

# 基于深度学习的工业分拣机器人快速视觉识别与定位算法

伍锡如<sup>1,2,3</sup>, 黄国明<sup>1,2</sup>, 孙立宁<sup>3</sup>

(1. 桂林电子科技大学电子工程与自动化学院, 广西 桂林 541004;

2. 桂林电子科技大学广西自动检测重点实验室, 广西 桂林 541004; 3. 苏州大学机电工程学院, 江苏 苏州 215006)

**摘要:** 针对工业分拣机器人识别复杂工件慢、精确度低以及定位不准等问题, 提出一种基于深度学习的快速识别定位算法。通过工业高精度相机获取目标图像信息, 经过图像灰度化、图像滤波、Otsu 二值化处理, 再经边界像素检测算法定位并分割目标图像。运用已训练的深度卷积神经网络(CNN)对目标进行识别, 得到目标所在的位置坐标以及所属类别, 实现工业机器人分拣。实验测试中以纹路复杂的象棋为例进行定位识别, 结果表明定位算法误差小于 0.8 mm, 最快识别速度达 0.049 秒/个, 在实验环境中识别精度能保持在 98% 以上, 表明算法具备良好的准确性和稳定性。

**关键词:** 深度学习; 卷积神经网络; 视觉识别; 定位算法; 工业分拣机器人

中图分类号: TP242

文献标识码: A

文章编号: 1002-0446(2016)-06-0711-09

## Fast Visual Identification and Location Algorithm for Industrial Sorting Robots Based on Deep Learning

WU Xiru<sup>1,2,3</sup>, HUANG Guoming<sup>1,2</sup>, SUN Lining<sup>3</sup>

(1. College of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin 541004, China;

2. Guangxi Key Laboratory of Automatic Detection, Guilin University of Electronic Technology, Guilin 541004, China;

3. College of Electrical and Mechanical Engineering, Soochow University, Suzhou 215006, China)

**Abstract:** To overcome the problems of slow recognition, low accuracy and inaccurate positioning for industrial sorting robots, a fast visual identification and location algorithm based on deep convolutional neural network (CNN) is proposed. Firstly, the target image information is obtained by an industrial precision camera, and the target image is located and segmented through graying, filtering, Otsu binarization and boundary detection of the images. Secondly, the target object is identified by using a trained CNN, and its position coordinate and class are obtained. Thus, target sorting by industrial robots is realized. Finally, the Chinese chess with complex lines are taken in test experiments to verify the identification and location algorithm. Experimental results show that the locating error is lower than 0.8 mm, the fastest recognition speed can reach 0.049 seconds per target, and the identification accuracy can be kept over 98% in the experimental environment. So, the proposed algorithm has good accuracy and stability.

**Keywords:** deep learning; convolutional neural network; visual identification; location algorithm; industrial sorting robot

## 1 引言 (Introduction)

深度学习的概念于 2006 年由多伦多大学人工智能专家 Hinton 提出<sup>[1-2]</sup>, 他提出了一种快速训练深度神经网络的算法, 开启了人工智能领域研究深度学习的热潮。深度学习通用的数学模型有采用监督学习方式的深度卷积神经网络, 以及采用混合监督学习方式的堆栈式自编码网络和深度置信网络。深度学习模型已被 Bengio<sup>[3]</sup> 等人证明具有比 BP (反向传播) 神经网络、支持向量机 (SVM) 等

浅层网络更优越的非线性逼近能力和泛化能力, 在复杂模式识别场合 (如语音识别、复杂图像识别、复杂控制系统建模) 表现出强大的性能<sup>[4-7]</sup>。斯坦福大学人工智能实验室吴恩达<sup>[8]</sup> 等人采用深度学习的方法让计算机自主产生“猫”的概念。纽约大学 LeCun<sup>[9]</sup> 等人基于卷积神经网络开发的手写字符识别系统成功投入商用。2016 年 3 月, Google 的 Deep Mind 团队基于深度学习算法设计的 AlphaGo 电脑棋手在围棋挑战赛中以 4:1 的比分击败著名韩国棋手李世石<sup>[10]</sup>, 引起社会各界对深度学习技术

基金项目: 国家自然科学基金 (61603107); 广西自然科学基金 (2015GXNSFAA139297, 2016GXNSFDA380001); 广西自动检测技术与仪器重点实验室基金 (YQ16108); 智能综合自动化高校重点实验室基金 (2016); 桂林电子科技大学研究生教育创新计划 (2016YJCX04)。

通信作者: 黄国明, 414079745@qq.com 收稿/录用/修回: 2016-05-11/2016-11-02/2016-11-03

的广泛关注。

机器人通过图像检测识别技术实现视觉系统高速目标定位和识别作业<sup>[11-14]</sup>。传统检测识别技术有基于分割的方法、特征分析方法、图像识别决策分类方法、模式学习和形状匹配方法等<sup>[13]</sup>，这些方法在工业领域得到广泛应用。王丹<sup>[15]</sup>等人在人类视觉系统处理机制的基础上，利用 HOG（方向梯度直方图）算子结合 SVM 提出一种人体行为仿生识别与分类的方法，实验证明了该算法对差别较大的行为的识别效果好于常用方法，而对于相似行为的识别仍有待提高。陈守煜<sup>[16]</sup>等人利用基于相对差异

函数的可变模糊方法，建立一种用于工件识别的多传感器信息融合的可变模糊识别模型。Schmitt<sup>[17]</sup>等人针对机器人的运动策略以及离散余弦变换提出基于动态环境的工件识别算法，实验证明该算法具有良好的性能。聂海涛<sup>[18]</sup>等人针对传统人脸识别系统在复杂背景情况下不能实时准确进行人脸识别的问题，提出一种基于快速 SIFT（尺度不变特征变换）算法结合模糊控制的人脸识别方法。通过引入闭环模糊控制系统减少 SIFT 特征误匹配，提高了人脸识别率，实验表明，在复杂环境下人脸识别精度提高 10%。

