

基于 OpenCV 和 TensorFlow 的自动分类识别系统设计

刘 斌¹, 程方毅¹, 龚德文²

(1. 华南理工大学高分子先进制造教育部重点实验室, 广州 510641; 2. 广东昌恒智能装备科技有限公司, 广东 东莞 519000)

摘 要:介绍了基于 OpenCV 和 TensorFlow 的自动分类识别系统的构建及其关键技术,包括图像处理、生成数据集、构建训练和保存卷积神经网络模型。首先对原始图像进行预处理,利用 OpenCV 图像处理技术从图像中提取出目标区域,运用灰度化、滤波降噪、二值化、形态学运算、连续轮廓检测及矩形边界框方法,研究结果表明,该流程在物流工业环境下能分割出较为完整的物体区域;然后,利用 TensorFlow 将获取的图像保存为 tfrecord 数据集,构建、训练轻量级卷积神经网络模型,该模型包含 4 个卷积层和 3 个全连接层,每个卷积层后面连接着一个最大池化层,训练结果被保存为离线 pb 模型,结果表明,随着迭代次数的增加,模型的准确率趋于稳定;最后将训练好的模型文件拷贝至工控机上,并在可移植平台 QT 平台上进行 GUI 开发,实现自动分拣识别。整个自动分类识别系统采用模块式开发,并且该模型还有较好的扩展性,可以实现物流自动化输送线上多种类商品的自动识别。

关键词:机器视觉;OpenCV;图像分类;自动识别;卷积神经网络;TensorFlow

中图分类号: TP391.41;F253.4

文献标志码: A

文章编号: 1003-7241(2026)10-0103-07

Design of automatic classification and recognition system based on OpenCV and TensorFlow

Liu Bin¹, Cheng Fangyi¹, Gong Dewen²

(1. The Key Laboratory of Polymer Advanced Manufacturing, Ministry of Education (South China University of Technology), Guangzhou 510641, China; 2. Guangdong Changheng Intelligent Technology Co., Ltd., Dongguan 519000, Guangdong, China)

Abstract: This article introduces the construction and key technologies of an automatic classification and recognition system based on OpenCV and TensorFlow, including image processing, generating datasets, constructing and training convolutional network models, and saving them. Firstly, the original image is preprocessed by using OpenCV image processing technology to extract the target area from the image. Grayscale transformation, filtering and denoising, binarization, morphological operations, continuous contour detection, and rectangular bounding box methods are applied. The research results show that this process can segment relatively complete object areas in the logistics industry environment; Then, using TensorFlow, the obtained images are saved as a tfrecord dataset to construct and train a lightweight convolutional neural network model. The model consists of 4 convolutional layers and 3 fully connected layers, with each convolutional layer followed by a max pooling layer. The training results are saved as an off line PB model, and the results show that as the number of iterations increases, the accuracy of the model tends to stabilize; Finally, it copies the trained model file to the industrial computer and develops a GUI on the portable QT platform to achieve automatic sorting and recognition. The entire automatic classification and recognition system adopts modular development, and the model has good scalability, which can achieve automatic recognition of multiple types of goods on logistics automation conveyor lines.

Keywords: machine vision; OpenCV; image classification; automatic recognition; CNN; TensorFlow

物料信息采集是自动化物流系统中必不可少的一个环节。信息的采集需要快速、准确,同时为了减少中间环节和人工干预,应尽量利用物流包装上的现有信息进行采集。目前,物流系统中大多采用人工录入、条码或射频识别等方法采集物料信息,但是,在一些食品、医药等行业,货物种类有限,且外表面没有张贴条码或射频卡等信息载体,因此,研发一种运算速度快、识别率高的自动分类识别系统直接识别货物类别信息十分必要^[1]。

