

基于卷积神经网络的焊接装配特征识别研究

陈建强 秦娜

西南交通大学电气工程学院 成都 610031

西南交通大学系统科学与技术研究所 成都 610031

(cjqsyghr@my.swjtu.edu.cn)

摘要 为实现高铁白车身焊接拼装技术的智能化与自动化,解决焊接过程中特征区域小、背景干扰多等问题,提出了基于迁移学习和卷积神经网络的焊接装配特征快速识别算法。首先采用二值化等传统图像处理算法确定待提取特征的粗略位置,在此基础上再使用 sobel、腐蚀、霍夫线段检测确定特征区域的精确位置。其次,考虑到不同环境下,精确定位后特征区域表现不同,故采用基于卷积神经网络的分类模型以增强预测模型的鲁棒性和准确性。最后,选择基于迁移学习的视觉几何群网络(VGG16)来解决样本量不足以训练整个模型参数的问题。实验结果表明,本文所提的识别算法能够准确识别型材的状态,且在识别检测速度上优于 YOLOV3,在准确率上劣于 YOLOV3,算法满足使用场景下的实时性要求。

关键词: 迁移学习;卷积神经网络;特征快速识别;霍夫线段检测;视觉几何群网络(VGG16)

中图分类号 TP391.4

Recognition Algorithm of Welding Assembly Characteristics Based on Convolutional Neural Network

CHEN Jian-qiang and QIN Na

School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, China

Institute of Systems Science and Technology, Southwest Jiaotong University, Chengdu 610031, China

Abstract In order to realize the intellectualization and automation of welding and assembling technology for high-speed white body, the problems of small feature area and multi-background interference in welding process are solved, a novel fast recognition algorithm of welding assembly based on migration learning and convolution neural network is proposed. Firstly, the traditional image processing algorithms such as binarization are used to determine the rough position of the feature to be extracted. On this basis, Sobel, corrosion and Hough line detection are used to determine the precise position of the feature area. Secondly, considering the different performance of feature regions in different environments, a classification model based on convolution neural network is adopted to enhance the robustness and accuracy of the prediction model. At last, Visual Geometry Group Network (VGG16) based on transfer learning is selected to solve the problem that the number of the samples is not enough to train the parameters of the whole model. The experimental results show that the recognition algorithm proposed in this paper can accurately identify the state of profile, and the detection speed is better than YOLOV3, and the accuracy is inferior to YOLOV3. The algorithm can meet the real-time requirements in the use scene.

Keywords Transfer learning, Convolution neural network (CNN), Fast feature recognition, Hough line segment detection, Visual geometry group network (VGG16)

1 序言

高铁白车身侧墙板和枕梁等典型构件在焊接前需要判断两块型材是否匹配,焊接不匹配的型材将对整个焊接过程甚至是整条生产线造成重大影响。目前型材焊接前的检查依赖于人工检查,该方式耗时长、效率低,且人力成本高。为了实现高铁白车身焊接拼装技术的智能化与自动化,对侧墙、枕梁典型构件自动、准确地完成焊接前的识别是自动化生产非常关键的一步。

图像识别技术是利用视觉传感器获取物理对象描述信

息,通过分析数字信息使视觉传感器能够模仿人类视觉,自动处理获取的数字信息^[1]。分析数字信息的方法可分为传统图像处理和深度学习。传统图像处理分析图像的步骤为数据特征分析、图像预处理、特征提取、模式识别,处理过程中采用人工设计的特征,在条件复杂、对对象属性不了解的情况下,模式识别输入的特征不能有效地表示物体独特的属性,因而识别效果较差。卷积神经网络^[2-3]把特征提取和模式识别紧密结合在一起,从低层的纹理特征、几何特征到深层的语义特征,通过损失函数^[4]优化模型,得到图像向任务目标的映射机制,并且该机制的泛化能力极其强大。训练卷积神经网络需

基金项目:2017年度国家重点研发计划“智能机器人”重点专项(2017YFB1303402,2017YFB1303402-03);国家自然科学基金项目(61603316,61773323);四川省科技计划(2019YJ0210,2019YFG0345)

This work was supported by the 2017 National Key R&D Program “Intelligent Robot” Key Special Project(2017YFB1303402,2017YFB1303402-03), National Natural Science Foundation of China (61603316, 61773323) and Sichuan Science and Technology Plan (2019YJ0210, 2019YFG0345).

