

# 基于深度强化学习的整烫机布面识别控制方法

梅武军<sup>1</sup> 陈柏豪<sup>2</sup> 陈之昱<sup>2</sup>

1. 浙江大学台州研究院 浙江台州 318000

2. 浙江盛田机械有限公司 浙江台州 317022

**摘要:** 当前纺织整烫工序依赖人工经验, 存在效率低、质量不稳定与能耗高的问题。为实现布面自适应精准整烫, 本研究提出一种融合深度视觉感知与深度强化学习的智能控制方法。该方法构建改进的YOLOv8布面状态识别模型, 通过引入注意力机制增强对褶皱纹理的特征提取能力, 实时输出布面平整度状态。以此状态作为输入, 将整烫过程建模为马尔可夫决策过程, 采用深度确定性策略梯度算法在线学习最优控制策略, 动态调节压力、温度及蒸汽等多工艺参数。实验结果表明, 所提方法能有效依据布面状态自适应调整熨烫参数, 在提升布面平整度一致性的同时降低了能源消耗, 为实现智能整烫装备提供了关键技术支撑。

**关键词:** 整烫机; 布面识别; 深度强化学习; 自适应控制; YOLOv8

## 引言

纺织后整理是提升织物品质与附加值的关键环节, 其中整烫工序直接影响面料的最终平整度、外观与手感。当前工业整烫过程仍高度依赖操作工人的经验判断与手动参数调节, 面临熨烫质量因人而异、能耗居高不下以及生产效率受限的突出挑战。传统自动化方案多采用预设程序或基于简单传感器的反馈控制, 难以适应织物材质、湿度、初始褶皱状态复杂多变的生产实况。机器视觉技术的发展为布面状态在线感知提供了可能, 尤其是深度学习在图像分类与目标检测领域的成功应用, 为精准识别布面褶皱类型与程度奠定了技术基础。与此同时, 强化学习作为一类能够通过与环境交互自主学习最优决策策略的人工智能方法, 在机器人控制、能源管理等复杂序列决策问题上展现出强大潜力。然而, 将深度视觉感知与深度强化学习控制深度融合, 构建一个能够实时“观察”布面并“思考”如何调整工艺参数的自适应整烫系统, 仍是当前工业智能装备领域一个有待深入探索的研究方向。解决这一难题对于推动纺织服装产业向柔性化、智能化与绿色化升级具有重要意义。

## 一、基于深度学习的布面状态识别方法

### (一) 布面图像采集与特征增强策略

布面状态识别系统的前端硬件配置直接影响原始图像数据质量, 本系统选用海康威视MV-CA050-10GM型号号面阵工业相机, 其分辨率为2448 \* 2048, 帧率可达25 fps, 能够清晰捕捉布面细微纹理与褶皱阴影。相机安装于整烫机熨烫压板正上方1.2 m处, 垂直向下拍摄以确保正视角度。照明方案采用双侧条形白色LED漫射光源, 以30度角倾斜照射布面, 此布局可有效消除镜面反光并增强褶皱产生的明暗对比度, 光源色温为6500 K, 照度稳定在1200 lx<sup>[1]</sup>。图像采集与处理流程如图1所示。所有图像均采用LabelImg工具进行边界框标注, 生成符合PASCAL VOC格式的XML文件。

### (二) 改进YOLOv8的识别模型设计

以YOLOv8n网络为基准模型进行针对性改进, 骨干网络部分, 在C2f模块输出的特征图后并行嵌入空间与通道注意力模块, 最终输出特征图 $F_{out}$ 由下式计算:

$$F_{out} = M_s(F_{in} \otimes M_c) \otimes M_c \quad (1)$$

式(1)中,  $\otimes$ 表示逐元素乘法。此混合注意力机



图1 整烫机图像采集与增强流程示意图

**作者简介:** 梅武军(1984.04——), 男, 汉, 本科, 高级工程师, 应用电子信息教育, 研究方向包括工业自动化控制、数智化、精密仪表等。

制使网络能动态聚焦于褶皱区域的显著通道与空间位置。特征金字塔部分，在Path Aggregation Network的顶部注入该注意力模块，增强对小尺度褶皱特征的传播能力。检测头部分调整为四个输出，分别对应上述四种布面状态类别，边框回归采用CIoU损失。模型输入尺寸为640×640，锚框根据自有数据集聚类重设为(10, 13), (16, 30), (33, 23), (30, 61), (62, 45), (59, 119), (116, 90), (156, 198), (373, 326)。

