

基于视觉注意 VGGNet 的中厚板低碳钢 多层多道熔池分类模型

周浩, 陈善本

(上海交通大学, 智能化机器人焊接技术实验室, 上海, 200240)

摘要: 重大装备制造中厚板机器人多层多道焊 (multi-layer and multi-pass welding, MLMPW) 一直是热点和难点, 而实现机器人 MLMPW 的核心是对其熔池的获取、监控并分类. 为了提高 MLMPW 的自动化和智能化, 有必要开发一个熔池图像在线分类系统. 针对焊接过程中的熔池图像提出了一种新的 MLMPW 熔池分类方法——基于视觉注意的 (SENet)VGGNet 熔池分类方法. 为了提高效率和精度, 引入迁移学习中的预训练模型到网络训练过程中. 因为针对中厚板多层多道熔池研究较少, 导致熔池公开数据集较少, 为了应对这一问题, 需要对数据集进行增广. 结果表明, 提出的模型可快速有效的对七类 MLMPW 熔池进行准确分类, 预测精度可达到 98.39%.

创新点: (1) 建立 MLMPW 熔池视觉监控系统, 采集到高质量的熔池图像.

(2) 提出了基于视觉注意的 VGGNet 的 MLMPW 熔池分类模型, 可准确预测七种不同类型熔池.

关键词: 多层多道焊; 熔池; VGG16; 视觉注意; 分类模型

中图分类号: TG 409 **文献标识码:** A **doi:** 10.12073/j.hjxb.20240710001

A MLMP welding pool classification model for medium-thick low-carbon steel plates based on a VGGNet with a visual attention mechanism

ZHOU Hao, CHEN Shanben

(Intelligentized Robotic Welding Technology Laboratory, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: The multi-layer and multi-pass welding (MLMPW) of medium-thick plates in large equipment manufacturing has always been a challenging and important research topic. The core of achieving robotic MLMPW lies in acquiring, monitoring, and classifying the weld pool. To enhance the automation and intelligence of MLMPW, it is necessary to develop an online classification system for weld pool images. A novel MLMPW weld pool classification method is proposed for weld pool images during the welding process, utilizing a VGGNet with a visual attention mechanism (SENet). To improve efficiency and accuracy, pre-trained models from transfer learning are introduced into the network training process. Due to the lack of research on the medium-thick plate multi-layer and multi-pass molten pool, there are few open datasets of molten pool. In order to deal with this problem, data augmentation is necessary to applied to the dataset. The results indicate that the proposed model can quickly and effectively classify seven types of MLMPW weld pools with an accuracy of 98.39%.

Highlights: (1) An MLMPW weld pool visual monitoring system to capture high-quality weld pool images was developed.

(2) An MLMPW weld pool classification model based on VGGNet with a visual attention mechanism was proposed, which can accurately predict seven different types of weld pools.

Key words: multi-layer multi-pass welding; weld pool; VGG16; visual attention; classification model

0 序言

在重大装备制造中,多层多道焊 (MLMPW) 技术一直是重点和难点^[1]. MLMPW 一般需要进行离线规划^[1-3],但是在焊接过程中不可避免的产生偏差,如机器人姿态、焊丝偏差造成的焊道填充不均匀,以及一些常见的焊接缺陷^[4-5](未熔合,未熔透,夹渣,气孔等)形成的瓶颈难题严重影响焊接质量,因此需要在焊接过程中进行实时调整.在一般情况下,实时调整的依据是通过熔深对焊缝成形进行控制,然而,从视觉传感器中能直接获取的只有正面熔池的特征几何信息,只能从正面熔池信息预测熔深.中厚板对接中的 MLMPW 熔池两边侧壁因为受到坡口以及已焊焊道的约束,熔池的特征信息将会有较大差异^[6],不同类型的熔池对应不同的熔深进而对应不同类型的成形,从而可以对焊接过程进行较为精确的动态控制.文中主要的研究内容是通过基于视觉注意 (SENet^[7]) 模型的 VGGNet^[8] 架构对多层多道焊接熔池进行精确分类.

