

联合损失监督的高频工件深度学习识别算法

杨涛¹, 欧阳¹, 苏欣², 吴学杰¹, 李柏林¹

(1. 西南交通大学 机械工程学院, 四川 成都 610031;
2. 中国电子科技集团公司第十研究所, 四川 成都 610036)

摘要:针对高频工件种类多、类间相似度较高造成的识别准确率低的问题,提出一种联合损失监督的深度学习识别算法。搭建基于卷积神经网络的图像特征向量编码模型,采用角度余量损失替换 SoftMax 损失,以减小工件类内特征之间的距离,完成同类工件的鲁棒性表示;引入隔离损失以增大异类工件特征之间的距离,实现异类工件的良好性区分。实验结果表明:该方法相较于传统的图像识别方法,识别准确率更高;相较于单一的角度余量和隔离损失,识别准确率分别提高了 3.97% 和 13.88%。

关键词:工件识别;联合损失;监督学习;卷积神经网络

中图分类号:TP391.41 **文献标志码:**A **文章编号:**1671-5276(2023)01-0030-04

High Frequency Workpiece Deep Learning Recognition Algorithm Based on Joint Loss Supervision

YANG Tao¹, OU Yang¹, SU Xin², WU Xuejie¹, LI Bailin¹

(1. School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China;
2. The 10th Research Institute of CETC, Chengdu 610036, China)

Abstract: To improve the low recognition accuracy caused by wide varieties of high frequency artifacts and high similarity between classes, a deep learning algorithm with joint loss supervision is proposed. An image feature vector encoding model is built based on convolutional neural network, and the SoftMax loss is replaced by the angle margin loss to reduce the distance between the features within the workpiece class and complete the robust representation of similar workpieces. The isolation loss is introduced to increase the distance between the features of heterogeneous workpieces and achieve good discrimination of heterogeneous workpieces. The experimental results show that the recognition accuracy of the proposed method is higher than that of the traditional image recognition method, with the single angle margin increasing by 3.97% and isolation loss 13.88% respectively.

Keywords: workpiece recognition; joint loss; supervised learning; convolutional neural network

0 引言

在中国制造 2025 目标的引领下,随着我国科学技术的发展,制造业实现自动化升级的任务急需完成。高频工件是我国航空航天硬件设备的重要组成部分,不仅种类繁多,而且加工工序较为繁琐。不同高频工件在热处理后对应的加工工序不同,所以需要在热处理后正确识别高频工件,以确保加工质量,保障航空航天设备的安全可靠。目前主要采用人工方式识别热处理后的高频工件,人眼的疲劳会降低工作效率,甚至影响工件图号识别的准确性。所以,迫切需要对高频工件的识别进行自动化升级。

图像识别作为一种先进的自动化技术,近年来已广泛应用于工业领域中。图像识别技术可以大致分为基于人工特征的传统方法和基于深度学习的智能方法。传统的图像识别方法中,主要是采用图像编码的方式,手动提取图像特征用于识别。如文献[1-2],采用局部二值模式

(local binary pattern, LBP) 和方向梯度直方图(histogram of oriented gradient, HOG)特征来训练支持向量机,最后使用分类器对图像进行识别。传统的图像识别方法虽然能识别类间差别较大的高频工件,但对于类间差异较小的工件,其泛化性能较差,鲁棒性有限。

随着深度学习在图像分类^[3]、目标检测^[4]、图像分割^[5]和人脸识别^[6]等研究领域的发展和进步,对工业图像的识别也得到了重视和研究。伍锡如等^[7]和郑健红等^[8]针对工业机器人抓取目标区域的识别问题,使用特征对神经网络进行训练,最后用训练好的卷积神经网络来对分拣目标进行识别,取得了不错的效果;王一鸣等^[9]和常海涛等^[10]针对工件缺陷的问题,采取训练好的卷积神经网络提取目标特征,达到提高缺陷检测率的目的。上述方法检测工件种类较为单一,类间差异较大,无法良好适用于本文研究的高频工件。此外,上述神经网络模型的训练困难,所需存储空间比较大,训练时间长,不适合用在高频工件识别的嵌入式自动化系统中。