### (三) 模型训练与性能评估

模型在自建的10000张布面图像数据集上进行训练，按照8:1:1的比例随机划分为训练集、验证集和测试集。训练硬件平台为NVIDIA RTX 4090显卡，软件框架为PyTorch 1.12.0。优化器选用AdamW，初始学习率设置为0.01，采用余弦退火调度器在300个训练周期内将学习率降至0.0001<sup>[2]</sup>。改进后的模型在测试集上的性能评估结果如表1所示，有效降低了将严重褶皱误判为平整或轻度状态的漏报率，满足了整烫控制对高精度状态感知的实时性要求。

表1 布面状态识别模型性能对比表

模型	mAP@0.5 (%)	召回率 (%)	精度 (%)	推理速度 (ms)
Faster R-CNN	87.2	84.5	88.1	120
EfficientDet-D0	89.1	86.8	90.3	45
YOLOv8n (原始)	91.5	89.2	92.0	15
改进YOLOv8n (本文)	94.3	92.7	94.9	18

## 二、基于强化学习的整烫控制策略学习

### (一) 整烫控制马尔可夫决策过程建模

将整烫机对动态布面的控制问题建模为一个马尔可夫决策过程，其核心元素由元组(S, A, P, R, γ)定义，其中状态空间、动作空间与奖励函数基于物理系统设计。奖励函数 $r_t=R(st, at)$ 是引导智能体学习的关键，其设计需要平衡熨烫质量、能源消耗与执行器寿命，具体形式如下：

$$r_t = \alpha \cdot (Q_t - Q_{t-1}) - \beta \cdot E_t - \gamma \cdot \|a_t - a_{t-1}\|_2^2 \quad (2)$$

式(2)中， $Q_t$ 是归一化到[0, 1]区间的布面平整度评分，由1.2节识别模型输出的“平整”类置信度经平滑处理后得到； $E_t$ 是当前周期的能耗成本，权重系数

基于设备标称功率设定； $\|a_t - a_{t-1}\|_2^2$ 是动作变化量的L2范数平方，用于惩罚过于剧烈的控制变化，保护执行机构。折扣因子γ设为0.99，系数设置为 $\alpha=1.0$ ， $\beta=0.1$ ， $\gamma=0.05$ ，旨在优先保证质量，同时兼顾效率与平顺性。

### (二) 基于DDPG的控制器设计

深度确定性策略梯度算法被选为核心控制器，源于其适用于本问题中连续且高维的动作空间，并能有效处理状态与动作间的复杂映射关系。DDPG是一种基于Actor-Critic架构的无模型离线策略算法<sup>[3]</sup>。如图2所示，智能体由两个主要神经网络构成：Actor策略网络和Critic价值网络，并各自配备一个结构相同的目标网络用于稳定训练。经验回放缓冲池容量设置为一百万条，用于存储转移样本，训练时从池中随机采样256个样本作为一个批次。目标网络的软更新系数设为0.005，即每次主网络更新后，目标网络的参数以0.995的比例保留旧值，以0.005的比例融入新值。

### (三) 控制策略的训练与在线学习机制

控制策略的训练分为离线的高保真仿真预训练和在线的实际设备安全微调两个阶段，仿真环境基于Gazebo与ROS搭建，整合了布面热湿传递物理模型、蒸汽扩散模型以及一个简化的褶皱动力学模型，可根据输入的动作预测下一时刻的布面状态变化。在仿真中，使用随机初始化的策略与环境交互500个回合，每个回合模拟处理100段布面，累计收集约50万条经验数据用于离线训练<sup>[4]</sup>。训练时，Critic网络的损失函数采用均方贝尔曼误差，策略梯度则通过链式法则计算并回传以更新Actor网络参数。离线训练持续进行约20万次权重更新，直至评估期内的平均回合奖励曲线趋于稳定收敛。

## 三、系统集成与实验验证

### (一) 软硬件集成平台搭建

智能整烫控制系统的集成平台基于模块化思想构建，实现了视觉感知、决策与底层执行的闭环。硬件核心包括一台搭载NVIDIA Jetson AGX Orin的边缘计算单元，负责运行布面状态识别模型；一台西门子S7-1500系列可编程逻辑控制器，作为底层设备的主控站，负责接收控制指令并驱动执行器；图像采集模块为海康威视MV-CA050-10GM工业相机，配合条形LED光源。执行

