

DOI:10.19344/j. cnki. issn1671-5276.2024.05.048

基于改进卷积网络的铸件视觉检测

颜梦玫^{1,3}, 杨冬平^{2,3}

(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:为实现对铸件各种不同缺陷的自动诊断,设计一种视觉检测系统对铸造产品进行外观检测。利用多头自注意力模块对图像识别网络的 downsampling 过程进行增强,使得经典图像分类网络能够获得全局特征信息,以提升模型对铸件体外观上细小裂纹和大范围毛刺的识别能力。实验结果表明:改进得到的卷积网络对铸件缺陷具有较高的识别准确率,能够较准确地判断铸件缺陷类型。

关键词:缺陷检测;图像识别;卷积神经网络;深度学习;自注意力机制

中图分类号:TP391.4 **文献标志码:**A **文章编号:**1671-5276(2024)05-0229-05

Casting Visual Detection Based on improved Convolutional Network

YAN Mengmei^{1,3}, YANG Dongping^{2,3}

(1. School of Advanced Manufacturing, Fuzhou University, Jinjiang 362251, China; 2. Zhijiang Lab of Research Center

for Human-Machine Augmented Intelligence, Research Institute of Artificial Intelligence, Hangzhou 311101, China;

3. Haixi Institute of Quanzhou Equipment Manufacturing Reserch Center, Chinese Academy of Sciences, Jinjiang 362216, China)

Abstract: For automatical diagnosing different defects in castings, a visual inspection system is designed to inspect the appearance of casting products. The multi-head self-attention module is used to enhance the downsampling process of the image recognition network, enabling the classic image classification network to obtain global feature information and promote the model's ability to identify small cracks and large-scale burrs on the appearance of castings. The experimental results show that the improved convolutional network has a higher recognition accuracy for casting defects, and can more accurately determine the type of casting defects.

Keywords: defect detection; image recognition; convolutional neural network; deep learning; self-attention mechanism

0 引言

铸件是用各种铸造方法获得的金属或者合金材料成型零件,常用作机械、建筑、航空航天和汽车等领域的主要零部件。由于铸件的特殊成型方式,在制造过程中常出现气孔、裂纹和夹渣等质量问题。这些问题不仅影响机械零件的性能,还可能导致它的使用寿命大大缩短。为及时发现铸件的缺陷,基于 X 射线的无损检测^[1]广泛应用于铸件检测中,能够准确诊断铸件内部缺陷。但是由于 X 射线的无损检测效率较低、成本较高,通常会采用外观检查等人工方式进行初步筛选。为了高效、准确地筛选出有缺陷的铸件,可通过图像分类网络对所有铸件进行分类,实现对气孔、细小裂纹和毛刺等缺陷的初步筛选。相比人工外观检查

等低效的筛选方式,图像检测系统^[2]能够更加高效、准确地检测铸件的外观缺陷。

随着计算机视觉技术和深度学习方法的快速发展,通过图像数据对物体进行准确分类已成为可能。2012 年 ILSVRC 图像分类大赛中,深度卷积结构实现的 AlexNet^[3] 一举夺冠,2014 年 GoogLeNet^[4] 实现了 74.8% 的 top-1 准确率,并且 VGGNet^[5] 也实现了相同的精度。此外,ResNet^[6] 提出的残差连接使得训练这些极深的网络更加容易,表现也更好。

近几年来,研究者们利用不同的改进方法来提升图像分类网络的准确率。Transformer^[7] 是一种与卷积结构不同的模型,最开始用于解决自然语言处理(NLP)任务,因其表现出的卓越性能而迅速成为主流架构。ViT^[8] 是视觉任务中应用

基金项目:国家自然科学基金项目(12175242)

第一作者简介:颜梦玫(1997—),女,四川泸州人,硕士研究生,研究方向为图像识别和语义分割,yanmengmei@qq.com。

2) 为提升深度卷积网络对细小裂纹的识别准确率,将 Conformer 中的方法加以简化和改进,提出增强 Transformer 结构,将该结构插入 ResNet 的下采样模块中能够显著提升准确率;并对此方法进行对比试验,以找到最佳融合方式。