通信作者:秦娜(qinna@swjtu.cn)

要大量的数据集,但是焊接装配图的数量有限,难以训练整个模型。迁移学习^[5]理论实现了将深度学习应用于小样本对象中,在目标识别和检测领域有较多成功应用的例子,为解决本文数据量少的问题提供了一个切实可行的方法。

本文首先在型材焊接前对图像进行预处理,其中主要研究如何定位用来分类的特征区域,并采用了一系列的传统图像处理方法,如二值化、sobel^[6]算子等,得到了少量线段的二值图,便于霍夫变换检测直线^[7-8]定位到特征区域坐标。在原图上截取特征区域,得到的局部图输入卷积神经网络中的训练模型,最后通过对比实验验证本文算法在焊接装配特征识别任务中的优越性。

2 焊接装配特征快速识别方法

本文采用的智能化高铁白车身装配特征检测装置如图 1 所示,面向的待检测的装配特征如图 2 所示,本文提出的基于传统图像处理和深度学习技术的焊接装配特征快速识别方法流程图如图 3 所示。从图 3 可以看出,测量方法分为两个阶段:训练阶段和测试阶段。在训练阶段,使用大量自动截取的特征图片作为数据集训练卷积神经网络。测试时,对测试图片进行预处理操作,定位特征区域,将该区域输入到已经训练好的网络中,得到识别结果。

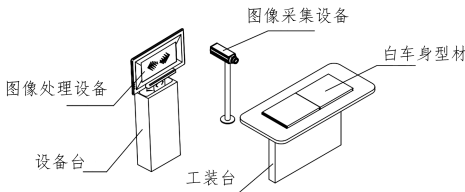


图 1 智能化高铁白车身装配质量检测装置

Fig. 1 Intelligent high-speed rail body-in-white assembly quality detecting device

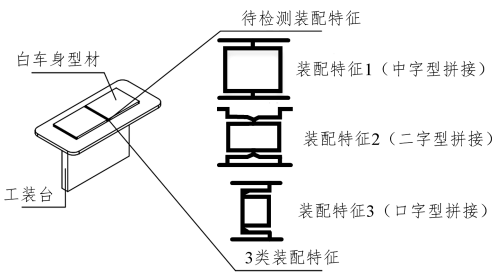


图 2 3类待检测的型材拼接状态

Fig. 2 Three types of profile splicing states to be tested

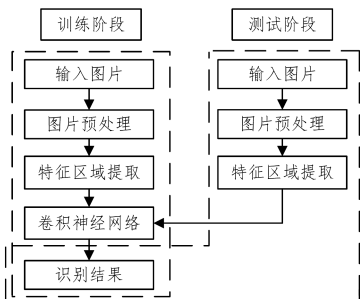


图 3 基于传统图像处理和深度学习技术的焊接装配特征快速识别方法流程图

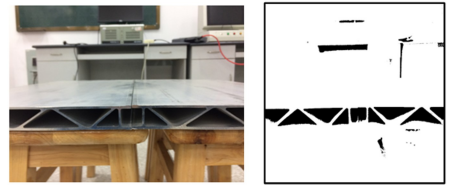
Fig. 3 Flow chart of rapid identification method for welding assembly features based on traditional image processing and deep learning techniques

2.1 图片预处理

2.1.1 图像缩放、二值化

本文获取的图像尺寸为 $3264 * 2448$ 。为了减小计算量,加快算法运行速度,将图片的尺寸缩小到 $800 * 800$ 。由于型材区域存在大量像素值较低的点,为了在图片上突出型材,对图片进行二值化处理。式(1)表示对第 x 行 y 列的像素点二值化, $f(x, y)$ 为该点的像素值,式(1)中的 40 是根据大量图片处理结果得到的,图片处理结果如图 4 所示。

$$f(x, y) = \begin{cases} 0, & f(x, y) \leq 40 \\ 255, & f(x, y) > 40 \end{cases} \quad (1)$$



(a)输入图像

(b)二值化后图片

图 4 输入图像二值化处理结果

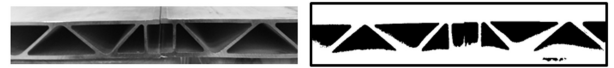
Fig. 4 Input image binarization result

2.1.2 初步定位特征区域

由于型材区域为全图的局部区域,为了精确定位特征区域,减少无关区域的干扰以及减少计算量,初步定位特征区域,并截取该区域。

1)从二值图像最后一行从下往上遍历图像,获取每行像素值为 0 的像素点个数,使用大量图片进行统计,当像素数大于 400,得到型材底部行坐标。

2)继续遍历图像,当某行的像素数小于 400,得到型材顶部行坐标。初步截取图像如图 5 所示。



(a)初步截取灰度图

(b)二值图像截取图

图 5 初步截取图像

Fig. 5 Preliminary capture of images

2.2 特征区域提取

2.2.1 Sobel 理论

Sobel 算子是线性滤波器,其主要作用是能够分别从 x 、 y 方向检测边缘,或者结合 x 、 y 方向的梯度信息检测边缘^[9]。该算子基于两个方向线性滤波器, x 、 y 方向最简单的滤波器结构如图 6 所示。