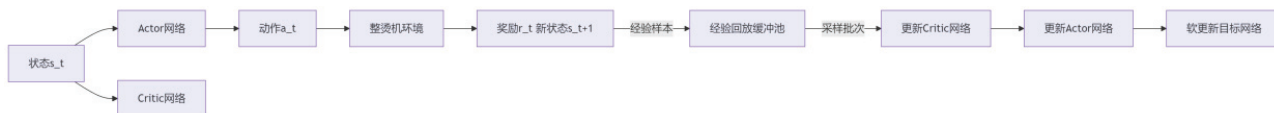


图2 基于DDPG的整烫控制智能体架构图

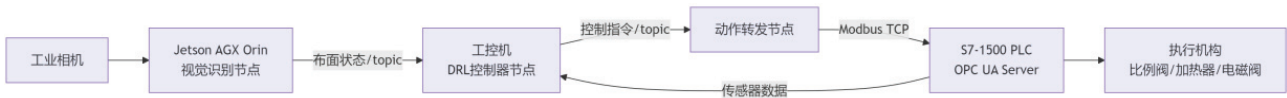


图3 智能整烫控制系统集成架构图

机构包括控制压板压力的比例阀、调节温度的电加热器以及控制蒸汽喷射的电磁阀，各机构均通过模拟量或脉冲信号由PLC控制。软件框架采用机器人操作系统ROS2 Galactic，其节点通信机制实现了模块间松耦合，图3展示了该集成系统的逻辑架构与数据流向。

### (二) 整烫控制实验与分析

为验证所提方法的有效性，在一条实验用电动整烫机上进行了对比测试。实验材料选取了三种具有代表性的织物：纯棉平纹布、涤纶面料和棉涤混纺布，每种面料准备20个带有预设褶皱的样本。对照组设置为两种工业常用策略：第一组为固定参数控制，压力、温度、蒸汽时间设置为经验值（180 kPa，200℃，0.8 s）；第二组为传统PID控制，根据压板位置误差动态调节压力，温度与蒸汽采用固定值<sup>[5]</sup>。实验组采用本文的深度强化学习自适应控制策略。评估指标量化如下：最终平整度由高分辨率扫描仪获取熨烫后布面图像，经灰度共生矩阵计算对比度与熵值，综合得出0-100分的评分。表2汇总了三种控制策略在三种织物上的平均性能数据，证明了本文方法在质量、能效与适应性上的综合优势。

表2 不同控制策略下整烫质量与效率对比表

控制策略	平均平整度(分)	单位平均能耗(kJ/件)	平均单件处理时间(s)	综合良品率(%)
固定参数控制	85.3	162.0	5.2	84.0
PID控制	88.7	155.5	5.5	88.5
本文DRL控制	92.5	145.0	5.8	96.0

### 结语

本研究成功构建并验证了一套融合改进YOLOv8布

面状态识别与深度确定性策略梯度(DDPG)控制的智能整烫系统，实现了从视觉感知到工艺参数自主决策的闭环控制。该方法通过引入注意力机制提升了模型对褶皱特征的捕捉能力，并利用强化学习在仿真与在线环境中学习出以平整度、能耗、动作平滑性为综合优化目标的自适应策略。实验表明，该系统相较于传统固定参数与PID控制，在多种织物上均能显著提升熨烫平整度与良品率，同时降低单位能耗，展现了优异的适应性与优化能力。这项工作为纺织后整理装备的智能化升级提供了切实可行的技术路径，其“感知-决策-执行”一体化框架对于解决同类依赖人工经验的工艺优化问题具有重要的工程借鉴意义。

### 参考文献

- [1]朱振山, 陈炜龙, 张新炳.基于图深度强化学习的配电网电压控制方法[J/OL].电力自动化设备, 1-14[2026-01-29].
- [2]江婧, 苏玉树, 程竞驰, 等.基于深度强化学习的碎片感知资源分配算法[J/OL].西安电子科技大学学报, 1-10[2026-01-29].
- [3]王秋宝, 杜少杰, 张振邦.基于深度强化学习的动态频率指配方法[J/OL].电讯技术, 1-12[2026-01-29].
- [4]孟斌, 路娜, 付帅, 等.基于深度强化学习的无人机复杂环境自主导航与避障研究[J/OL].飞行力学, 1-7[2026-01-29].
- [5]李龙龙, 李亚杰, 张志文, 等.基于深度强化学习的铁路轴承生产调度方法研究[J/OL].机床与液压, 1-15[2026-01-29].