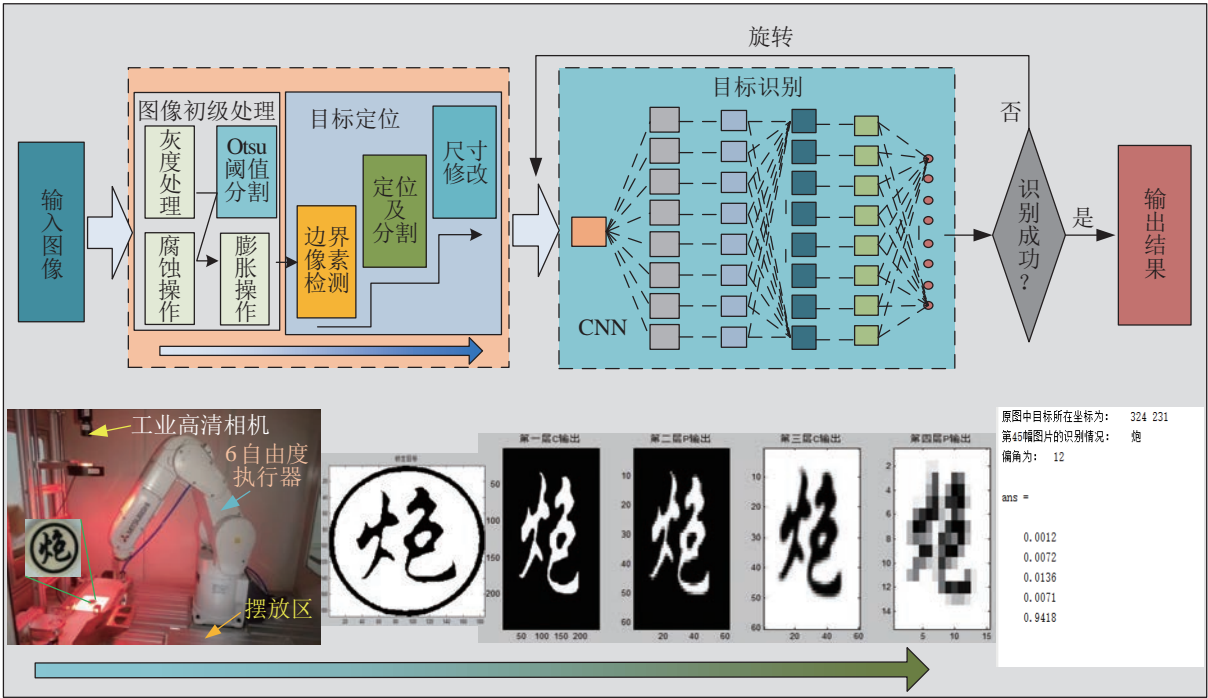


图 1 快速视觉识别与定位算法框架

Fig.1 The architecture of the fast visual identification and location algorithm

王红涛<sup>[19]</sup>等人提出一种基于边缘匹配的工件识别方法，采用 Canny 算子提取的边缘信息作为匹配特征，将改进 Hausdorff 距离作为图像匹配的相似性度量，在搜索过程中应用自适应代沟替代策略的遗传算法。实验证明该算法提高了匹配过程，能有效解决平移、遮挡及部分遮挡情况下的工件识别问题。耿庆田及其团队<sup>[20]</sup>基于颜色模型的候选火焰图像元素分类算法，先对 RGB 模型各通道求平均值获得新的图像样本，再利用 YCbCr 颜色空间建立火焰图像元素分类模型。新规则削弱了图像亮度变化产生的干扰，提高了火焰像素的检测率。

然而，现有的研究多数针对轮廓简单的工件展开，随着工业器件的复杂化，传统的定位识别算法面临着定位误差大、识别速度慢和准确率低的挑

战。目前，复杂工件的分拣仍处于人工操作阶段，为实现工业机器人对复杂工件的自动分拣，本文提出了一种基于深度学习的视觉识别与定位算法，通过结合图像处理技术，采用边界像素检测算法以及深度神经网络识别训练算法对目标进行精确定位与图像分割，应用 CNN 模型构建定位识别算法，仿真实验证明所提出算法具有良好的稳定性和准确性。

## 2 分拣机器人快速视觉算法模型（Fast visual algorithm model for the sorting robot）

本文以识别纹路较为复杂的象棋展开算法设计。根据分拣作业的需求，算法以识别定位获取工件的坐标、偏转角度、类别信息为目的，兼顾算法

速度、定位精度及识别准确率。快速视觉识别与定位算法如图 1 所示, 算法由目标定位算法和目标识别算法两部分组成。

算法的具体步骤如下:

步骤 1: 激光传感器触发传送带停止运动, 同时工业相机采集制定区域图像;

步骤 2: 将图像灰度化操作, 并进行灰度均衡化处理, 提高图片对比度;

步骤 3: 基于 Otsu 阈值分割将图像二值化处理;

步骤 4: 通过腐蚀操作滤除图像中的颗粒噪声;

步骤 5: 采用膨胀操作适度还原被腐蚀目标的边缘;

步骤 6: 通过边界像素检测确定象棋边界;

步骤 7: 从图片中分割提取出仅有象棋的局部图像并确定象棋的中心坐标;

步骤 8: 修改图像尺寸为识别作准备;

步骤 9: CNN 进行目标识别分类;

步骤 10: 识别成功则输出结果, 否则左旋 10° 并返回步骤 9。

2.1 边界像素检测定位算法

运送工件的传送带通常为白色橡胶带, 为提高目标定位的速度, 本文采用边界像素检测的方法。定位算法如图 2 所示。

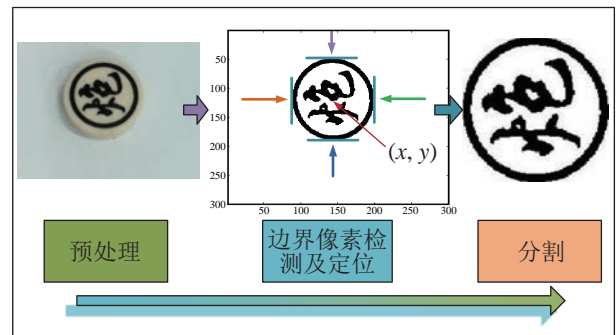


图 2 边界像素检测定位算法框架  
Fig.2 The architecture of the boundary pixel detection algorithm for localization

算法首先对输入图像预处理, 滤除传送带上其他颗粒的干扰, 保证得到的二值图像只包含象棋, 然后进行快速像素检测, 确定象棋 4 个方向的边界, 并根据象棋为圆的特点对边界点矫正, 最后实现定位及分割。