图像分类是计算机视觉领域中的基本任务之一,它通

过提取图像的判别特征将不同的图像区分开来,在工业领域有着重要的应用。近年来,由神经网络发展起来的深度学习在图像分类上取得了显著的成绩,其最突出的特点就是处理图片时的智能化。与传统图像处理技术不同,在基于深度学习的图像分类技术中,不需要进行复杂的图像处理过程就可以达到自己的目的^[2]。

通过利用 OpenCV 和 TensorFlow 实现对物流自动化输送线上产品的自动识别。其中,OpenCV 是一个开源的、跨平台的计算机视觉库,它由一系列 C 函数和少量

收稿日期:2026-01-06;录用日期:2026-02-24

作者简介:刘 斌(1969—),男,工学博士,教授,研究方向为模具 CAD、CAE、CAM 和材料成型装备及数控技术。

引用本文:刘斌,程方毅,龚德文.基于 OpenCV 和 TensorFlow 的自动分类识别系统设计[J].自动化技术与应用,2026,45(10):103-108,183. (Liu Bin, Cheng Fangyi, Gong Dewen. Design of automatic classification and recognition system based on OpenCV and TensorFlow[J]. Techniques of Automation and Applications, 2026, 45(10): 103-108, 183.)

C++类构成,轻量且高效,同时提供了 Python、Ruby、Matlab 等语言的接口。OpenCV 实现了图像处理和计算机视觉方面的很多通用算法,被广泛应用于人机互动、物体识别、图像分割、运动跟踪、机器视觉及汽车自动驾驶等领域。TensorFlow 是谷歌开源的一个计算框架,具有兼容性好、支持并行运算以及支持细粒度网络层等优点。该计算框架可以很好地实现各种深度学习算法,被广泛应用于图像识别、语音识别及自然语言处理等场景。

1 自动分类识别系统框架

自动分拣识别流程为:当输送带上的产品触发光电开关时,通过相机对产品进行拍照得到原始图像信息;运用计算机图像处理技术对原始图像进行处理,获取产品的相关信息;经过信息采集后,根据产品信息,通过 PLC 控制产品流向指定的分拣道口。分拣信息的识别速度和准确率对物流自动化输送系统的正常运行至关重要^[3-4]。自动分类识别系统,其应用框架流程图如图 1 所示,主要可

分为 3 步。

第 1 步 获取数据集。利用工业相机自动采集图像数据,将数据集按比例划分为训练集和验证集;然后使用 tfrecord 数据格式存储图像数据。tfrecord 格式可以统一不同的原始数据格式,对不同的属性实现更加有效的管理,并且极大地节省内存开销。一个保存了 1 000 个 100×100 像素图像数据的 tfrecord 文件,大小仅为 28.6 MB。

第 2 步 得到模型文件。从 tfrecord 文件中批量读取数据,输入卷积神经网络,训练模型,将训练得到的神经网络模型持久化,方便离线调用。在将网络模型保存为离线文件时,只需要知道如何从神经网络的输入层经过前向传播计算得到输出层即可。

第 3 步 自动分类识别。工业相机自动获取图像,将原始图像数据经过数字图像处理输入离线模型文件,经过前向传播计算得到输出结果。输出结果为一维矩阵,其每一列的值代表着图像为每种类别的概率,选取置信度最大的类别为识别结果。