基金项目:四川省重大科技专项(18ZDZX0140)

第一作者简介:杨涛(1997—),男,四川雅安人,硕士研究生,研究方向为图像处理与识别模式。

因此,本文基于角度余弦损失^[11]可以解决同类工件特征相距太远的问题,以及隔离损失^[12]可以改善不同类别工件之间的距离情况,提出了在改进深度可分离卷积 MobileNetV2 网络^[13]的基础上,采用角度余弦损失和隔离损失的联合损失函数来监督训练过程的方法,简称为 JLS-DL。本文方法的主要贡献包括:1)采用角度余量损失以改善同类工件特征之间的距离,完成类内工件的鲁棒性表示;2)使用隔离损失以实现异类工件的良好性区分;3)相较于传统的图像识别方法和单一损失函数的深度学习模型,有效降低了网络的参数量和

计算量,并且识别准确率得到显著提高。

1 联合损失函数的神经网络训练方法

本文提出的高频工件识别模型框架(JLS-DL)如图1所示,其中 λ 是正则化参数,用于调节两个损失的权重,将在实验章节中进行讨论。模型的输入为经过预处理后的高频工件图像,图像的大小为 224×224 ,第二个全连接层后输出工件的识别结果。

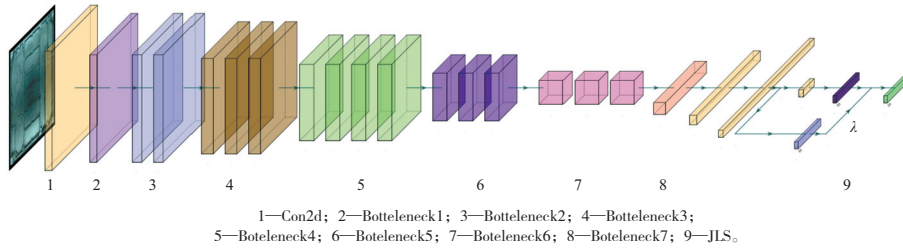


图1 基于联合损失函数的高频工件识别框架(JLS-DL)

工业上零件识别一般采用的是 SoftMax 损失函数。SoftMax 损失函数第一步将预测结果转化到指数函数上,保证了概率的非负性。第二步将转化后的结果除以所有转化后结果之和,通过最大化预测概率,将不同类别区分开。SoftMax 损失函数定义为

$$l_s = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{w_j^T x_i + b_j}}{\sum_{j=1}^C e^{w_j^T x_i + b_j}} \quad (1)$$

式中: x_i 表示神经网络提取的特征向量; y_i 表示特征向量所对应的标签; w_j 表示神经网络全连接层中第 j 列的权重; b_j 表示神经网络中的偏置项; N 表示神经网络的批次大小; C 表示类别个数。

然而,SoftMax 损失函数对类内差异较大的类别效果不好。图2展示了在高频工件数据集上用 SoftMax 损失函数训练网络得到的特征向量经降维后的二维特征分布图,其中高频工件的不同类别用不同的颜色表示。从图2红框1中可以看出,经过 SoftMax 损失函数监督训练之后还是有3类高频工件没有明显地区分开。由红框2和红框3可以看出,同一类高频工件由于摆放位置的随机性和光照等原因,类内距离没有减小。从而不能很好地区分高频工件的类别,造成识别准确率低。针对图2所示问题,本文提出一种由角度余量损失和隔离损失组成的联合损失函数,由联合损失函数监督网络训练,实现最小化类内距离和最大化类间距离(本刊为黑白印刷,如有疑问请咨询作者)。