-1	0	1
-2	0	2
-1	0	1

(a)x 方向滤波器

-1	-2	-1
0	0	0
1	2	1

(b)y 方向滤波器

图 6 x 、 y 方向滤波器结构

Fig. 6 x 、 y direction filter structure

较常使用的滤波器为 $3 * 3$ 、 $5 * 5$ 、 $7 * 7$,每个尺寸下的两个滤波器分别在 x 和 y 两个方向求图像的梯度变化率,其中 X 为需要处理的图像, G 为某个方向上的梯度变化率图, $*$ 为二维卷积操作, W_1 、 W_2 分别为 x 、 y 方向上的滤波器^[9]。

1)水平梯度变化率图:

$$G_x = X * W_1 \quad (2)$$

2)垂直梯度变化率图:

$$G_y = X * W_2 \quad (3)$$

2.2.2 算法步骤

二值化图像使用 sobel 算子对图像的竖直线进行锐化,采用的 sobel 算子的内核为 5 * 5。

1)在结果图基础上进行腐蚀操作,腐蚀的像素点在竖直线之外。结果如图 7 所示。



图 7 锐化及腐蚀图像

Fig. 7 Sharpen and corroded images

2)采用霍夫变换检测线段,得到线段坐标精确定位特征区域。图 8 为处理后的结果图。

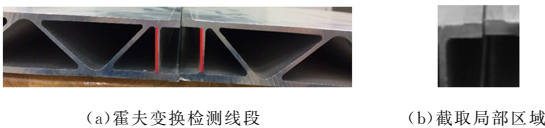


图 8 霍夫检测及截取图像

Fig. 8 Hough detection and interception of images

2.3 特征区域图片训练卷积神经网络

本文采用的卷积神经网络结构如图 9 所示,模型输入图片尺寸为 50 * 50 * 3,视觉几何群网络^[10](Visual Geometry Group Network,VGG16)。

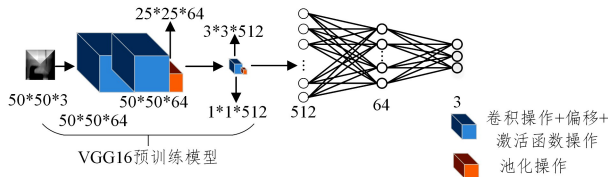


图 9 卷积神经网络结构图

Fig. 9 Convolutional neural network structure

本文卷积神经网络参数多达为 14747715,在不采用已训练好的参数初始化模型的部分卷积层的情况下,训练模型时为了避免过拟合,数据集原则上需要大于参数数量。在现实条件下获取如此大的数据集耗时长、成本高,而各种分类任务中采用的卷积神经网络的特点为:网络浅层提取的特征为基础纹理、几何特征,深层提取的特征为对特定分类任务有效的语义特征,不同分类任务中网络浅层的参数是可以复用的,即采用迁移学习方法,将 VGG16 中已训练好的部分模型参数初始化并将初始化后的层冻结(不训练)。本文选择的已训练模型为 VGG16,VGG16 由 Karen Simonyan 和 Andrew Zisserman 于 2014 年开发,整个网络已在大型数据集 ImageNets^[11]分类任务上进行训练,VGG16 学到的特征的空间层次结构可以有效地作为视觉世界的通用模型。文本采用的卷积神经网络冻结了除全连接层外的其余层,冻结的和需训练的参数数量分别为 14714688,33027,微调添加 2 层全连接层,得到图像向任务目标的映射机制。

移除全连接层后的模型作为特征提取层,在原 VGG16 的全连接层位置上加 2 层全连接层,卷积神经网络中各层的参数如表 1 所列。

表 1 卷积神经网络参数图

Table 1 Convolutional neural network parameter map

模型结构	层名称	输出尺寸	参数数量
VGG16	输入层	(20,50,50,3)	0
	卷积层 1	(20,50,50,64)	1792

	卷积层 3	(20,3,3,512)	2359808
	池化层 3	(20,1,1,512)	0
微调全连接层	全连接层 1	(20,512)	0
	全连接层 1	(20,64)	32832
	全连接层 2	(20,3)	195