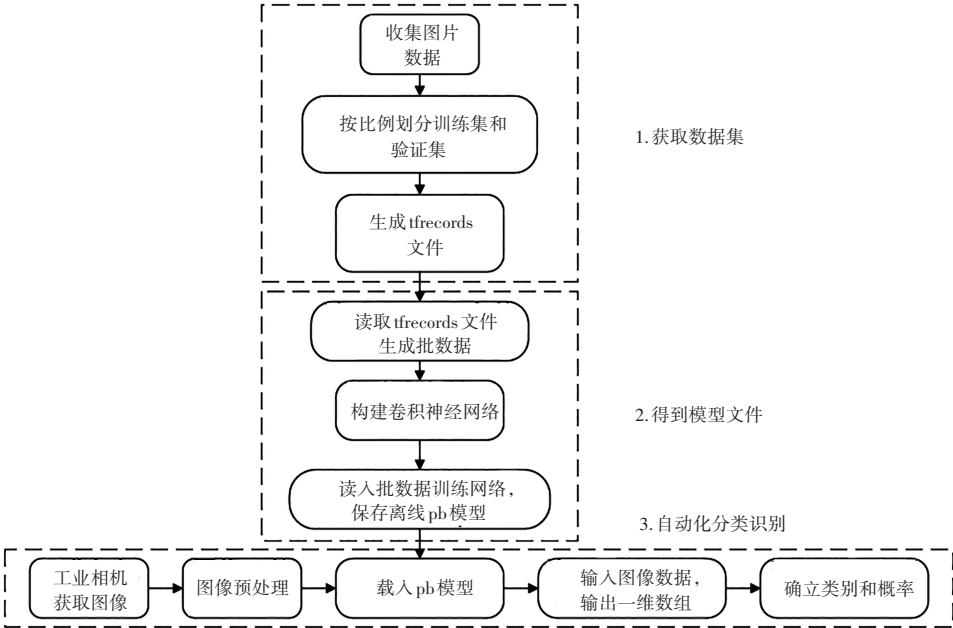


图 1 自动分类识别系统的应用框架流程

Fig. 1 Application framework process of automatic classification and recognition system

2 图像处理步骤

采用 OpenCV 对原始图像数据进行处理,图像处理基本步骤为获取图像、图像预处理、ROI 提取、载入模型和输出结果。

2.1 获取图像

图像识别第 1 步为获取图像,训练模型和进行图像分类识别也需要图像。在物流自动化输送线上,当物体触发光电开关时,光电开关向相机发送脉冲信号,在光源的照射下,相机通过光学镜头将产品表面成像于相机传感器上,光信号先转换成电信号,进而转换成计算机能处理的数字信号,从而完成产品表面图像的采集^[5]。采集到的原始图像实例如图 2 所示。

2.2 图像预处理

在图像分类过程中,输入图像的质量会在一定程度上影响分类结果,通常,也需要图像预处理来消除图像中的不相关信息,加强可用信息^[6]。一般需要将目标图像从背景中提取出来,灰度化、降噪及二值化是为了降低后续步骤的难度。

2.2.1 灰度化

对于计算机来说,经过灰度化处理后的图像可以极大减少处理图像过程中所占用的资源。对于图像的灰度化,一般有最大值法、平均法和加权平均法 3 种方法。通过采用加权平均法,调用 OpenCV 中的 `cv::cvtColor` 函数,将相机获取的彩色三通道图像转换为灰度化单通道图像。进行灰度化处理后的图像实例如图 3 所示。