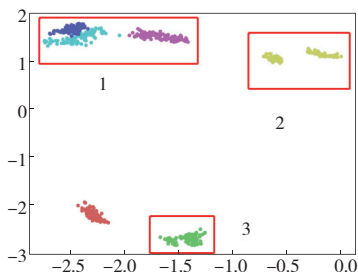


图2 用 SoftMax 训练网络得到的二维特征可视化

1.1 角度余量损失监督同类工件的特征距离

在高频工件的识别任务中,由于工件摆放的角度、光线的变化等原因会造成工件类内差别较大。图3展示高频工件的类别C中不同视角的图像。

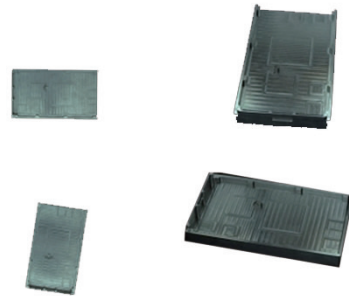


图3 相同类别工件图像差异大

结合图2可知,提高识别准确率的首要措施是减小类内距离。本文采用角度余量损失函数来实现,角度余量损失函数的定义为

$$l_{AC} = -\frac{1}{N} \sum_{i=1}^N \log \left(\frac{e^{s \cdot (\cos \theta_{y_i} - m)}}{e^{s \cdot (\cos \theta_{y_i} - m)} + \sum_{j=1, j \neq y_i}^n e^{s \cdot \cos \theta_j}} \right) \quad (2)$$

式中: m 表示余弦余量,用来将神经网络权重 w_j 和特征向量 x_i 的距离调节为 $\cos \theta_{y_i} - m$; s 表示缩放因子; N 表示卷积神经网络的批次大小。

角度余量损失是在 SoftMax 损失函数的基础上改进而来的,相比较 SoftMax 损失函数,角度余量损失函数通过 m 约束特征向量之间的距离,可以监督训练到更小的类内距离。

1.2 隔离损失监督异类工件的特征距离

图4展示了不同类别工件的微小差异,其中A、B、C三类高频工件主要在图中红框标注处存在微小差异。因

此,使用角度余量损失函数来减少类内距离时,会不可避免地减少类间的距离。所以,本文采用隔离损失函数来增大类间距离,以提高特征向量的表征能力。隔离损失函数定义如下:

$$l_{IL} = \frac{1}{n} \sum_{i=1}^n \| \mathbf{x}_i - c_{y_i} \|^2 + \lambda_1 \sum_{c_j \in A, c_k \in A} \sum_{c_k \neq c_j} \left(\frac{c_k \cdot c_j}{\| c_k \|_2 \| c_j \|_2} + 1 \right) \quad (3)$$

式中: c_{y_i} 表示与第*i*个样本具有相同类别标签的所有样本的中心; A 表示高频工件类别的集合; c_k 表示具有 L_2 范数的 $\| c_k \|_2$ 的第*k*个工件类别中心; c_j 表示具有 L_2 范数的 $\| c_j \|_2$ 的第*j*个工件类别中心; λ_1 表示调节因子,用来调节式(3)中前后项的比例。

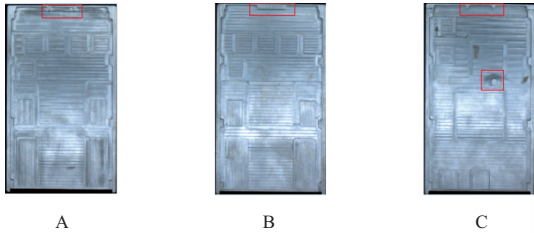


图4 不同类别工件相识度高

式(3)在保证工件类内距离小的同时,通过惩罚工件之间的相似性,增大工件类间的距离。因此,本文网络模型采用结合角度余量损失和隔离损失的联合损失函数来监督训练,减小相同类别工件之间的差异,同时增大不同类别工件之间的差异,从而提高工件识别的准确率。联合损失函数定义如下:

$$l = l_{AC} + \lambda l_{IL} \quad (4)$$

式中: λ 为正则化参数,用于调整角度余量损失和隔离损失的权重。

2 实验结果与分析

2.1 实验配置

本实验使用GPU(显卡型号为NVIDIA GeForce GTX 1660 SUPER)进行训练和测试,训练迭代次数为5 000次,批处理大小为32,初始学习率为0.001,每间隔1 000次迭代,学习率衰减1/2。在Windows 10平台上,采用Pytorch1.2框架来搭建本文提出的JLS-DL高频工件识别模型。