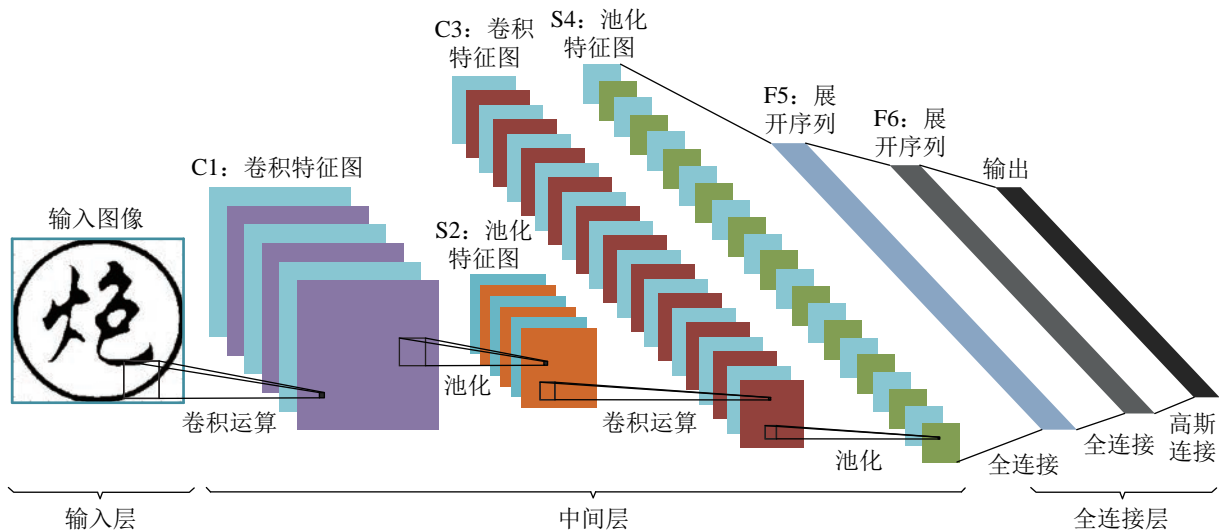


图 3 CNN 结构模型  
Fig.3 The architecture of CNN

2.2 CNN 视觉识别

CNN 起源于多层感知机, 模仿动物视觉局部提取信息的机制对输入进行特征提取。该原理是 Fukushima 在研究猫的视觉皮层时发现的<sup>[21]</sup>。在卷积神经网络中, 卷积算子 (亦称为滤波器) 只与输入图像的局部连接, 通过卷积运算提取输入特征信息<sup>[22-23]</sup>。CNN 一般为多层结构模型, 图 3 所示为 8 层结构的 CNN 模型, 分为 3 部分: 输入层、中间

层、全连接层。

(1) 输入层通常输入为二值图像或是 RGB 彩色图像, 在输入之前一般对图像进行滤波、尺寸修改等预处理操作以提高网络识别效果。

(2) 中间层是由卷积层和池化层交替组成。CNN 作为深度学习模型, 其深度主要体现在中间层中的卷积层和池化层交替的次数。当 CNN 识别图片的复杂度越大时, 中间层交替的次数也会越

多, 类似于 GoogLeNet 一类的大型卷积网络, 用于学习超多分类图片会使用到十几层卷积层和池化层。

卷积操作如图 4 所示, 其中卷积算子  $W$  与输入图像  $X$  的局部连接, 定义该局部为  $X_{\text{small}}$ , 则输出为  $f = \sigma(WX_{\text{small}} + b)$ ,  $\sigma$  一般选择 Sigmoid 函数。卷积算子相当于一种滤波器, 能够实现图像滤波及边缘增强等操作。输入图像被分为多个子图后以同一卷积算子处理, 这种将全连接改为局部连接的权值共享方式成功降低了网络训练难度。

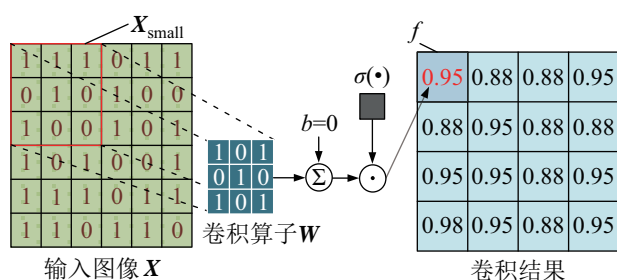


图 4 卷积运算

Fig.4 Convolutional operation

当输入为高分辨率图片时, 网络运行会占用大量计算资源, 降低了网络的实时性, 同时输入图像包含太多特征易导致网络学习出现过拟合现象。为了解决这些问题, 本文在卷积的基础上进行了池化运算 (如图 5 所示), 有效降低了特征维度, 压缩了数据量, 提高了网络的学习效率以及实时性能。

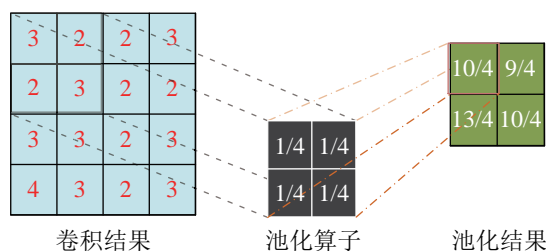


图 5 池化运算

Fig.5 Pooling operation

(3) 全连接层一般为三层结构的浅层分类器, 比如 SVM、SoftMax 多分类器等, 包含输入层、隐层、输出层。输入层为 CNN 中间层中最后的池化层所得特征图拉长得到的列向量; 隐层为一层非线性映射层; 输出层也是 CNN 的结果输出层, 一般输出为目标分类的分数值, 并取最高分为识别结果。

### 2.3 深度 CNN 训练

为使网络识别速度快且精度高, 在构造样本时先对样本进行加工, 去掉样本中无关的特征。在不同环境光下获取样本图片, 经边缘像素检测算法锁定图片中的象棋, 通过像素区域选择去除象棋文字

