DOI: 10.3901/JME.2019.17.022

基于卷积神经网络的 GTAW 熔透预测

李海超 刘景风 谢吉兵 王 昕

(哈尔滨工业大学先进焊接与连接国家重点实验室 哈尔滨 150001)

摘要: 焊接熔池信息能反映熔透状态,但建立熔池与熔透状态的关系十分困难。针对该问题,提出一种基于卷积神经网络(Convolution neural networks, CNN)的 GTAW 背面熔透预测模型。通过基于被动视觉的熔池二维图像采集,建立了用于 CNN 网络训练和测试的训练集和测试集。其次建立了 CNN 背面熔透预测模型,优化了学习率、batch-size、迭代次数等网络参数。研究发现:第一层卷积核尺寸为 9×9 ,最后一层含有 64 个卷积核使模型在预测准确率和训练时间上综合表现达到最佳。使用训练集数据训练模型,将训练完成的模型在测试集上对背面熔透进行预测,取得了高于 96% 的预测准确率。通过对预测模型的特征映射进行可视化分析,模型是通过熔池边缘,反光点位置和熔池尾部等特征来预测背面熔透情况。

关键词: 卷积神经网络; GTAW; 焊接熔透预测

中图分类号: TP39

GTAW Penetration Prediction Model Based on Convolution Neural Network Algorithm

LI Haichao LIU Jingfeng XIE Jibing WANG Xin

(State Key Laboratory of Advanced Welding and Joining, Harbin Institute of Technology, Harbin 150001)

Abstract: The penetration state can be reflected by the information of the molten pool, but it is difficult to establish a function between the molten pool and the penetration state. To solve this problem, a penetration prediction model based on convolution neural network (CNN) is proposed. Based on the introduction of CNN principle, a molten pool sensing system based on passive vision is designed to collect 2D images of the molten pool. The acquired images are preprocessed to generate the training set and test set for CNN training and testing. Then the prediction mode is built, and the network parameters such as learning rate, batch-size and iterations are optimized. It is found that the model can achieve the best comprehensive performance in prediction accuracy and training time when the size of the first layer convolutional kernel is 9×9 and the last layer contained 64 convolutional kernels. After training the model with the training set, the trained model is used to predict the penetration state on the test set, and the prediction accuracy is higher than 96%. By visualizing the feature mapping of the prediction model, it is found that the model predicts the penetration state through the features of edge, the position of reflective point and the molten pool tail, thus explaining how the model judges the penetration state.

Key words: convolutional neural networks; GTAW; welding penetration prediction

0 前言

熔透控制是实现焊接质量控制和焊接过程自动化的焦点和难点,而实现熔透控制关键在于如何有效地获取表征焊接熔透状态的信号^[1]。熔池图像包含着最为丰富的信息,能够反映焊接过程的稳定性、焊接熔透状态等情况^[2]。电弧焊中,一个熟练

的焊工通过观察液体熔池表面进行判断,进而调整焊接参数,控制焊接过程达到预定的熔透状态,从而得到合适的焊缝形状和焊接质量^[3-4]。如果通过视觉检测系统模拟人工操作,通过正面熔池检测预测熔透状态,进而实现闭环控制,具有很大的实用价值^[5]。

人们在将人工神经网络用于实际熔透预测研究中进行了许多探索。肯塔基大学的 KOVACEVIC 等^[6]基于神经网络建立了熔池尺寸信息与背面熔宽的模型。山东大学高进强^[7]使用普通 CCD 摄像机拍

摄出 GTAW 熔池图像, 并提取熔池宽度、熔池半长、熔池后部面积和熔池后拖角等参数, 基于神经网络建立正面熔池几何形状参数与背面熔宽的关系模型。刘新锋等^[8-9]进一步改进 CCD 摄像机镜头, 拍摄出完整的 GTAW 焊熔池图像, 提取熔池长度、熔池宽度、熔池面积、熔池后拖角作为神经网络的输入, 背面熔宽作为输出, 建立的神经网络模型能够精确预测熔池背面熔宽的变化。以上研究只能建立熔池特征参数与熔透状态的关系, 并不能充分利用熔池内包含的丰富信息。