本文所采用的卷积神经网络具体设置如表1所示。图像的输入大小表示为长×宽×通道数,Bottleneck1-Bottleneck7表示7种不同的反残差模块, C_0 表示输出通道数, N 表示在网络中的重复次数,FC表示全连接层; k 表示工件的类别数。

表1中的反残差块如图5所示,包括1)膨胀层,用于扩展维度;2)可分离卷积层,用于提取深度特征;3)投影层,用于压缩数据。

表1 模型的参数设置

输入大小	层名称	C_0	N	步长
224×224×3	Con2d	32	1	2
112×112×32	Bottleneck1	16	1	1
112×112×16	Bottleneck2	24	2	2
56×56×24	Bottleneck3	32	3	2
28×28×32	Bottleneck4	64	4	2
14×14×64	Bottleneck5	96	3	1
14×14×96	Bottleneck6	160	3	2
7×7×96	Bottleneck7	320	1	1
7×7×320	Con2d	1 280	1	1
1×1×1280	FC1	1 280	1	—
1×1×6	FC2	k	1	—

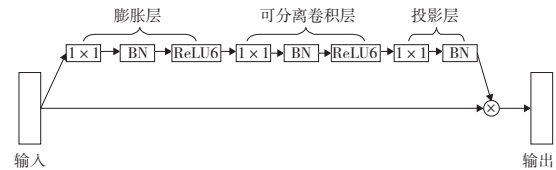


图5 反残差块

在实际场景中采集了3 420张不同类别的工件图像,建立工件数据集,包括36类高频工件。从数据集中随机抽取80%的图像作为训练集,剩余的20%作为测试集。图6展示了部分工件在两个不同视角下的部分实验数据。

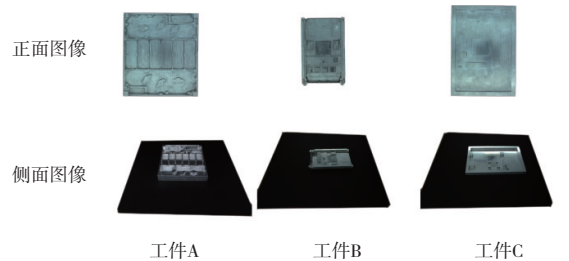
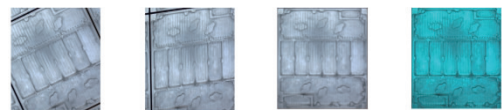


图6 不同视角下的部分实验样本

本文首先采用数据增强,以扩大样本的多样性,可以有效抑制在训练时发生的过拟合,并且有效提高了模型的泛化能力,显著提升模型的鲁棒性。数据增强的方式包括对原始工件图片数据进行随机的强相关性的裁剪、翻转、旋转、缩放和扭曲等变换,部分增强结果如图7所示。



(a) 移位尺度旋转 (b) 弹性变换 (c) 模糊 (d) 随机通道丢失

图7 部分数据增强结果

2.2 识别性能对比实验

为了验证本文提出的JLS-DL模型的识别性能,采用

多种图像分类方法进行对比,包括两种模型驱动的图片识别方法和一种数据驱动的深度学习识别方法。实验结果如表 2 所示,相较于模型驱动的 LBP+SVM 和 HOG+SVM 方法,本文方法对 36 类不同高频工件的识别准确率提升了约 30%;此外,相较于数据驱动的 MobileNetV3 模型^[14],本文的 JLS-DL 方法对高频工件的识别正确率也提升了约 4%。因此,实验结果说明本文提出的 JLS-DL 模型具有更优的性能,能够有效地区分热处理前后不同类别的高频工件。

表 2 不同方法的识别准确率比较

方法	识别准确率/%
LBP+SVM	56.22
HOG+SVM	67.68
MobileNetV3	92.27
JLS-DL	96.33

2.3 损失函数正则化参数 λ 的取值

λ 用于调节两个损失的权重,以优化模型的性能,其取值范围为 0~1。表 3 展示了当 λ 取不同值时,其准确率的变化情况。可以看出, λ 的值对模型的识别性能有较大影响。其中当 $\lambda=0$ 时,即只采用角度余量损失时,模型的识别准确率为 92.36%;当 $\lambda=1$ 时,即只采用隔离损失时,模型的识别准确率为 82.45%;当 $\lambda=0.003$ 时,本文提出的 JLS-DL 模型的识别准确率为 96.33%。本文将 λ 值设置为 0.003。