1 数据收集

1.1 试验系统和试验参数设置

文中涉及的多层多道 MAG 焊熔池图像数据是通过视觉传感系统采集的,如图 1 所示.试验系统包括 Fanuc M-10iA 型机器人(包括控制器和示教器),工控机, Xiris XVC-1000 型高动态范围 (HDR) 相机(图 2), Fronius TPS 400i 焊接电源和保护气(80% 氩气和 20% 二氧化碳)组成.试验母材为厚度 30mm 坡口为 30° 的 Q235 普通碳钢,使用直径为 1.2mm 的焊丝完成多层多道对接填充试验.

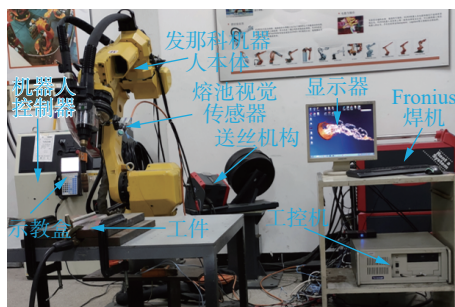


图 1 MLMPW 熔池视觉传感系统

Fig. 1 MLMPW Molten Pool Vision Sensing System

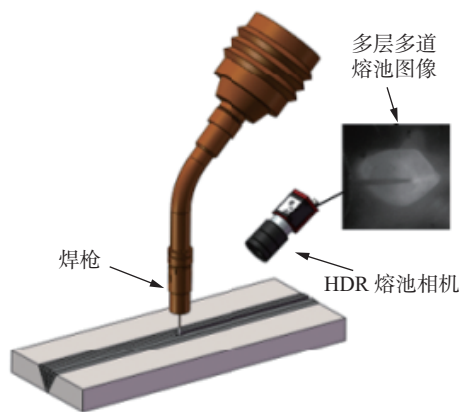


图 2 HDR 相机位置关系

Fig. 2 Positional relationship of HDR camera

如图 3 所示,遵循先两边后中间的 MLMPW 顺序,通过试验探索出焊接顺序和对应的焊接参数. MLMPW 的焊接方式和工艺参数如表 1 所示.

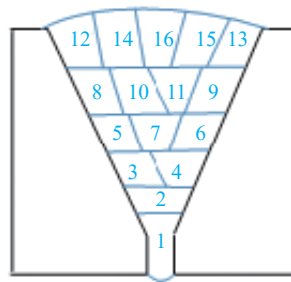


图 3 MLMPW 排序

Fig. 3 MLMPW sequence

表 1 MLMPW 参数
Table 1 MLMPW parameters

焊接类型	焊接速度 $v/(\text{mm} \cdot \text{min}^{-1})$	打底焊电流 I_1/A	填充焊电流 I_2/A	焊枪倾斜角 $\theta/^\circ$	气流量 $q/(\text{L} \cdot \text{min}^{-1})$	母材尺寸 $L \times W \times H/\text{mm}$
CMT	240	170 ~ 185	210 ~ 280	-15 ~ 15	20	400 × 150 × 30

1.2 试验系统和试验参数设置

如图 4 所示,机器人 MLMPW 理论模型分为三种:三角形成形、梯形成形和平行四边形成形.焊道 1 为三角形成形只对应于打底焊,焊道 2, 4, 7, 11, 16 为梯形成形,焊道 3, 5, 6, 8, 9, 10, 12, 13,

14, 15 为平行四边形成形.细分的情况为打底焊为类型 1(T1); 类型 2(T2, 焊道 2), 类型 3(T3, 焊道 4), 类型 4(T4, 焊道 7, 11, 16) 为梯形成形中两侧为坡口、一侧为焊道一侧为坡口和两侧为焊道的情况; 类型 5(T5, 焊道 3, 5, 6, 8, 9)、类型 6(T6, 焊道 10,