• 230 •

特征图 X 依次经过多头自注意力层 (multi head self-attention, MHSA)、layer normalization (LN) 层和多层感知机 (multi layer perceptron, MLP), 计算过程如式 (1) 所示。

$$X_T = \text{MLP}(\text{LN}(\text{MHSA}(X))) \quad (1)$$

多头自注意力层中的特征图 X 分别生成 query、key 和 value, 再经过计算得到同维度的输出结果, 计算过程如式 (2) 所示。这个过程中, 特征图上的全局特征会进行信息交互, 并对感兴趣的特征进行加权, 从而捕捉到特征图上的全局关键信息。

$$\text{Attention}(Q, K, V) = \text{softmax}(QK^T)V \quad (2)$$

特征图 X_T 重新变形重组为 $B \times C \times H \times W$ 的特征图, 再经过一个 3×3 卷积层后和主干特征图相加。这个卷积层的作用是将特征图 X_T 的维度变为和主干特征图的维度一致。整个计算过程可以由式 (3) 表示, 其中 $C(\cdot)$ 表示卷积操作, X_l 和 X_{l-1} 分别表示当前卷积层的输出特征图和前一个卷积层的输出特征图, Y 表示双分支的合并结果。

$$Y = X_T + C_{1 \times 1}(X_l + C_{1 \times 1}(X_{l-1})) \quad (3)$$

通过这种双分支的方式, 网络模型在第 4 个 Stage 能够感受到图片上的局部特征和全局特征, 因此对大范围的毛刺和小尺寸的裂纹都能有较好的识别效果。并且, 增强 Transformer 结构只在低分辨率阶段与卷积进行融合, 不会导致计算速度大幅下降。实验结果表明: 增强 Transformer 带来准确率提升的同时, 并没有大幅度降低吞吐率, 这种方法可以为图像缺陷诊断系统带来更好的检测性能。

2 实验

2.1 实验设置

实验中每个网络均使用 SGD 优化器进行训练, 优化器的 weight decay 和 momentum 分别设置为 0.000 1 和 0.9。在训练过程中, Batchsize 设置为 128, 初始学习率固定为 0.1, 随着实验进行, 学习率逐渐衰退。

在各种网络的训练过程中, 对所有训练图像运用相同的数据增强方法。这里用到的数据增强方法为随机裁剪和随机水平翻转。所有实验代码基于 pytorch 框架实现, 训练和测试过程在多张 NVIDIA GeForce RTX 3090 上采用并行计算完成。

2.2 对比实验

1) 诊断模型对比实验

为了让诊断模型能够准确判断出铸件缺陷的

类型, 以便后期能针对性地加以改进, 需制作包含不同类型缺陷的训练数据集。从各相关数据集中挑选各种不良铸造轴承图片, 将其铸造缺陷类别主要分为孔状缺陷、毛刺和细微划痕等 3 类, 加上正常铸件类别组成一共 4 个类别的数据集。其中, 每个类别的训练图片为 190 张, 验证图片都是 10 张, 另外准备少量跨域场景图片对诊断模型进行泛化能力测试。

采用 2.1 节中的实验参数设置, 将 ResNet、VGG、ResNeXt^[11]、DenseNet^[12] 等主流分类网络和这里的 ResNet (ET) 在铸件诊断训练数据集上进行训练和验证, 其中 BoT 为 CNN 和 Transformer 结合网络, 主要特点为精度和计算速度表现都较好, 实验结果如表 1 所示。相比 ResNet 和 VGG 等分类网络, 具有 ET 结构的 ResNet 网络实现了更高的准确率和更快的计算速度, 表明 ET 结构可以作为缺陷检测系统的增强方法, 能使模型准确、快速地对铸件缺陷类型进行分类判断。

表 1 铸件诊断模型准确率对比

模型	验证准确率/%	帧率/(帧/s)
ResNet-50	96.89	38
VGG	96.77	35
ResNeXt	96.93	33
DenseNet	97.38	29
BoT	95.20	36
ResNet-50 (ET)	97.50	40