网络的输入是来自输入图像的特征区域。模型中的逐层特征抽象是通过在每层应用多卷积核,然后对每个局部区域响应进行下采样来实现^[12]。第 Q 个卷积核在第 P 个卷积层对输入特征的局部区域卷积后输出为:

$$H(P,Q) = F(\sum_{c=1}^N X_{p,c} * W_{p,c} + b_{p,q}) \quad (4)$$

其中,*表示二维卷积,卷积核通道与卷积层输入的特征图通道为 N,卷积核逐通道卷积,c 为当前卷积通道,遍历所有通道,在每个卷积核得到的响应基础上加上一定的偏移值。F()是非线性激活函数,VGG16 采用的是 ReLU 函数。在对输入特征图所有区域使用全部卷积核得到的特征图采用最大池化下采样处理^[12]。

损失函数刻画了预测向量和期望向量概率分布的距离,卷积神经网络通过不断向损失函数的极小值点逼近,从而得到特征图像到特征类别目标的映射机制,在 2 分类问题上交叉熵是较为常用是损失函数^[13]。p 为期望向量的概率分布,q 为预测向量的概率分布,通过 q 来表示 p 的交叉熵为:

$$H(p,q) = -\sum_x p(x) \log q(x) \quad (5)$$

3 实验结果

本文研究了 3 种不同型材拼接特征区域定位问题,实验结果如图 10 所示,精确截取图(d),(e),(f)作为卷积神经网络第一种输入。

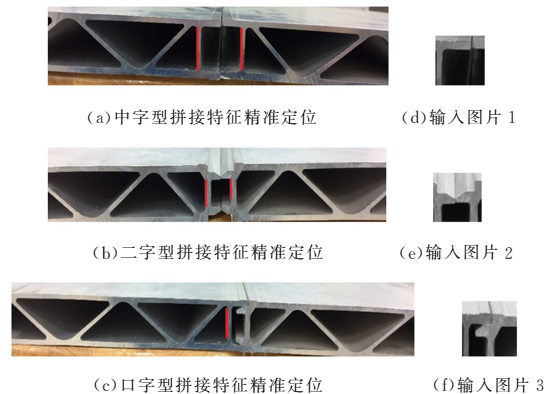


图 10 霍夫检测直线结果及其截取图片

Fig. 10 Hough detects straight line results and their captured pictures

精准截取图数据集分为训练集、验证集和测试集,图片由相机拍摄得到,训练集共有 2426 张图片,其中 1 类、2 类、3 类图片数分别为 724,743,959,验证集共 256 张,其中 1 类、2 类、3 类图片数分别为 78,80,98,测试集共 344 张,其中 1 类、2 类、3 类图片数分别为 110,133,101。一轮 121 次遍历全部样本,总共有 16 轮。

本文提出的算法在处理图片的过程会产生大量的初步截取图,以初步截取图作为模型的第二种输入,如图 11 所示,数据集的数量分配与精确截取图数据集一致。初步截取图与精确截取图分别作为模型的输入训练网络,最后得到 2 个模型,对比 2 个模型的性能,选择最佳的输入。

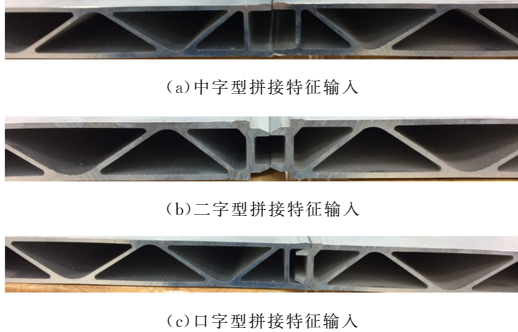
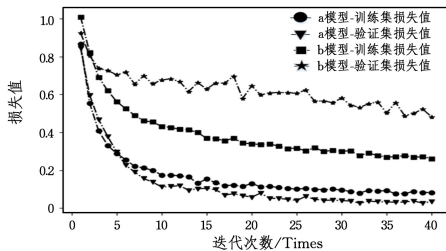


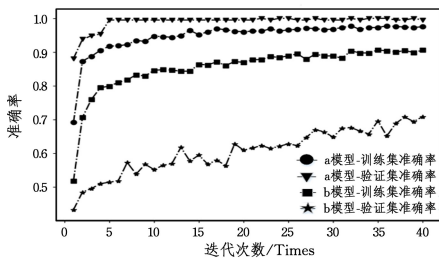
图 11 第二种输入图片

Fig. 11 Second input picture

初步截取图、精准截取图分别作为输入训练模型,训练结果如图 12 所示。



(a) a, b 模型实验训练和验证损失



(b) a, b 模型训练和验证准确率

图 12 精准截取图、初步截取图训练模型损失值与准确率

Fig. 12 Accurate interception map, preliminary interception map training model loss value and accuracy