卷积神经网络(Convolution neural network, CNN)是一种前馈神经网络, 它由一个或多个卷积层和顶端的全连通层组成, 同时也包括关联权重和池化层。这一结构使得卷积神经网络能够利用输入数据的内在结构。与其他深度学习结构相比, CNN 在图像分类, 目标识别等方面能够给出更好的结果^[10]。

因此, 本文以 GTAW 正面熔池图像为对象, 建立了基于卷积神经网络的背面熔透预测模型。利用获取的熔池正面图像训练了卷积神经网络并验证了背面熔透准确率。通过试验找到了最优的卷积核数量和尺寸并达到了最佳熔透预测准确率。

1 卷积神经网络结构

1.1 卷积层

卷积层使用卷积核(kernel)在输入图像上通过滑动窗口计算而得, 卷积核中的每一个参数都相当于传统神经网络中的权值参数, 与对应的局部像素相连接, 将卷积核的各个参数与对应的局部像素值相乘之和即可得到卷积层上的结果。卷积层的最大优势在于权值共享, 因为卷积核在输入图像上滑动时可以判断局部的特征而忽略其余部分的信息。同时, 权值共享还可以减少训练参数, 降低对于硬件计算能力的要求。卷积运算的实现如式(1)所示

$$(I \times K)_{xy} = \sum_{i=1}^h \sum_{j=1}^w K_{ij} I_{x+1-i, y+j-1} \quad (1)$$

式中 I ——输入图像;

K ——卷积核, 大小为 $h \times w$;

x ——输出矩阵的横坐标;

y ——输出矩阵的纵坐标。

卷积计算的示意图如图 1 所示。卷积核与原图像中相同大小的一块区域的灰度值相乘, 然后滑动 1 个像素, 继续进行以上操作直至完成对整张图像的扫描。

图 1 卷积运算示意图

1.2 池化层

池化操作(Pooling)是一种降采样操作, 主要目的是在降低图片尺寸的同时保存了图片的纹理特征等关键信息。其中, 最大池化是最为常用的池化方法。最大池化的具体算法如图 2 所示。它从相邻的 2×2 窗口中选择最大的像素值代表整个窗口的特征。本文所采用的降采样方法即为最大池化方法。

图 2 最大池化算法

1.3 全连接层

在 CNN 中, 全连接层的作用是接收前面的卷积层和池化层输出的特征矢量并根据这些特征矢量进行分类。全连接层的输出维度与分类类别的个数相同, 本章的输出结果分三类, 因此全连接层共有三个输出神经元。将输入神经元与输出神经元进行全连接即形成了全连接层。同时, 全连接层后面往往连接着一个 softmax 层, softmax 层具有将输入归一化的功能, 即将输入转化为概率分布, 从而有利于损失函数的建立。softmax 函数的表达式为

$$S_i = \frac{e^i}{\sum_j e^j} \quad (2)$$

式中 S_i ——第 i 个属于元素的 softmax 函数输出。

全连接层和之后的 softmax 层的示意图如图 3 所示。

图 3 全连接层和 softmax 层示意图

1.4 损失函数与反向传播算法

损失函数(Loss Function)是评判预测结果与真

实值偏差程度的函数。其中交叉熵损失函数(cross-entropy loss function)使用较为广泛,这也是本节使用的损失函数,它的表达形式为

$$C = -\frac{1}{n} \sum_{i=1}^n [y \ln a_i + (1-y) \ln(1-a_i)] \quad (3)$$

式中 C ——损失函数的值;

n ——神经元的个数;

y ——期望的输出;

a_i ——第 i 个神经元的实际输出。

式(3)对 a_i 求导,得到

$$C' = -\frac{1}{n} \frac{y-a_i}{a_i(1-a_i)} \quad (4)$$

由式(4)可以看出,神经元 a_i 输出越接近,损失函数越小,当 $a_i = y$ 时,原函数取得最小值。因此损失函数可以表征模型的预测准确程度。神经网络学习的目的就是损失函数的值减小,这一过程是通过反向传播算法实现的。具体而言,训练进行时首先通过正向传播计算训练数据的响应,然后将响应值代入损失函数,最后从后向前逐层计算关于权重的梯度值,并让原始权重值减掉梯度值和学习速率的乘积,从而完成一次权重系数的更新。神经网络通过重复以上权值更新的过程来提升表现。

