残差连接的改进

残差连接是当前神经网络的最重要结构之一,它保证了人工神经网络能够往更深的深度发展,并使网络更好地收敛。图 1 展示了 ResNet18 网络中残差连接的方式以及对残差连接结构的改进。图 1(a)表示普通的残差连接结构,输入经过卷积层后通道数改变,输出与输入直接相加前,原输入需要卷积成与输出相同的通道数;图 1(b)的结构表示对残差结构的改进,在残差连接中加入了储备池模块。

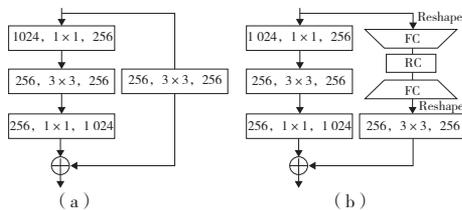


图 1 残差连接结构

#### 1.2 储备池模块

储备池计算由输入层、储备池以及读出层 3 部分构成。储备池的基本结构如图 2 所示,其中

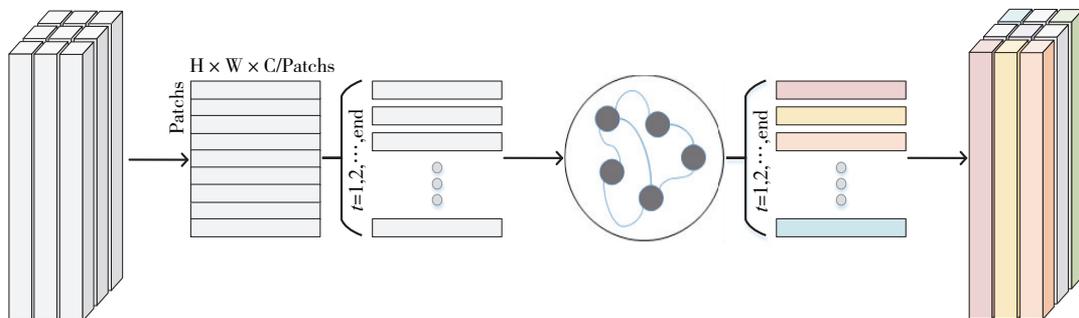


图 3 数据关联表征示意图

$W_{in}$ 表示输入权重,在 $[0,1]$ 的范围内随机均匀采样。处于中间层的储备池,是一个 RNN 网络,内部具有循环连接权重  $W_{res}$ ,神经元之间稀疏连接,其中  $x_i$ 表示神经元  $i$  的值。在储备池中,  $W_{in}$ 和  $W_{res}$ 均不训练,在初始化后保持固定不变。  $W_{out}$ 表示读出权重,它将储备池和读出层连接起来。

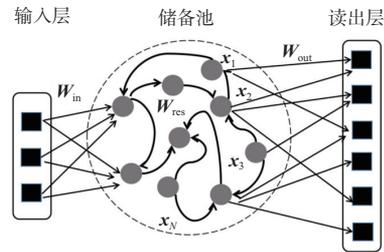


图 2 储备池模块结构

储备池内部的神经元状态更新方式如式(1)所示。

$$x(n+1) = (1-\alpha)x(n) + \alpha \tanh[W_{res}x(n) + W_{in}x_{in}(n+1)] \quad (1)$$

式中:  $x(n)$ 是第  $n$  时刻储备池神经元电压组成的状态列向量;  $\alpha$ 表示神经元的泄露率,在 $[0,1]$ 区间内取值;  $\tanh$ 是非线性激活函数。

#### 1.3 数据关联表征残差连接

本文将普通的残差连接更改为具有储备池模块的数据关联表征残差连接结构,如图 1 所示。

在残差连接的卷积层之前,加入了储备池模块。首先将输入数据按照区域分为一个个的小块,每一个小块按照从左到右、从上到下的顺序依次进行拉平,在某个时刻输入给储备池,每个时刻只输入一个小图片块。小图片块输入给储备池后,网络在混沌边缘的状态下,将不同部分的小数据块互相关联起来,相当于一个全局的注意力操作,数据被表征出了更多的特征,变为了高维的网络状态。随后网络状态经过一个线形层读出信息后再进行一个相反的操作,把网络状态变为输入的原始形状。图片分块输入给储备池的过程如图 3 所示。

## 2 实验

### 2.1 数据集

为了提高模型在对于工业零件识别的准确率和有效性,首先从公开数据集中选取了螺栓(bolt)、定位销(locatingpin)、螺母(nut)、垫片(washer)等。除此之外,为了验证提出的模型在多物体数据集上的物体识别能力和泛化能力,本文还在 CIFAR-10、CIFAR-100、Tiny-ImageNet、FLOWER 数据集上与其他的网络模型进行了比较。

### 2.2 实验环境和参数设置

本文实验基于 Pytorch 框架,在 NVIDIA V100 上完成网络的训练和测试过程。实验使用 SGD 优化器,采用 0.01 的学习率,学习率随着实验过程逐渐减小,优化器中的参数 weight-decay 设置为  $1 \times 10^{-4}$ 。损失函数选择了交叉熵损失 CrossEntropyLoss。储备池的特征值谱半径设置为 1,储备池神经元数目设置为 1 000。