由图 12 可知,经过一定次数的迭代后, a 模型训练集和验证集的损失值可达到一个较小的数,且两条损失曲线相对接近; b 模型训练集和验证集的损失值和两者差值较大,存在过拟合现象。 a, b 模型在检测准确率上也存在较大差异, a 模型的准确率较高,迭代后期能够达到 98% 以上,两条曲线非常接近,且在验证集上的准确率高于在训练集上的准确率,原因在于为了增强模型的健壮性和泛化性,训练集图片较验证集图片的拍摄条件更加恶劣,更难以辨别;而 b 模型迭代后期的准确率较低,不能达到实际工程应用的要求。

加载训练好的 a, b 模型,预测对应的测试集, a, b 模型预测准确率分别为 99.8%, 80%, a 模型比 b 模型提高了 19.8 个百分点。精确、初步截取图训练的模型效果区别如此大的原因在于初步截取图的无关因素太多,训练时模型容易陷入局部最小值,而精确截取图减少了大部分无关因素,只聚焦于

能够用于分类的特征,由此训练的模型泛化性更加强大。

针对同样的数据集,分别采用本文提出的算法与 YOLOV3^[14] 检测特征。YOLOV3 算法的输入图片如图 13 所示,图片输入尺寸为 400 * 320。两种算法实验结果如表 2 所列。

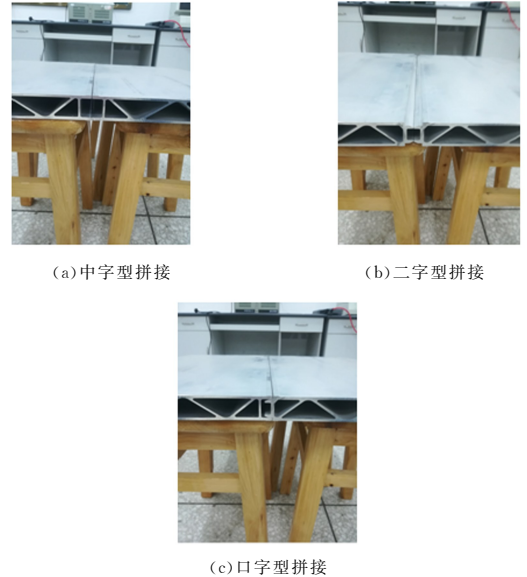


图 13 模板匹配算法第 2 组输入图片

Fig. 13 Template matching algorithm group 2 input picture

表 2 两种算法的试验结果

Table 2 Test results of two algorithms

算法	输入	测试集数量	准确率	每张检测耗时/ms
YOLOV3	图 13	344	0.99	26
本文方法	图 10(a 模型)	344	0.92	10

YOLO 系列算法采用了分而治之、端到端训练、批数据归一化、多尺度训练等模式。而 YOLOV3 在前面系列算法的基础上将卷积、批数据归一、激活函数结合起来作为算法的基本组件,并采用残差网络 resn 作为网络的主干,网络的深度达到了 252 层,算法结构清晰、实时性好。分析实验结果可知,本文提出的算法与 YOLOV3 在准确率和时间上都表现出了巨大的优越性。YOLOV3 层数太多,时间、空间复杂度较高。本文提出的算法采用传统图像处理和深度学习相结合的方式,本文使用的传统图像处理算法的处理速度快于 YOLOV3,在卷积神经网络方面, VGG16 层数较少,且基础组件复杂度比 YOLOV3 低。综合分析可知,本文提出的算法实时性好,但由于算法结构简单,提取的特征有限,因此准确率较低。

结束语 本文主要研究了多干扰环境下高铁白车身智能焊接装配特征快速识别的问题,主要包括初步、精确定位特征区域,特征区域分类 3 部分。首先初步、精确定位减少了无关因素的干扰,且初步定位提高了精确定位的准确性和快速性,同时两步定位为下一步训练和测试卷积神经网络奠定了坚实的基础。由于卷积神经网络端到端的工作方式,目标分类效果较好,因此后续的特征区域分类采用卷积神经网络。数据集不足、训练参数超千万的采用了迁移学习初始化模型,使模型训练过程更快速、准确。在图片处理过程中会产生初步截取图与精准截取图,比较两种不同输入训练得到的模型的性能

(下转第 235 页)