14, 15) 和类型 7(T7, 焊道 12, 13) 为平行四边形中一侧为坡口、一侧为已经填充焊道和盖面焊的情况。

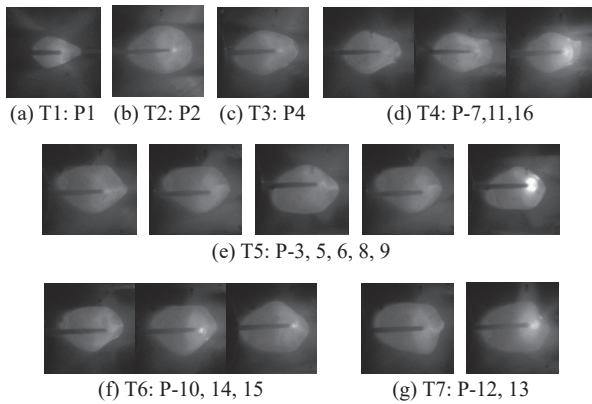


图 4 MLMPW 7 种类型熔池

Fig. 4 MLMPW 7 types of molten pool. (a) T1: P1; (b) T2: P2; (c) T3: P4; (d) T4: P-7,11,16; (e) T5: P-3, 5, 6, 8, 9; (f) T6: P-10, 14, 15; (g) T7: P-12, 13

1.3 数据集增广

MLMPW 熔池公开的数据集极少,为了训练熔池分类模型,必须自建数据集.为了应对自建数据集规模较小的问题,必须丰富数据集的多样性和数量,进而对其进行数据增广,包括旋转、翻转、随机平移和缩放的组合操作.如图 5 所示,随机选取第四层第三道(第七道)对应的熔池作为样例,对数据增广进行展示。

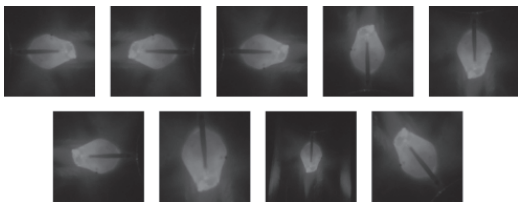


图 5 熔池数据增广. 第 1 幅为原图, 第 2-6 幅为翻转和顺时针旋转三次 90° 的图像, 第 7-9 幅图像为随机旋转、缩放和平移组合的图像。

Fig. 5 Melt pool data augmentation. The first image is the original image, the second to sixth images are flipped and rotated three times 90° clockwise, and the seventh to ninth images are images with random rotation, scaling, and translation.

2 模型与算法

本研究使用基于视觉注意深度学习模型对多层多道熔池进行分类.根据焊接不同的层道特性分为七种熔池模型^[6].如图 6 所示的研究框架,通过 MLMPW 试验获得熔池图像并进行标记;对数据集

进行增广并设定训练集和测试集.为了提高训练效率和精度,分类模型加载预训练权值到分类网络中.试验涉及四个分类模型:(1)从零开始训练 VGG16 网络 (M1), (2) 训练 Pre-trained VGG16 网络 (M2), (3) 从零开始训练 SENet-VGG16 网络 (M3), (4) 训练 Pre-trained SENet-VGG16 网络 (M4).