表 3 λ 取值对识别准确率的影响

λ 取值	识别准确率/%
0.000	92.36
0.001	94.21
0.003	96.33
0.010	93.87
0.030	93.08
0.100	85.66
0.300	84.92
1.000	82.45

2.4 消融实验

为验证本文方法中采用的联合损失的有效性,即采用两项损失相对于单一损失项的性能变化,在本节进行消融实验。将采用角度余量损失的模型命名为 LAC-DL;将采用隔离损失的模型命名为 LIL-DL。LAC-DL 和 LIL-DL 和所提出的 JLS-DL 三种模型在不同迭代次数下的高频工件识别准确率如图 8 所示。从图 8 中可以看出:1) 三种模型都需要一定的迭代次数才能收敛,分别实现各模型的最优识别准确率;2) LAC-DL、LIL-DL 和 JLS-DL 分别收敛于 92.36%、82.45% 和 96.33%。

因此可以得出结论:1) 三种模型对高频工件的识别都有一定的有效性;2) 本文提出的 JLS-DL 模型对高频工

件的识别性能最高。这是因为 JLS-DL 模型不仅能够减小小同类别高频工件的类内距离,还能够增大不同类别高频工件的类间距离。

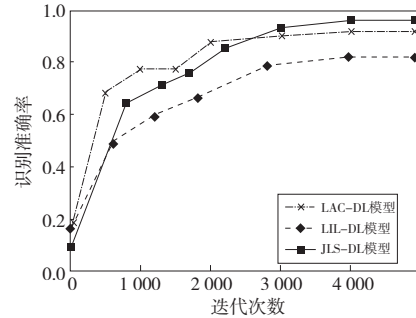


图 8 准确率随迭代次数变化图

为进一步说明本文所提出模型对各类高频工件特征向量的区分性,在图 9 中对 JLS-DL 学习到的特征向量进行降维展示。初始特征向量在未学习时,在特征空间中具有很大的重叠,很难区分;随着模型的优化,学习到的特征向量逐渐具有区分性;当模型收敛时,对各类高频工件最终学习到的特征向量具有明显的代表性,能很好地区分各类高频工件。

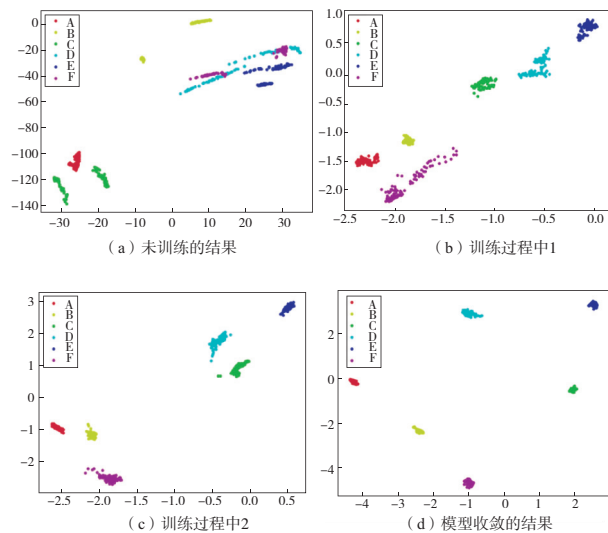


图 9 工件的特征向量

3 结语

针对传统高频工件识别方法和单一损失深度学习方法鲁棒性差、工件识别准确率低的问题,提出了结合角度余量损失和隔离损失的 JLS-DL 高频工件识别模型。模型采用角度余量损失函数减小工件类内距离,克服了由于工件类内距离较大造成的错误识别;同时利用隔离损失函数在控制类内距离的基础上增大类间距离,提升模型的识别准确率。实验结果表明:本文方法能很好地收敛并且可以取得 96.33% 的识别准确率,相比于单一的角度余量和隔离损失,识别准确率分别提高了 3.97% 和 13.88%,验证了本文所提出的联合损失函数的有效性。

(下转第 47 页)