此外, 为了验证本文方法对细小划痕缺陷识别的增强效果, 随机选择包含细小划痕的铸件图片进行验证和分析。将缺陷铸件图片输入改进诊断模型计算, 并得到计算时的热力图, 如图 3 所示。图 3(a) 为原图, 图 3(b) 是网络发现毛刺和细小气孔缺陷区域的热力图, 这正说明它依赖于这些特征进行缺陷分类。

2) 公开数据集对比实验

为进一步验证改进方法对分类网络的有效性, 让 ResNet、VGG、ResNeXt、DenseNet 等分类网络和 ResNet (ET) 分别在公开分类数据集上进行对比实验。实验数据集主要包括: CIFAR^[13] 和 Tiny-ImageNet。训练参数与 2.1 节中的一致, 为了让分类网络适应小分辨率数据的尺寸, 将深度卷积模型都进行相同的修改, 将下采样次数由原本的 5 次减少为 3 次。

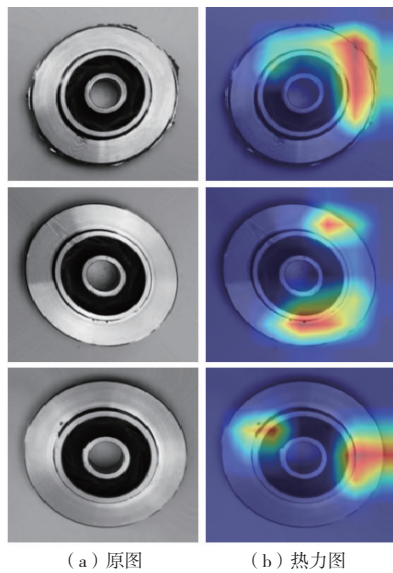


图3 热力图

实验结果如表2所示,提出的 ResNet-50 (ET)在 CIFAR10、CIFAR100 和 Tiny-ImageNet 上分别比 ResNet-50 高 0.52 个百分点、0.85 个百分点和 0.02 个百分点。并且在一系列公开数据集中,ResNet-101 (ET) 也比 ResNet-101 表现出更高的准确率。此外, DenseNet 在 CIFAR10 和 CIFAR100 两个数据集上的表现都稍优于 ResNet (ET),但是在 Tiny-ImageNet 数据集上的准确率表现较差,并且从表1可以看出它的计算速度较慢。

表2 分别在 CIFAR 和 Tiny-ImageNet 上实验的准确率
单位: %

模型	CIFAR10	CIFAR100	Tiny-ImageNet
ResNet-50	96.06	77.10	66.25
ResNeXt	96.41	77.29	67.10
VGG16	96.25	76.85	67.13
DenseNet	96.82	79.82	65.61
ResNet-50(ET)	96.58	77.95	66.27
ResNet-101	97.60	78.75	67.35
ResNet-101(ET)	97.65	78.80	67.41

改进方法能够提升分类网络的准确率,但是也会引入更多参数量和计算量,因此只在第4个 Stage 使用改进方法。为了对比改进方法在网络不同层插入带来的改变,将改进方法分别插入 ResNet-50 第4个 Stage 的第1、第2、第3个 Bottleneck 中,并在 CIFAR100 和 Tiny-ImageNet

上训练和验证。实验结果如表3所示。增强 Transformer 的位置放在下采样 Bottleneck 的前面能得到较好的准确率,并且插入位置越靠近顶部,模型的准确率会越低。

表3 插入位置对比表 单位: %

模型	Bottleneck	Tiny-ImageNet	CIFAR100
ResNet (ET)	1	66.27	77.95
	2	66.13	77.44
	3	65.88	75.66

3 结语

通过在 ResNet 下采样过程中添加增强 Transformer 得到改进模型,利用图像识别对铸件外观缺陷检测。实验表明改进后的图像识别模型能够更精确地对缺陷图像进行分类。在深度卷积结构中插入 Transformer 模块,使得网络可以进行局部或全局特征信息交互,从而增强特征信息,对大范围毛刺和细小划痕缺陷都能准确识别。ResNet(ET) 不仅能够在图像识别任务中获得更高的准确率,同时它的计算速度并未下降,ResNet (ET) 相比同类型的网络架构表现出更好的泛化性能。