外围的圆圈, 将样本整理成  $250 \times 250$  像素的二值图像。通过 CNN 对人为规则化的样本进行监督学习, 获取到最有区分度的特征信息。

在训练过程中, 通过有监督的均方差反向传播算法修正网络参数<sup>[24-25]</sup>, 即最小均方差的方法 (MMSE)。对于样本数据为  $D = (x_i, y_i^*)_{1 \leq i \leq m}$  的网络, 均方误差 (MSE) 可表达为

$$J(W, b; x, y^*) = \frac{1}{2} \|h_{w,b}(x) - y^*\|^2 \quad (1)$$

其中  $h_{w,b}(x)$  为网络输出值,  $h$  为激活函数, 选择 Sigmoid 函数,  $W$  为网络的权值矩阵,  $b$  为每层偏置  $b$  构成的偏置矩阵,  $x$  为输入样本矩阵,  $y^*$  为期望输出。最小化非凸函数  $J(W, b; x, y^*)$  便可得到网络的最优解。通常采用梯度下降的方法对网络迭代运算更新权值:

$$W_{ij}^l = W_{ij}^l - \alpha \frac{\partial}{\partial W_{ij}^l} J(W, b; x, y^*) \quad (2)$$

$$b_i^l = b_i^l - \alpha \frac{\partial}{\partial b_i^l} J(W, b; x, y^*) \quad (3)$$

其中  $\alpha$  为学习率。传统梯度下降的方法容易使网络收敛于局部最优解且训练速度慢, 为获得理想的收敛结果, 本文采用随机梯度下降算法。该算法具有收敛速度快、多次迭代能够获得全局最优解的特点。基于随机梯度下降的误差反向传播算法训练网络过程如图 6 所示, 其中运算符  $\odot$  为互相关运算, 运算符  $\circ$  表示元素相乘运算。

### 2.4 识别算法实现

本文所设计算法模型中 CNN 为 12 层结构, 分别为: 输入层 ( $250 \times 250 \times 1$ ), 一级卷积层 C1 ( $3 \times 3 \times 1$ ), 一级池化层 P1 ( $2 \times 2$ ), 二级卷积层 C2 ( $5 \times 5 \times 1$ ), 二级池化层 P2 ( $2 \times 2$ ), 三级卷积层 C3 ( $5 \times 5 \times 1$ ), 三级池化层 P3 ( $2 \times 2$ ), 四级卷积层 C4 ( $5 \times 5 \times 1$ ), 四级池化层 P4 ( $2 \times 2$ ), Softmax 分类器输入层、隐层、输出层。图像预处理及定位部分采用 Matlab 图像处理库函数实现, 采用 Deep Learning Toolbox 的操作函数实现 CNN, 并在 CPU 为 Intel® Core™ i3-2100 CPU @3.10 GHz、内存为 12G 的台式计算机上训练网络。算法训练过程中, 以识别 5 种象棋棋子为例, 每个象棋对应 240 张样本, 总共 1200 张样本图像, 并采用可视化的方式分析网络性能。如图 7、8 所示, 以学习“马”和“炮”为例子给出输入层和中间层的特征图像。在训练 CNN 的过程中, 拟以不同的学习率  $\alpha$  进行。MSE 下降曲线如图 9 所示, 其中横轴为网络训练时迭代的样本数 (迭代次数为 10 次), 纵轴表示 MSE, 图 9 中分别给出了学习率  $\alpha$  取不同值的训练效果。