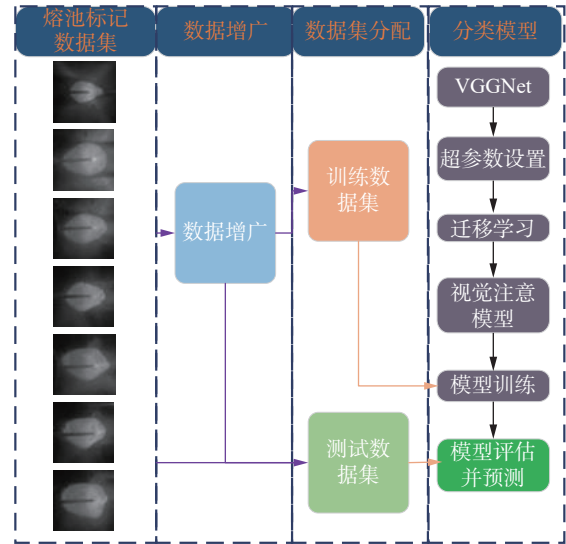


图 6 MLMPW 熔池分类总体 workflow

Fig. 6 Overall workflow of MLMPW melt pool classification

2.1 VGGNet

如图 7 所示, VGG16 是由一系列连续的 3*3 卷积层、2*2 的池化层、三个完全连接层 (FC) 和 Softmax 输出层组成。

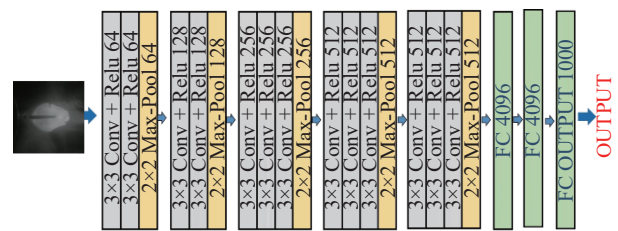


图 7 VGG16 分类架构

Fig. 7 VGG16 classification architecture

2.2 SENet

如图 8 所示, 作为通道注意力机制代表的是 squeeze-and-excitation network (SENet)^[7], 基本原理包括“挤压”(squeeze)和“激励”(excitation)两个操

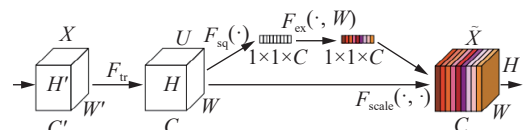


图 8 SENet 模型

Fig. 8 SENet Model

作, 分别对特征图进行全局信息的嵌入和自适应重标定, 设定不同的通道权值表达视觉注意力影响力。

2.3 迁移学习

为了应对自建数据集中数据相对较少的问题, 必须引入迁移学习到 MLMPW 熔池分类网络中. 训练数据集规模较小会引起网络欠拟合, 并且从零开始训练网络时间过长, 这些因素会引起网络鲁棒性较差. 引入预训练模型权值到熔池分类网络中可以

应对这些潜在不利因素, 提升网络分类性能。

3 结果和讨论

学习算法中的超参数的合理设定是分类网络准确与否的关键因素. 如表 2 所示, 根据试验经验设置超参数. 经过训练, M1、M2、M3 和 M4 的分类准确率为 93.41%、97.73%、96.78% 和 98.39%, 所设计的分类网络达到最优分割性能。

表 2 超参数设置
Table 2 hyperparameter configuration

优化算法	Momentum	学习率	损失函数	最大epochs ε	Batch size
SGD	0.9	0.001	cross entropy	50(eps)	8(幅图)

熔池分类网络的综合评价标准为精确率 (Pr)、召回率 (Re)、 $F1$ 分数 ($F1$). 可表示为

$$Pr = \frac{TP}{TP + FP} \quad (1)$$

$$Re = \frac{TP}{TP + FN} \quad (2)$$

$$F1 = \frac{2Pr \times Re}{Pr + Re} \quad (3)$$

式 (1)~式 (3) 中的物理量如表 3 所示. TP 表示预测值为 1, 实际值也为 1; FP 表示预测值为 0, 实际值为 1; FN 表示预测值为 0, 实际值为 0; TN 表示预测值为 1, 实际值为 0。

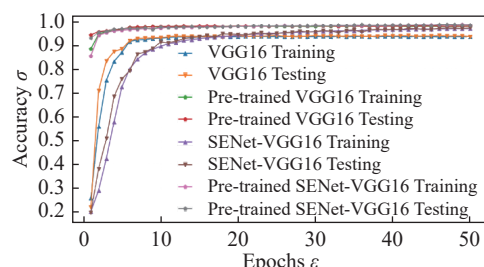
表 3 混淆矩阵
Table 3 Confusion Matrix

Confusion matrix	Actual value	
Predicted value	TP	FP
	FN	TN

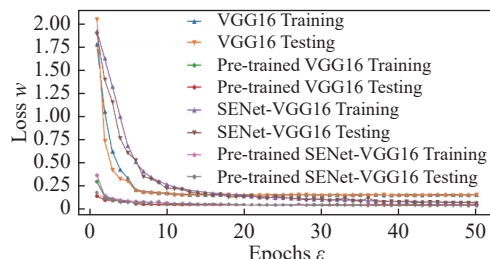
学习曲线可以可视化分类器的学习性能. 网络的训练集和测试集评估模型的准确率和损失的曲线如图 9 所示, 从 Loss 曲线的波动性情况分析, 所有模型的训练曲线未有波动的情况, 这说明训练过程没有陷入局部极小值的情况. 引入预训练权值模型的初始精度较高, 高于 0.85, 同时初始的 Loss 值较低, 低于 0.4。