2 数据采集

2.1 熔池视觉传感系统

熔池图像采集系统主要由微视公司生产的 RS-A1003 GM50 CMOS 相机、日本 RICOH 公司生产的 FL-HC1214-2M 工业相机镜头、Lincoln 公司生产的 INVERTEC V270-T 脉冲 TIG 焊机以及三轴运动机构构成。依据图 4 的结构示意图组建了熔池图像采集系统。相机与工件的夹角约为 40° ,镜头到钨极的距离为 23 mm,焦距长度 12 mm,光圈 F1.4。

图 4 熔池视觉传感系统示意图

2.2 数据采集结果

本文试验所使用的母材为尺寸 150 mm×55 mm×1.5 mm Q235 碳钢板,接头形式为对接,背面悬空,无衬板。为了保证进行深度神经网络训练时有足够数量而具有代表性的数据集,试验通过改变焊接电流或焊接速度,在 70 个参数组合下,如表 1 所示,获取了熔池正面图像。由于在起弧和收弧时,焊接速度会发生变化,在这些情况下拍摄的图片不具有代表性,因此这些图片被剔除,最终的图片集仅包含处于稳定焊接状态时采集的图像。经过筛选后,共采集熔池正面图像 11 197 张,我们将图像分为 3 类:其中背面焊漏对应的熔池正面图像来自 28

表 1 试验参数与熔透情况

焊接 编号	焊接 电流 /A	焊接速度 /(mm/min)	熔透状态	焊接 编号	焊接 电流 /A	焊接速度 /(mm/min)	熔透状态
1	100	40	未焊透	36	150	40	未焊透
2	100	45	恰好焊透	37	150	45	未焊透
3	100	50	焊漏	38	150	50	恰好焊透
4	100	55	焊漏	39	150	55	恰好焊透
5	100	60	焊漏	40	150	60	焊漏
6	100	65	焊漏	41	150	65	焊漏
7	100	70	焊漏	42	150	70	焊漏
8	110	40	未焊透	43	160	40	未焊透
9	110	45	恰好焊透	44	160	45	未焊透
10	110	50	焊漏	45	160	50	未焊透
11	110	55	焊漏	46	160	55	恰好焊透
12	110	60	焊漏	47	160	60	恰好焊透
13	110	65	焊漏	48	160	65	焊漏
14	110	70	焊漏	49	160	70	焊漏
15	120	40	未焊透	50	170	40	未焊透
16	120	45	恰好焊透	51	170	45	未焊透
17	120	50	恰好焊透	52	170	50	未焊透
18	120	55	焊漏	53	170	55	恰好焊透
19	120	60	焊漏	54	170	60	恰好焊透
20	120	65	焊漏	55	170	65	焊漏
21	120	70	焊漏	56	170	70	焊漏
22	130	40	未焊透	57	180	40	未焊透
23	130	45	未焊透	58	180	45	未焊透
24	130	50	恰好焊透	59	180	50	未焊透
25	130	55	恰好焊透	60	180	55	未焊透
26	130	60	焊漏	61	180	60	恰好焊透
27	130	65	焊漏	62	180	65	恰好焊透
28	130	70	焊漏	63	180	70	焊漏
29	140	40	未焊透	64	190	40	未焊透
30	140	45	未焊透	65	190	45	未焊透
31	140	50	恰好焊透	66	190	50	未焊透
32	140	55	恰好焊透	67	190	55	未焊透
33	140	60	焊漏	68	190	60	未焊透
34	140	65	焊漏	69	190	65	恰好焊透
35	140	70	焊漏	70	190	70	恰好焊透

个参数,共 5 244 张,背面恰好焊透对应的熔池正面图像来自 18 个参数,共 2 655 张,背面未焊透对应的熔池正面图像来自 24 个参数,共 3 298 张。图 5 从左至右分别为背面焊漏、恰好熔透和未焊透情况下对应的正面熔池图像。