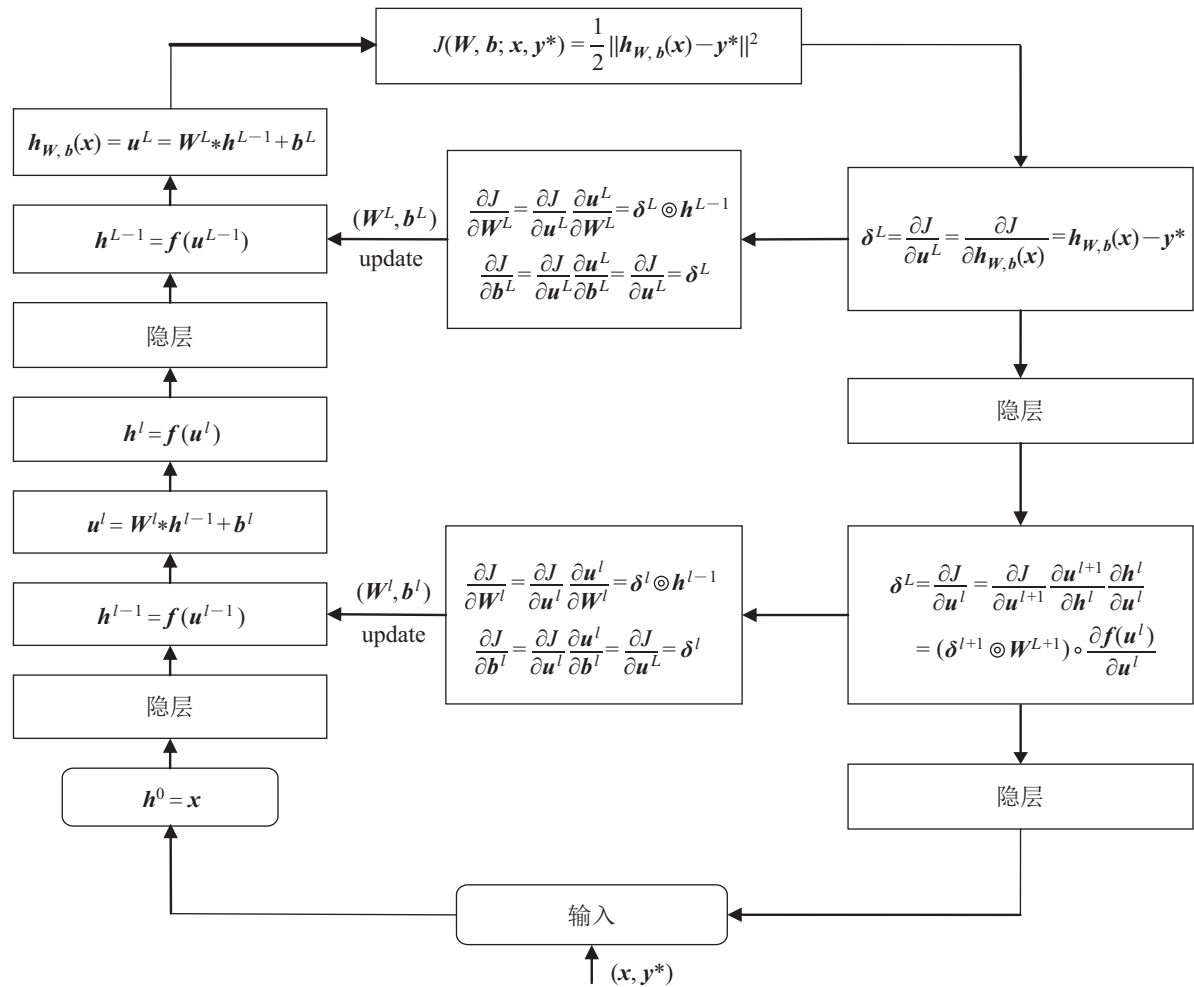


图 6 基于 MMSE 的 CNN 训练示意图  
Fig.6 The CNN training based on MMSE

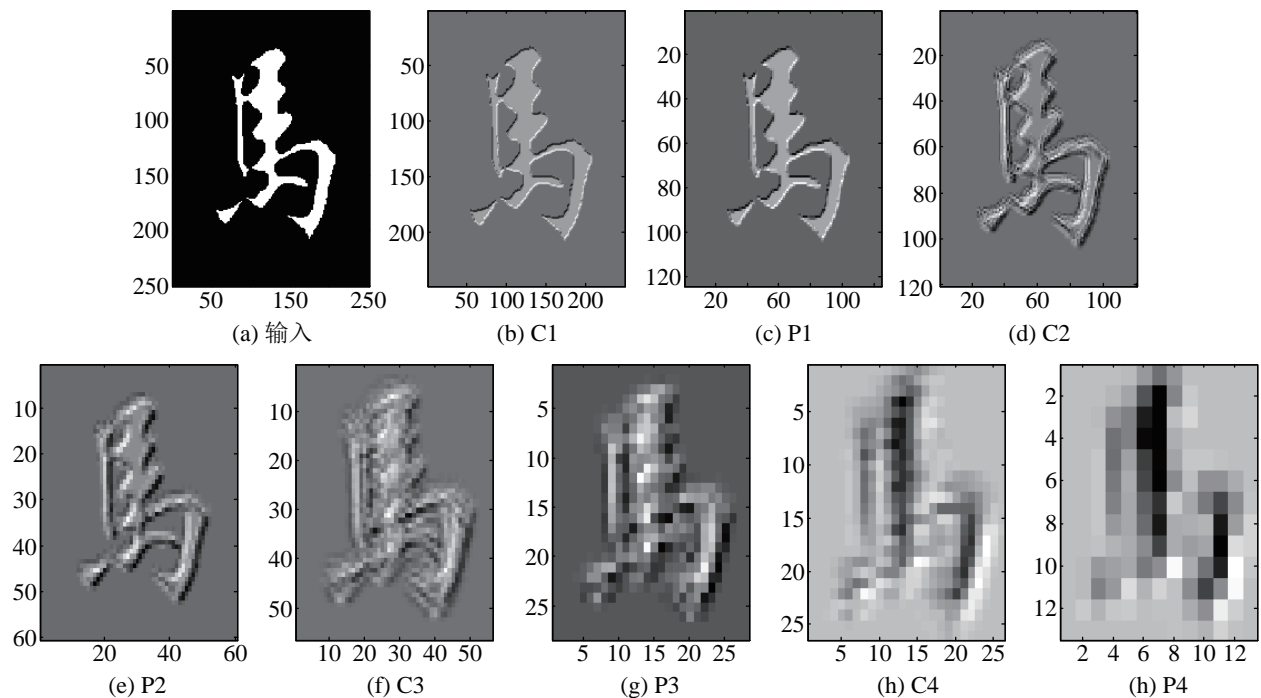


图 7 马的训练特征图  
Fig.7 The feature map of MA

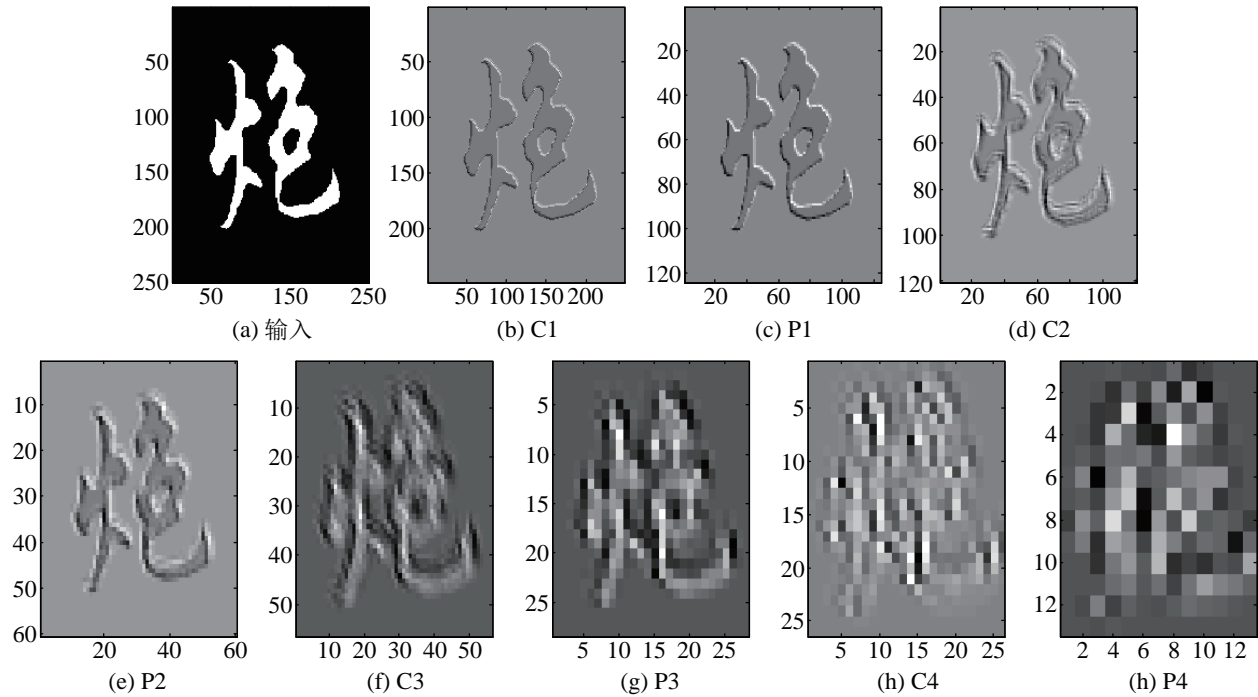


图 8 炮的训练特征图

Fig.8 The feature map of PAO

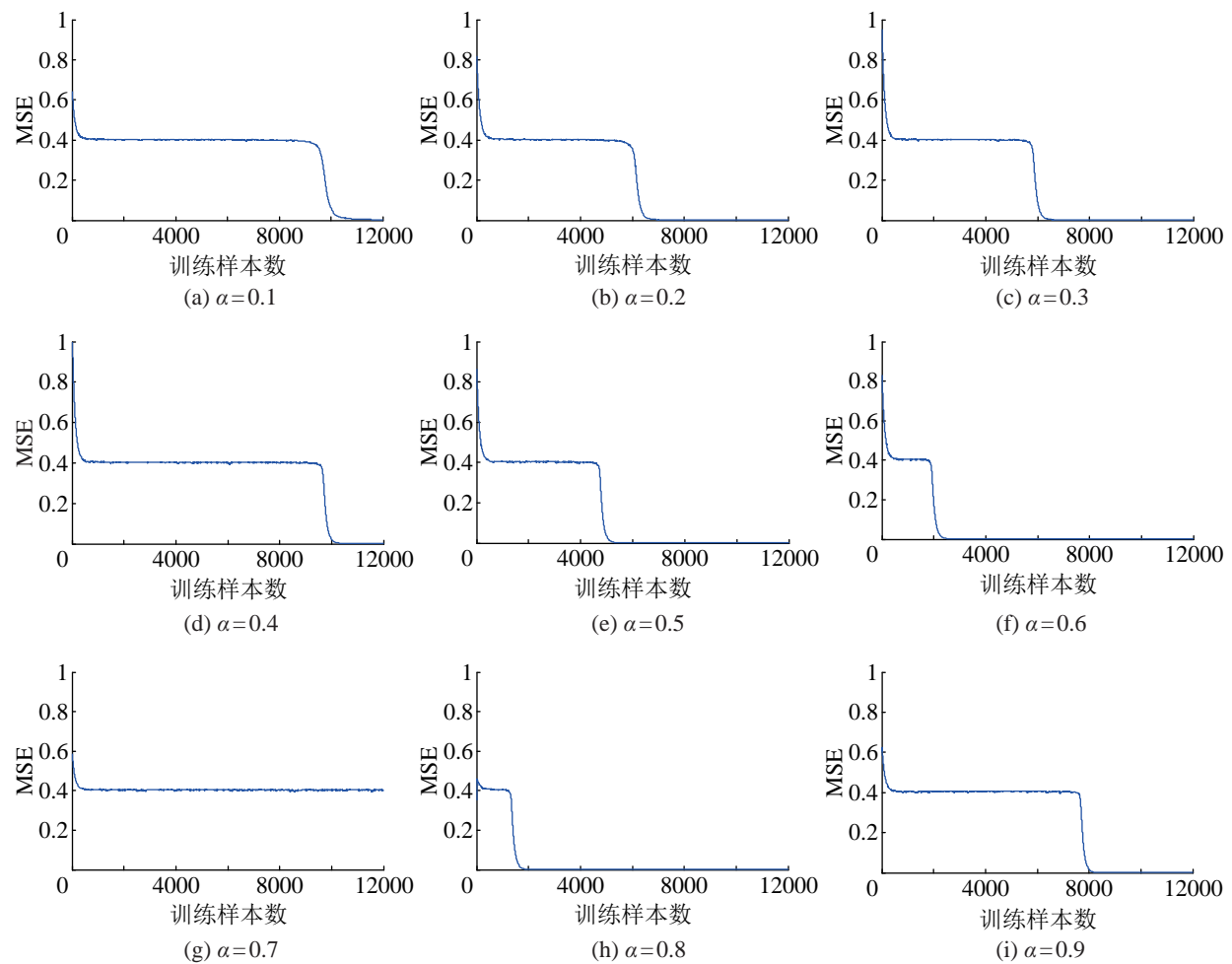


图 9 均方误差下降曲线图

Fig.9 The decline graph of MSE

随着 MSE 下降, 特征图中文字的边缘越加清晰, 网络收敛至最优时 (误差跌至 0.1 以下), 样本图像中的文字特征得到大幅度增强. 如图 7(d) 所示, 特征图具有较强的立体感, 边缘信息突出, 表明第 2 级卷积算子经调整权值后有效实现了特征提取. 根据图 9 可知, 训练过程中 MSE 的局部最小值约为 0.4, 当  $\alpha = 0.8$  时网络最快跳出局部最小值 (约迭代至 1700 个样本时), 并且迅速收敛于全局最小. 当网络训练迭代至 2100 个样本时, 基本趋近于 0 ( $MSE < 0.01$ ), 得到理想的训练效果.