如图 10 所示, 为进一步研究分类性能, 引入混淆矩阵可视化分类性能. 可以看出, 多层多道熔池分类整体准确率最高的是 M4, 达到最高的 98.39%, 验证了引入 SENet 可提高整体分类准确率。

综合性能指标如表 4 所示, 可以更全面的判断



(a) 准确度曲线



(b) Loss 曲线

图 9 训练和测试曲线

Fig. 9 Training and testing curves. (a) accuracy curve; (b) loss curve

分类器的性能. 可以看出, 除了 M1 之外, 其他模型的精确率都在 96% 以上. 在整体分类精度上, M4 是优于其他三个模型的, 特别是 Type3, M4 比 M2 高 1.03%, 提升较大; 然而, 在除了 Type3 和 Type6 外的其他类别分类性能出现了小幅度的局部退化, 分类性能需再平衡优化。

在所有模型中, 多层多道焊 Type1 为打底焊, 达到了近似于 1 的分类精度, 这是因为打底焊熔池与其他层道的熔池特征差距较大. 其中 Type6 的分类精度最低, 原因是同属于平行四边形成形对应的熔池, 有其相似性, 区别在于 Type6 的熔池前端更加平滑. 此时如果有弧光遮挡熔池的前端, 将加

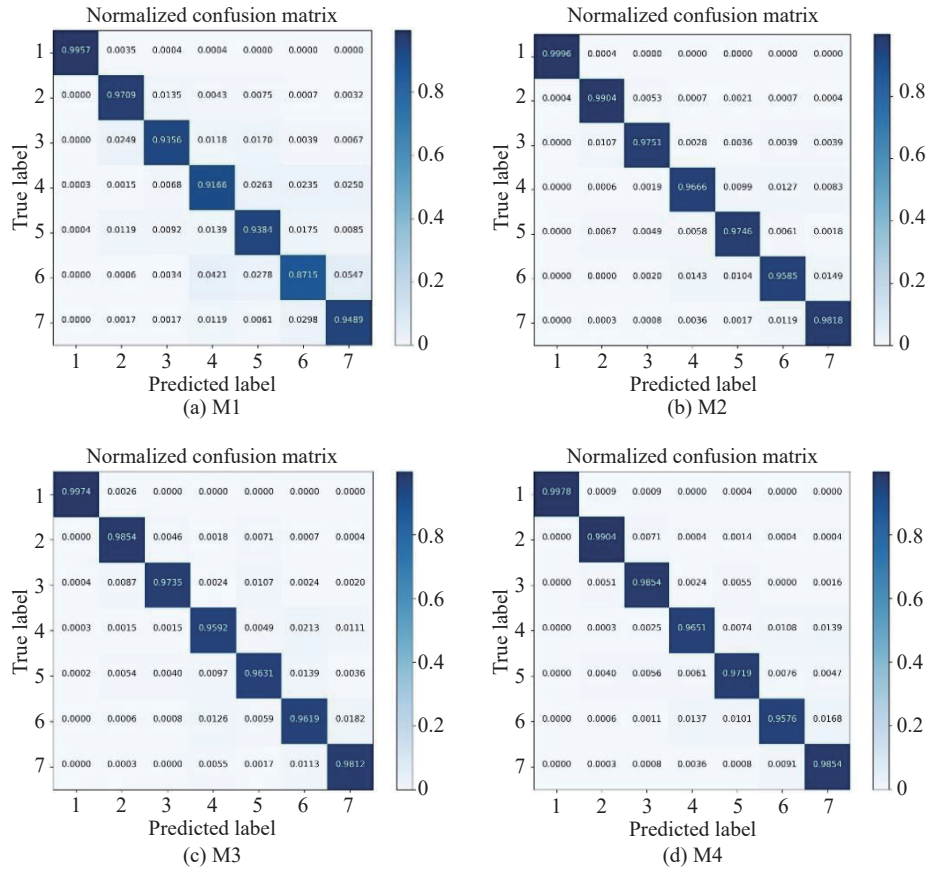


图 10 混淆矩阵

Fig. 10 Confusion Matrix. (a) M1; (b) M2; (c) M3; (d) M4

表 4 熔池分类评价价值 (%)