图 5 不同背面熔透对应的熔池正面图像

3 训练集与测试制作

3.1 训练集与数据集划分

在 CNN 的训练过程中,不平衡训练集,即各个分类中的样本数量相差较大的数据集对模型的准确性有较大的负面影响。因此,本方法建立训练集时保证了焊漏、恰好熔透和未焊透三种情况下具有等量的图片样本,且在每种熔透情况内部,在各个焊接参数下获得的图片数量也相同。

由于在同一参数下进行焊接采集的图片具有较强的相似性,且这种相似性来源于相同的焊接参数而不是相同的背面熔透情况,因此训练集和测试集必须使用在不同焊接参数下采集的图像,这样测试集才能有效地对模型预测表现进行验证。

综合以上两个因素,本试验按照训练集平衡的排除相似性等原则,对训练集和测试集进行了如表 2 的划分。

表 2 训练集与测试集划分

熔透情况	训练集	测试集
焊漏	1 200 张,分属 28 个参数,平均每个参数 42 张	764 张,分属 4 个参数
恰好焊透	1 200 张,分属 18 个参数,平均每个参数 67 张	351 张,分属 3 个参数
未焊透	1 200 张,分属 18 个参数,平均每个参数 67 张	457 张,分属 4 个参数
总计	3 600 张	1 572 张

3.2 生成训练集与测试集

本试验采集的熔池正面图像尺寸为 1280×1024 ,只截取大小为 160×160 的含有熔池的窗口进行分析。生成数据结构前首先将图像读入工作需要将图像读入工作区中并形成四维矩阵,矩阵的前两个维度代表图片的大小,为 160×160 。第三个维度代表通道数,使用灰度图,因此通道数为 1。最后

一维代表矩阵中所含图片数量,训练集和测试集含有的图片数量分别为 3 600 和 1 572。在读入图片后,需要对每一张图建立一个标签,这一标签代表了熔池正面图像对应的背面熔透情况。然后将所有标签合成一个标签矢量。分别将训练图像与测试图像与对应的标签读入,即得到训练和测试 CNN 所需的训练集和测试集。最终形成的训练集和测试集结构如表 3 所示。

表 3 训练集与测试集的数据结构

类别	图像矩阵	标签矢量
训练集	(160, 160, 1, 3 600)	(3 600, 1)
测试集	(160, 160, 1, 1 572)	(1 572, 1)

4 CNN 模型结构及性能分析

4.1 参数设定

进行神经网络训练之前,首先要指定一系列训练参数(也称超参数),包括学习速率, batch-size 和迭代次数等。采取使用动量的随机梯度下降(SGDM)优化算法通过试验确定了网络训练最合适的学习速率为:初始学习速率 10^{-4} 每 4 个周期降低 20%。其他超参数对模型训练结果也有影响,迭代次数过多则时间消耗大; batch-size 过大超过硬件的运算能力限制,而 batch-size 过小则可能造成模型梯度下降方向不稳定,难以收敛,进行以下参数设置,如表 4 所示。

表 4 参数设定

参数	设定值
训练 Epoch 数	50
迭代次数	1 400
Batch-Size	128

4.2 模型预测性能分析

卷积神经网络的预测性能受到很多因素的影响,如全连接层数,卷积核尺寸等,通过对以上变量进行研究,找出最适用于背面熔透预测任务的网络参数。第一层卷积核承接输入图像,决定了首次卷积时感受野的大小。本文设定了 6 组参数进行熔透准确性以及训练时间对比。最终得到的结果如图 6 所示。可以看出,当卷积核尺寸为 9×9 时,熔透预测模型的准确率最高。这是由于当卷积核过小时,模型无法从较为宏观的角度提取熔池所具有的特征,而由于输入图像大小为 160×160 ,分辨率较小,不含有大尺度的复杂信息,因此使用过大的卷积核不能正确地提取图像的特征。从训练时间角度,如图 7 所示,网络训练时间随卷积核尺寸的增加而递

增,但卷积核尺寸从 3×3 至 9×9 时,网络训练时间仅增长了19%,当卷积核尺寸上升至 13×13 时,网络训练时间大幅上升,达到 11×11 卷积核训练时间的2.98倍。