3 试验及分析 (Experiments and analysis)

3.1 试验

利用基于视觉的六轴柔性工业分拣机器人对算法进行识别定位测试, 测试平台如图 10 所示. 其中该测试平台中的计算机配置为: 处理器为 Intel® Core™ i3-2100 CPU @3.10 GHz、内存为 12 G. 该

机器人平台主要由小型工业传送带、工业摄像机、6 自由度机械臂组成. 测试中, 机器人的工作流程如下: 通过工件触发激光传感器将传送带停止, 使工件置于摄像头视觉区域, 同时工业摄像机采集图像, 经算法处理后获得工件的所属类别、坐标位置及其旋转角度, 并将属性发送至机械臂控制端, 机械臂快速将物品拾取摆放至预定位置, 若无法判断工件类别, 则放弃分拣该工件.

工业生产车间中分拣机器人作业光环境受到诸多因素的影响, 比如车间窗户投入的户外光、车间生产照明灯以及电焊的弧光等. 为提高视觉识别算法的鲁棒性, 实验测试在 LED 红色光源 (光源位置在图 9 中右上部分高清工业相机竖直下方约 15 cm 处) 产生的稳定光环境下进行, 通过打光直射能有效提高图像的对比度, 有利于分割及定位目标. 通过与文 [14,26-28] 对比试验定位结果进行算法分析及讨论.