参考文献:

- [1] 张国宝,杨为,赵恒阳,等. 基于 X 射线三维成像技术的在役 GIS 盆式绝缘子缺陷检测[J]. 高压电器, 2022,58(10):230-236.
- [2] 周祺智,马万太. 基于局部分类的铝合金低倍组织图像分割[J]. 机械制造与自动化, 2022, 51(4): 129-132.
- [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]. (2015-04-10)[2022-10-20]. <http://arxiv.org/abs/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] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all You need [EB/OL]. (2017-06-12) [2022-10-20]. <https://arxiv.org/abs/1706.03762>.
- [8] 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) [2022-10-20]. <https://arxiv.org/abs/2010.11929>.
- [9] SRINIVAS A, LIN T Y, PARMAR N, et al. Bottleneck transformers for visual recognition [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE, 2021:16514-16524.
- [10] PENG Z L, HUANG W, GU S Z, et al. Conformer: local features coupling global representations for visual recognition [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, QC, Canada: IEEE, 2022:357-366.
- [11] XIE S N, GIRSHICK R, DOLLÁR P, et al. Aggregated residual transformations for deep neural networks [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017:5987-5995.
- [12] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017:2261-2269.
- [13] KRIZHEVSKY A. Learning multiple layers of features from tiny images [R]. Toronto, Canada: University of Toronto, 2009.

收稿日期:2023-03-08

(上接第212页)

3 结语

本文介绍了气动载荷下航空发动机机匣表面传感器布置方法,并基于某单转子整机试验器进行了应用分析。建立了气动载荷的表征模型来计算整机模型机匣表面的响应,并基于响应计算结果进行响应关于气动载荷灵敏度的分析,再结合有效独立法,综合考虑了灵敏度与独立性。最后在单转子整机试验器上选择了机匣表面的两个最佳测点,并与对照组进行了响应对比。结果表明,最佳测点的响应明显高于对照组,证明了该方法的有效性。

参考文献:

- [1] 吴伟力,陈大力. 航空发动机振动故障诊断技术及发展趋势[J]. 航空动力,2021(3):69-72.
- [2] 蔡肇云,金六周. 航空发动机强度设计、试验手册:叶片强度与振动计算[M]. 北京:第三机械工业部第六研究院,1980:34-50.
- [3] MAILACH R, LEHMANN I, VOGELER K. Periodical unsteady flow within a rotor blade row of an axial compressor—part II: wake - tip clearance vortex interaction[J]. Journal of Turbomachinery, 2008, 130(4): 1587-1597.
- [4] STUMMANN S, JESCHKE P, METZLER T. Circumferentially non-uniform flow in the rear stage of a multistage compressor[C]//Proceedings of ASME Turbo Expo 2015: Turbine Technical Conference and Exposition, Montreal, Quebec, Canada: [s. n.], 2015: 42935-42944.
- [5] 敬彤,臧朝平,张涛,等. 复杂时变激励下失谐叶盘瞬态强迫响应分析[J]. 航空学报, 2021, 42(9): 310-321.
- [6] KAMMER D C. Sensor placement for on-orbit modal identification and correlation of large space structures[J]. Journal of Guidance, Control, and Dynamics, 1991, 14(2):251-259.
- [7] 刘渊,臧朝平,刘忠华. 搜索算法对有效独立法优化振动模态测量位置的影响[J]. 航空动力学报,2012, 27(2):312-318
- [8] 梁昊天. 航空发动机缩尺动力学相似整机试验器设计与建模研究[D]. 南京:南京航空航天大学,2020.
- [9] 蒋锂军,臧朝平,张磊,等. 航空发动机机匣结构减缩建模的方法[J]. 机械制造与自动化,2023,52(1): 119-122.
- [10] 谭元球,臧朝平,周标,等. 失谐叶盘受迫响应的灵敏度分析方法[J]. 航空学报,2017,38(12):192-202.

收稿日期:2023-03-15