图6 第一层卷积核尺寸对预测准确率的影响

图7 第一层卷积核尺寸对训练时间的影响

最后一个卷积层含有的卷积核数量决定了全连接层的输入矢量大小。全连接层具有对卷积层提取的特征进行分类的作用。采取10组不同的卷积核数量对网络进行训练,分别得到了准确率与卷积核数量的关系以及训练时间与卷积核数量的关系,如图8所示。当卷积核数量为64时,模型的预测准确率最高。造成这种现象的原因是,当卷积核数量较少时,卷积层对图像特征的提取不够充分,而当卷积核数量过多时,网络所需要训练的参数过多,容

图8 最后一层卷积核数量对训练时间的影响

易出现过拟合(over-fitting)现象,即网络对训练集有较好的拟合结果,但是泛化能力较弱,因此无法对陌生的数据集进行有效预测。当卷积核数量为64时,既从熔池图像中提取了很多特征,又避免了过拟合现象。从训练时间角度,网络训练时间与卷积核数量几乎成线性关系,卷积核数量从最小值16到64,网络的训练时间仅增加17.7%。如图9所示。

图9 最后一层卷积核数量对预测准确率的影响

4.3 模型预测表现分析

CNN的首层卷积核尺寸为 9×9 ,末层卷积核数量为64时可取得最优的熔透预测表现。根据以上结论,设计了如图10所示的CNN结构,它包含3个卷积层,每个卷积层后加上一个池化层以降低计算量。前两个卷积层后还加上了 dropout 层以防止过拟合。在3个卷积层后是全连接层和 softmax 层,它们的作用是根据卷积层的输出对特征进行分类。建立模型后,根据训练参数对网络进行训练。

图10 CNN网络结构

和测试集的预测准确率变化如图11所示。从图11中可以看出,在前200次迭代中,模型在训练样

本和测试集上的准确率迅速上升, 其后, 准确率逐渐收敛, 曲线振荡的幅度逐渐减小。最终模型对训练集的预测准确率稳定在 100%, 而对测试集的准确率稳定在 96% 以上, 说明基于 CNN 的模型在训练后可以收敛。

图 11 CNN 模型训练过程

除了分析背面熔透预测的整体准确率外, 还需要对各个熔透情况下的预测准确率进行分析。如图 12 列出了模型在三种熔透情况下的表现。可以看出, 在背面焊漏, 恰好焊透和未焊透的情况下, 模型预测的准确率都超过了 94%。其中对于焊漏的情况具有最高的预测准确率, 即 98.82%。据此可以得出结论, 经过训练后, 基于 CNN 的熔透预测模型对各种背面熔透情况都有较好的预测效果, 可以达到对背面熔透进行准确预测的要求。

图 12 各种熔透情况下的预测准确率

4.4 特征可视化

在卷积神经网络中, 特征映射是使用一个卷积核(神经元)在前一层图像上进行卷积运算后得到的输出。因此, 一张特征映射图可以反映卷积核对上一层的哪些部分有较强的响应, 进而判断卷积核学习到了图像的哪些特征。分别将背面焊漏和背面未焊透对应的正面图片作为输入, 对第一个卷积层的特征映射进行可视化, 并分析 CNN 是通过熔池图像的哪些特征对背面熔透进行分类。

第一个卷积层的尺寸为 $9 \times 9 \times 3$, 因此具有三个不同的特征映射。三个特征映射和原图的对照如图 13 所示。其中第一个和第三个卷积核有部分区域比

较明亮, 表明这两个卷积核对于输入图像的响应较强, 而第二个卷积核的响应较弱。因此主要对这第一个和第三个卷积核对应的特征映射进行分析。从第一个卷积核的特征映射和原图的对比可以看出, 该图像在熔池上的反光点与熔池的左右两段边缘处亮度较高, 即响应较强。因此该卷积核成功识别出了熔池上的反光点和熔池的右上方的边缘两种特征。从第三个卷积核的特征映射图可以看出, 该卷积核对于原图像尾部大片的黑色区域响应较强, 即第三个卷积核成功识别了熔池塌陷的尾部。