Table 4 Molten pool classification evaluation value

模型		Type1	Type2	Type3	Type4	Type5	Type6	Type7
Pt_se_vgg16	Pr	100.00	98.69	97.58	97.02	98.14	97.07	96.46
	Re	99.78	99.04	98.54	96.51	97.19	95.76	98.54
	F1	99.89	98.87	98.05	96.76	97.66	96.41	97.49
se_vgg16	Pr	99.87	97.88	98.44	96.31	97.94	95.01	96.65
	Re	99.74	98.54	97.35	95.92	96.31	96.19	98.12
	F1	99.81	98.21	97.90	96.12	97.12	95.60	97.38
Pt_vgg16	Pr	99.96	97.86	97.90	96.93	97.97	96.53	97.29
	Re	99.96	99.04	97.51	96.66	97.46	95.85	98.18
	F1	99.96	98.45	97.70	96.80	97.71	96.19	97.73
vgg16	Pr	99.87	95.23	95.18	90.87	93.92	91.90	91.00
	Re	99.57	97.09	93.56	91.66	93.84	87.15	94.89
	F1	99.72	96.15	94.36	91.26	93.88	89.46	92.90

刷分类的误判。验证了引入 SENet 可提高整体分类准确率。

4 结论

(1) 为了控制 MLMPW 质量, 开发了熔池可视化监测系统。

(2) 通过试验比较, 带有视觉注意力机制的 VGG16 网络, 训练过程通过加载预训练权值获得了整体最优的分类性能。

(3) 精确的对熔池进行分类可以避免在 MLMPW 过程中出现难以修复的焊道填充缺陷。

参考文献

- [1] Zhang H, Lu H, Cai C, *et al.* Robot path planning in multi-pass weaving welding for thick plates [C]//Robotic Welding, Intelligence and Automation: RWIA'2010. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011: 351 – 359.
- [2] Xue K, Wang Z, Shen J, *et al.* Robotic seam tracking system based on vision sensing and human-machine interaction for multi-pass MAG welding[J]. *Journal of Manufacturing Processes*, 2021, 63: 48 – 59.
- [3] Yang C, Zhong J, Chen Y, *et al.* The realization of no back chipping for thick plate welding[J]. *The International Journal of Advanced Manufacturing Technology*, 2014, 74: 79 – 88.
- [4] 徐东辉, 孟范鹏, 孙鹏, 等. 基于深度学习的 GMAW 焊接缺陷在线监测 [J]. *焊接学报*, 2024, 45(3): 114 – 119.
- Xu D, Meng F, Sun P, *et al.* Online monitoring of GMAW welding defect based on deep learning[J]. *Transactions of the China Welding Institution*, 2024, 45(3): 114 – 119.
- [5] 高昶霖, 宋燕利, 左洪洲, 等. 基于动态权重的自适应 PSO-BP 神经网络焊接缺陷成因诊断 [J]. *焊接学报*, 2022, 43(1): 98 – 106.
- Gao C, Song Y, Zuo H, *et al.* Cause diagnosis of welding defects based on adaptive PSO-BP neural network with dynamic weighting[J]. *Transactions of The China Welding Institution*, 2022, 43(1): 98 – 106.
- [6] Zhou H, He Y, Chen H, *et al.* MLD classification model of visual features of multi-layer and multi-pass molten pool during robotic MAG welding of medium-thick steel plates[C]//International Conference on Robotic Welding, Intelligence and Automation. Singapore: Springer Nature Singapore, 2020: 56 – 81.
- [7] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition, Salt Lake City, 2018: 7132 – 7141.
- [8] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [J]. *arXiv preprint arXiv: 14091556*, 2014.

第一作者:周浩, 博士; 主要研究方向为图像处理模式识别, 焊接过程控制; Email: image_processing@sjtu.edu.cn.

通信作者:陈善本, 博士, 教授, 博士研究生导师; Email: sbchen@sjtu.edu.cn.

(编辑: 周珍珍)