### 2.3 工业零件分类实验结果

本文根据 2.2 节中的实验环境和参数设置,在工业零件图片数据集上进行了图片识别验证。对 ResNet18 的残差网络模型进行改进,在残差结构中加入 1.3 节所示的数据关联表征,提出了 ResNet18-RC 模型。该模型与其他模型在此工业零件数据集上的比较如表 1 所示。ResNet18-RC 达到了 99.67% 的准确率,能够准确地识别 4 种工业零件的图片。由表 1 可知,ResNet18-RC 与 ResNext、VGG16、DenseNet 和 ResNet18 等模型相比,具有更高的模型识别准确率。

表 1 工业零件识别模型准确率 单位:%

模型方法	准确率
ResNext	98.68
VGG16	92.84
DenseNet	98.51
ResNet18	99.50
ResNet18-RC	99.67

在此实验中,随机选择一些图片进行测试,模型分类零件的热力图如图 4 所示。图 4 表明提出的数据关联残差结构模型能够观察到各类不同零件上的关键特征。



图 4 模型分类零件的热力图

### 2.4 公开数据集分类实验结果

为了进一步验证本文提出的数据关联表征残差结构在模型上的效果,接下来在公开数据集上验证模型的识别准确率。本文继续使用 2.2 节中的实验环境和参数设置,在比 ResNet18 深度更深的残差神经网络 ResNet50 上对残差结构进行改进,提出了 ResNet50-RC 模型,在公开数据集 CIFAR-10、CIFAR-100、Tiny-ImageNet 以及 FLOWER 上与其他模型 VGG16、ResNext、DenseNet 以及 ResNet50 进行比较,在测试集中计算。实验结果如表 2 所示。

表 2 公开数据集图像识别模型准确率

模型方法	单位:%			
	CIFAR-10	FIFAR-100	Tiny-ImageNet	FLOWER
ResNext	96.25	77.46	67.01	74.81
VGG16	96.03	76.89	67.30	79.84
DenseNet	96.52	79.07	65.10	80.23
ResNet50	95.31	77.15	66.31	76.01
ResNet50-RC	95.66	77.77	66.85	77.32

如表 2 所示,无论是在 CIFAR-10、CIFAR-100 还是在 Tiny-ImageNet 数据集上,所提出的具有数据关联表征残差结构的模型 ResNet50-RC 和其他模型相比,均有更好地准确率。从表 2 中可以看出,ResNet50-RC 模型在 CIFAR-10、CIFAR-100、Tiny-ImageNet 以及 FLOWER 数据集上的准确率比 ResNet50 模型分别高了 0.35、0.62、0.54、1.31 个百分点。

表 1 和表 2 都表明:提出的具有数据关联表征残差结构在图片识别上能够提高网络的准确率,无论是在机械零件数据集,还是在公开数据集上,具有数据关联表征残差连接的网络均有很好的效果。

最后在 CIFAR-10 数据集上,对 ResNet50-RC 模型在不同的 Stage 上加入数据关联表征残差连接进行了消融实验,结果如表 3 所示。表 3 表明,在第 3、第 4 个 stage 改进残差连接的模型分类准确率越高。

**表 3 ResNet50-RC 改进残差连接位置的消融实验**  
单位: %

ResNet50-RC stage	准确率
1-2-3-4	94.51
2-3-4	95.41
3-4	95.66
4	95.11

### 3 结语

在自动化工业组装生产中,提高机械零件的识别率是降低生产事故、减少损失的关键之一。通过在残差神经网络的残差结构中引入储备池模块,对残差输入进行数据关联表征,实现了全局注意力的操作,增加了数据的特征,提高了网络识别图像的准确率。实验结果表明:深度残差网络结合基于储备池的数据关联表征残差连接,能够实现更高的图像识别准确率,无论是在工业机械零件数据集还是在公开数据集中都有一定的提升。

#### 参考文献:

- [1] 方舟,程筱胜,崔海华,等. 基于 GAPSO-SVM 的钣金零件图像识别方法[J]. 机械制造与自动化,2020,49(5):116-118,122.
- [2] RUSSAKOVSKY O, DENG J, SU H, et al. ImageNet Large scale visual recognition challenge [J]. International Journal of Computer Vision,2015,115(3):211-252.
- [3] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017,

60(6):84-90.

- [4] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA; IEEE,2015:1-9.
- [5] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. (2014-04-04) [2022-10-08]. <https://doi.org/10.48550/arXiv.1409.1556>.
- [6] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA; IEEE,2016:770-778.
- [7] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: transformers for image recognition at scale [EB/OL]. (2020-10-22) [2022-10-08]. <https://doi.org/10.48550/arXiv.2010.11929>.
- [8] SCHRAUWEN B, VERSTRAETEN D, CAMPENHOUT J V. An overview of reservoir computing: theory, applications and implementations [C]//ESANN '2007 Proceedings - European Symposium on Artificial Neural Networks, Bruges (Belgium): [s.n.], 2007: 471-482.
- [9] LUKOŠEVIČIUS M, JAEGER H. Reservoir computing approaches to recurrent neural network training [J]. Computer Science Review,2009,3(3):127-149.
- [10] SOMPOLINSKY H, CRISANTI A, SOMMERS H J. Chaos in random neural networks[J]. Physical Review Letters,1988,61(3):259-262.
- [11] LANGTON C G. Computation at the edge of chaos: phase transitions and emergent computation [J]. Physica D: Nonlinear Phenomena,1990,42(1/2/3):12-37.
- [12] KRIZHEVSKY A, HINTON G. Learning multiple layers of features from tiny images [R]. Technical report, University of Toronto, 2009.
- [13] LE Y, YANG X S. Tiny ImageNet visual recognition challenge[J]. [S.I:s.n.]. 2015.

收稿日期:2023-03-08