图 13 原始图像和三个卷积核的特征映射

根据以上分析, 可以看出, 训练后的神经网络是通过判断熔池正面图像的边缘, 尾部塌陷和反光点等特征来对背面焊漏、未焊透和恰好焊透的情况进行区分。这一判断过程与人类的判断过程非常相似。同时, 卷积核对于烟雾反光, 熔渣和图像干扰条纹没有做出较强的响应, 说明该模型对于干扰具有较强的抵抗能力。

5 结论

(1) 搭建熔池视觉传感系统获得了清晰的熔池图像, 并生成了训练集和测试集。

(2) 进行了 CNN 网络参数设定并分析了卷积核尺寸对预测准确性的影响, 并选定第一层卷积核尺寸为 9×9 , 最后一层含有 64 个卷积核的参数为最优网络参数。

(3) 建立了 CNN 网络模型并对模型的预测准确性进行了分析, 模型的总体预测准确率较高, 而且对各种背面熔透情况都有良好的预测效果。

(4) 对卷积层的特征映射进行了可视化, 发现了神经网络模型是通过边缘, 反光点和尾部塌陷等

信息做出响应和判断, 解释了熔透预测模型进行判断的根据。

参 考 文 献

- [1] 石珂, 张刚, 马晓骥, 等. 基于激光视觉的脉冲 GTAW 熔池振荡检测与分析[J]. 机械工程学报, 2012, 48(24): 28-32.
SHI Yu, ZHANG Gang, MA Xiaoji, et al. Detection and analysis of pulsed GTAW weld pool oscillation based on laser vision[J]. Journal of Mechanical Engineering, 2012, 48(24): 28-32.
- [2] 高进强, 王德民, 武传松. 基于视觉的 MAG 焊接熔池的标定[J]. 焊接学报, 2009, 30(2): 34-36.
GAO Jinqiang, WANG Demin, WU Chuansong. Visual calibration of MAG welding pool[J]. Transactions of the China Welding Institution, 2009, 30(2): 34-36.
- [3] CHEN J, CHEN J, FENG Z, et al. Dynamic evolution of the weld pool reflection during weld penetration development[C]// IEEE International Conference on Advanced Intelligent Mechatronics. IEEE, 2016: 548-553.
- [4] ZHANG W J, XIAO J, ZHANG Y M. A mobile sensing system for real-time 3D weld pool surface measurement in manual GTAW[J]. Measurement Science & Technology, 2016, 27(4): 045102.
- [5] 刘新锋. 基于正面熔池图像和深度学习算法的 PAW 穿孔/熔透状态预测[D]. 济南: 山东大学, 2017.
LIU Xinfeng. Prediction of penetration of PAW based on frontal molten pool image and depth learning algorithm[D]. Jinan: Shandong University, 2017.
- [6] KOVACEVIC R, ZHANG Y M, LI L. Monitoring of weld joint penetration based on weld pool geometrical appearance[J]. Welding Journal, 1996, 75(10): 317s-329s.
- [7] 高进强. 恒流 TIG 焊接熔池形状参数的视觉检测及智能控制[D]. 济南: 山东大学, 2000.
GAO Jinqiang. Visual detection and intelligent control of shape parameters of molten pool for TIG welding with constant current[D]. Jinan: Shandong University, 2000.
- [8] 刘新锋, 高进强, 姜锡瑞, 等. 连续电流 TIG 焊接熔池图像的低成本自动化检测[J]. 焊接学报, 2001, 22(6): 25-28.
LIU Xinfeng, GAO Jinqiang, JIANG Xirui, et al. Low-cost automatic detection of continuous current TIG welding pool image[J]. Transactions of the China Welding Institution, 2001, 22(6): 25-28.
- [9] 赵衍华. 连续电流 TIG 焊接熔池形状检测与神经网络建模[D]. 济南: 山东大学, 2002.
ZHAO Yanhua. Shape detection and neural network modeling of continuous current TIG welding molten pool[D]. Jinan: Shandong University, 2002.
- [10] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.
ZHOU Feiyan, JIN Linpeng, DONG Jun. A review of research on convolutional neural networks[J]. Chinese Journal of Computers, 2017, 40(6): 1229-1251.

作者简介: 李海超(通信作者), 男, 1975 年出生, 博士, 副教授。主要研究方向为机器人与智能焊接。
E-mail: lihaichao@hit.edu.cn