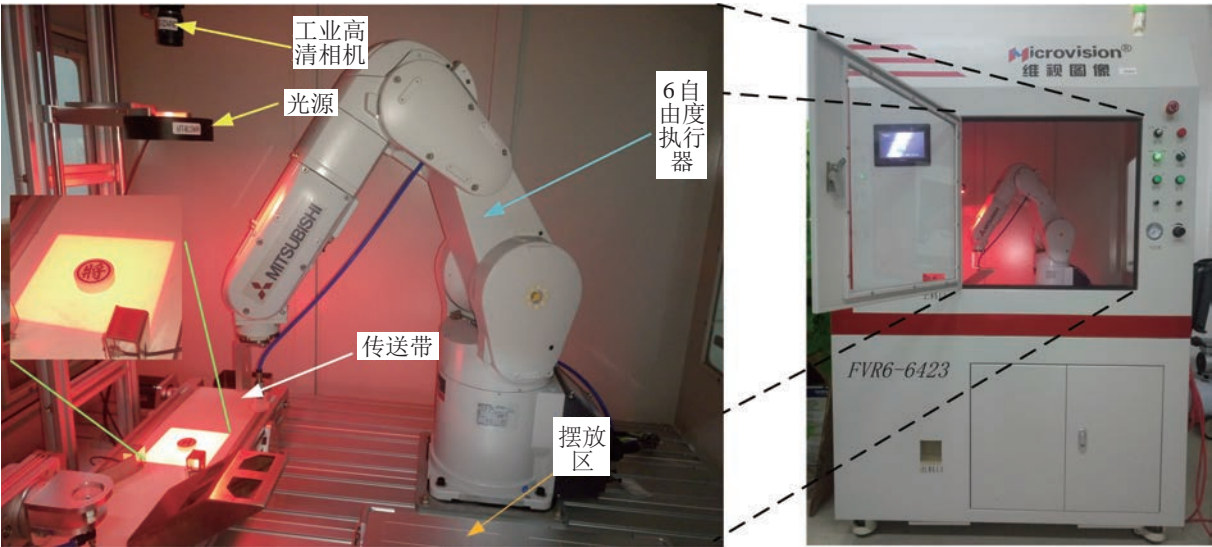


图 10 工业分拣机器人测试平台  
Fig.10 The test platform of the industrial sorting robot

实验结果以两部分展示, 一部分是定位测试, 结果如表 1 所示, 其中以计算物理坐标以及实测物理坐标的距离作为定位误差, 无偏角平均识别时间为象棋旋转角度为 0 的状态下算法识别所需的平均时间; 一部分是识别测试, 结果如表 2 所示. 定位测试主要考察目标定位结果的误差, 其中相机的视野为  $1000 \times 1000$  像素的正方形区域, 对应物理尺寸为  $100 \text{ mm} \times 100 \text{ mm}$  的区域 (图 9 中红光直射区域), 工件 (象棋) 为直径等于  $28 \text{ mm}$  的圆形木块. 实验以识别测试 6 种偏角放置物品为例考核算法的识别时间以及检测、识别准确率. 实验以每种偏角

进行 300 次, 放置的棋子为任意类型.

3.2 试验结果分析

表 1 的测试结果表明, 在实验环境下算法能快速定位象棋所在的位置, 定位误差控制在  $0.8 \text{ mm}$  以内, 该效果明显优于文 [14] 所得结果, 说明边界像素检测修正的方法能够胜任象棋定位的任务. 相对于 K 均值等传统聚类算法, 边界像素检测算法通过合理的预处理后无需迭代运算便能得到定位结果, 有利于提高算法的实时性能. 测试结果证明了算法设计的合理性和有效性.

识别测试如表 2 所示, 算法从获取照片到识别

表 1 定位测试结果  
Tab.1 The result of locating test

序号	图像像素坐标		计算物理坐标		实测物理坐标		定位误差
	col /pixel	row /pixel	X /mm	Y /mm	X' /mm	Y' /mm	R /mm
1	624	731	62.4	73.1	62.9	73.5	0.640
2	413.5	564	41.35	56.4	41.3	56.8	0.403
3	300	610.5	30.0	61.05	30.2	61.5	0.452
4	555.5	419.5	55.55	41.95	55.7	42.4	0.474
5	634	321	63.4	32.1	64.1	31.8	0.762
6	311.5	424.5	31.15	42.45	31.5	42.1	0.495
7	356	456.5	35.6	45.65	35.1	45.3	0.610
8	528.5	386	52.85	38.6	52.3	38.1	0.743
9	712.5	623.5	71.25	62.35	70.8	62.4	0.453
10	428	290.5	42.8	29.05	43.2	29.6	0.680

目标的最短时间达到 0.049 s，识别准确率不低于 98%，说明基于 CNN 的识别方法优于基于 BP 神经网络的识别方法<sup>[26]</sup>、基于 Hough 变化的识别方法<sup>[27]</sup>和基于年轮统计的识别方法<sup>[28]</sup>，体现了算法识别精度高、抗干扰能力强的特点，适合用于高稳定性要求的工业场合。通过检测准确度和识别准确率两项指标说明只要算法能够检测到目标，就能够快速准确地识别目标，表明对训练样本以及测试样本进行人为规则化处理有利于提高算法的学习和识别效果。由测试结果可知，本文的算法模型能满足当前工业分拣机器人的视觉作业需求。

表 2 识别测试结果  
Tab.2 The result of recognition test

偏转角 /(°)	平均识 别时间/s	正确检 测个数	检测准 确度/%	正确识 别个数	识别准 确率/%
0	0.049	300	100	300	100
30	0.063	299	99.67	299	100
60	0.095	297	99	297	100
90	0.134	294	98	293	99.66
120	0.189	296	98.67	292	98.65
150	0.223	293	97.67	290	98.98
180	0.262	295	98.33	292	98.98

4 总结 (Conclusion)

深度学习在工业、生活等领域具有优越的特征抽取性能和巨大的发展潜力。本文采用深度卷积网络和图像处理技术结合的方法实现分拣机器人的快速视觉识别与定位，该算法能够快速准确地识别与定位复杂目标物体，具有良好的稳定性。通过仿真实验证实了本文算法的有效性和准确性。

参考文献 (References)

[1] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527-1554.

[2] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.

[3] LeCun Y, Bengio Y, Hinton G E. Deep learning[J]. Nature, 2015, 521(7553): 436-444.

[4] Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position[J]. Biological Cybernetics, 1980, 36(4): 193-202.

[5] Sanchez-Riera J, Hua K L, Hsiao Y S, et al. A comparative study of data fusion for RGB-D based visual recognition[J]. Pattern Recognition Letters, 2016, 73(10): 1-6.

[6] Wang A R, Lu J W, Cai J F, et al. Large-margin multi-modal deep learning for RGB-D object recognition[J]. IEEE Transactions on Multimedia, 2015, 17(11): 1887-1898.

[7] Wu D, Wu J S, Zeng R, et al. Kernel principal component analysis network for image classification[J]. Journal of Southeast University, 2015, 31(4): 469-472.

[8] Markoff J. How many computers to identify a cat? 16000[N]. New York Times, 2012-06-25.

[9] LeCun Y, Boser B, Denker J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural Computation, 1989, 1(4): 541-551.

[10] 田渊栋. 阿法狗围棋系统的简要分析 [J]. 自动化学报, 2016, 42(5): 671-675.

Tian Y D. A simple analysis of AlphaGo[J]. Acta Automatica Sinica, 2016, 42(5): 671-675.

[11] 倪鹤鹏, 刘亚男, 张承瑞, 等. 基于机器视觉的 Delta 机器人分拣系统算法 [J]. 机器人, 2016, 38(1): 49-55.

Ni H P, Liu Y N, Zhang C R, et al. Sorting system algorithms based on machine vision for Delta robot[J]. Robot, 2016, 38(1): 49-55.

[12] 郝明. 机器视觉在机器人杂乱工件分拣中的应用 [D]. 沈阳: 沈阳工业大学, 2015.



- Hao M. An application of machine vision on sorting clutter workpiece by robot[D]. Shenyang: Shenyang University of Technology, 2015.
- [13] 王耀南, 陈铁健, 贺振东, 等. 智能制造装备视觉检测控制方法综述[J]. 控制理论与应用, 2015, 32(3): 273-286.
- Wang Y N, Chen T Z, He Z D, et al. Review on the machine vision measurement and control technology for intelligent manufacturing equipment[J]. Control Theory and Applications, 2015, 32(3): 273-286.
- [14] 王殿君. 基于视觉的中国象棋棋子识别定位技术[J]. 清华大学学报: 自然科学版, 2013, 53(8): 1145-1149.
- Wang D J. Recognition and positioning technique of Chinese chess based on vision[J]. Journal of Tsinghua University: Science and Technology, 2013, 53(8): 1145-1149.
- [15] 王丹, 张祥合. 基于 HOG 和 SVM 的人体行为仿生识别方法[J]. 吉林大学学报: 工学版, 2013, 43(S1): 489-492.
- Wang D, Zhang X H. Biomimetic recognition method of human behavior based on HOG and SVM[J]. Journal of Jilin University: Engineering and Technology Edition, 2013, 43(S1): 489-492.
- [16] 陈守煜, 胡吉敏. 可变模糊方法及其在工件识别中的应用[J]. 系统工程与电子技术, 2006, 28(9): 1325-1328.
- Chen S Y, Hu J M. Variable fuzzy method and its application in parts recognition[J]. Systems Engineering and Electronics, 2006, 28(9): 1325-1328.
- [17] Schmitt R, Cai Y. Recognition of dynamic environments for robotic assembly on moving workpieces[J]. International Journal of Advanced Manufacturing Technology, 2014, 71(5-8): 1359-1369.
- [18] 聂海涛, 龙科慧, 马军, 等. 基于快速 SIFT 算法和模糊控制的人脸识别[J]. 吉林大学学报: 工学版, 2016, 46(2): 549-555.
- Nie H T, Long K H, Ma J, et al. Face recognition based on fast scale invariant feature transform algorithm and fuzzy control[J]. Journal of Jilin University: Engineering and Technology Edition, 2016, 46(2): 549-555.
- [19] 王红涛, 傅卫平, 康业娜. 工件图像识别的边缘匹配方法研究[J]. 仪器仪表学报, 2008, 29(5): 986-991.
- Wang H T, Fu W P, Kang Y N. Study of edge matching approach to workpiece image recognition[J]. Chinese Journal of Scientific Instrument, 2008, 29(5): 986-991.
- [20] 耿庆田, 于繁华, 赵宏伟, 等. 基于颜色特征的火焰检测新算法[J]. 吉林大学学报: 工学版, 2014, 44(6): 1787-1792.
- Geng Q T, Yu F H, Zhao H W, et al. New algorithm of flame detection based on color features[J]. Journal of Jilin University: Engineering and Technology Edition, 2014, 44(6): 1787-1792.
- [21] Fukushima K, Miyake S. Neocognitron: Self-organizing network capable of position-invariant recognition of patterns[C]//5th International Conference on Pattern Recognition. Piscataway, USA: IEEE, 1980: 459-461.
- [22] Kavukcuoglu K, Sermanet P, Breuer Y L, et al. Learning convolutional feature hierarchies for visual recognition[C]//24th Annual Conference on Neural Information Processing Systems. Red Hook, USA: Curran Associates Inc., 2010.
- [23] Jaderberg M, Simonyan K, Vedaldi A, et al. Reading text in the wild with convolutional neural networks[J]. International Journal of Computer Vision, 2016, 116(1): 1-20.
- [24] LeCun Y, Bottou L, Orr G, et al. Efficient backProp[M]//Neural Networks: Tricks of the Trade. Berlin, Germany: Springer, 2012: 9-48.
- [25] 谢建斌. 视觉机器学习 20 讲[M]. 北京: 清华大学出版社, 2015: 170-184.
- Xie J B. Twenty lectures of machine learning for vision[M]. Beijing: Tsinghua University Press, 2015: 170-184.
- [26] 王春丽. 中国象棋嵌入式视觉识别算法和程序开发[D]. 北京: 北方工业大学, 2010.
- Wang C L. Design of Chinese chess recognition algorithm and program based on embedded vision system[D]. Beijing: North China University of Technology, 2010.
- [27] 莫妙桃. 基于 DSP 的智能象棋机器人视觉图像采集与识别研究[D]. 北京: 北方工业大学, 2009.
- Mo M T. Study on vision image grabbing system based on DSP and character recognition method for Chinese chess playing robot[D]. Beijing: North China University of Technology, 2009.
- [28] 朱一峰. 象棋机器人视觉识别算法研究[J]. 江汉大学学报: 自然科学版, 2013, 41(3): 51-56.
- Zhu Y F. Visual recognition algorithm of Chinese chess robot[J]. Journal of Jiangnan University: Natural Science Edition, 2013, 41(3): 51-56.

## 作者简介:

- 伍锡如 (1981-), 男, 博士, 副教授. 研究领域: 非线性系统控制, 神经网络, 机器人控制.
- 黄国明 (1992-), 男, 硕士生. 研究领域: 机器学习, 机器视觉, 深度学习.
- 孙立宁 (1964-), 男, 博士, 教授. 研究领域: 纳米级微驱动及微操作机器人, 工业机器人技术, 医疗机器人, 仿人手臂及机器人机构与控制.