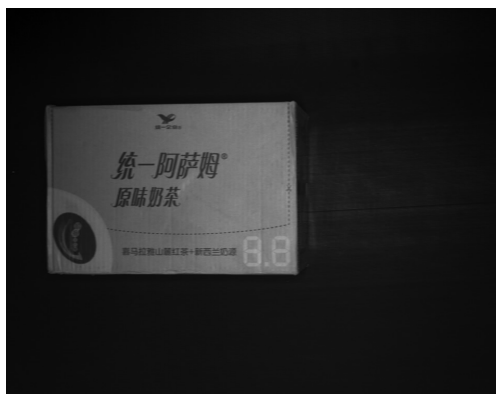


图 2 原始图像实例

Fig. 2 Example of the original image

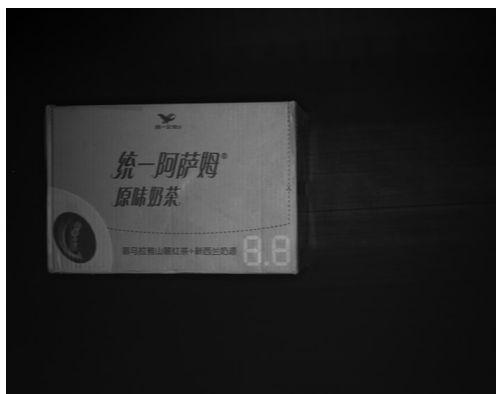


图 3 灰度化处理后的图像实例

Fig. 3 Example of an image after grayscale processing

2.2.2 图像滤波

由于现场环境、CCD 图像光电转换、传输电路及电子元件都会使图像产生噪声,这些噪声降低了图像的质量,从而对图像的处理和分析带来不良影响,所以要图像进行滤波以去噪。常见的图像滤波方式有中值滤波、均值滤波和高斯滤波。通过使用中值滤波^[7],调用 `cv::medianBlur` 函数对图像进行中值滤波处理,这对消除椒盐噪声非常有用。

2.2.3 二值化

图像的二值化,是指将一幅灰度图像转换为黑白的二值图像。其基本思想是确定一个阈值,然后把每个像素点的灰度值和阈值相比较,根据比较结果把该像素划分为两类:目标或背景,从而得到相应的二值图像^[8]。

关于阈值的确定方法,目前比较常用的有固定阈值法、自适应阈值法和多区域阈值法等。考虑到部分产品外表面颜色与黑色输送带之间灰度值区别不明显,采用自适应阈值法处理。自适应阈值法会将每个像素的值与邻域的平均值进行比较,如果像素的值与它的局部平均值差别较大,则会被当成异常值在阈值化过程中剔除。调用 `cv::adaptiveThreshold` 函数对灰度化图像进行自适应阈值化处理,将物体用白色表示,灰度值为 255,背景用黑色表示,灰度值为 0。处理后的图像实例如图 4 所示。



图 4 自适应阈值化处理后的图像实例

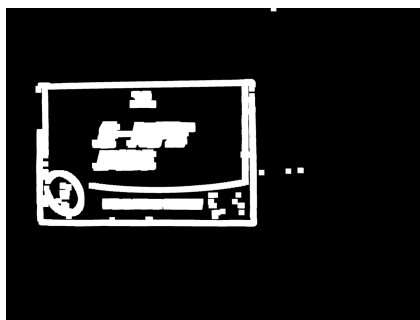
Fig. 4 Example of image after adaptive thresholding processing

2.2.4 形态学运算

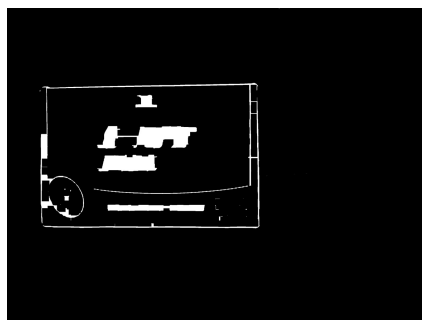
数学形态学是指以形态结构为基础对图像进行分析的数学工具,其基本运算有 4 个:膨胀、腐蚀、开运算和闭运算。其中,闭运算的定义是对图像先膨胀后腐蚀,运算效果为可以把错误分裂成小碎片的物体连接起来。调用 `cv::dilate` 函数对图像进行一次膨胀运算处理,结果如图 5(a) 所示,此时图像变粗,外围轮廓离散点连成一条线;调用 `cv::erode` 函数对经过膨胀后的图像进行一次腐蚀运算处理,结果如图 5(b) 所示,此时图像轮廓变细。

2.2.5 轮廓查找

进行图像分析的目的是提取出目标物体。生成二值图像后,可以提取出图像中连续区域的轮廓^[9]。调用 `cv::findContours` 函数可以检测所有连续外部轮廓,结果如图 6(a) 所示。通过 `cv::contourArea` 函数估算每个连续外部轮廓的面积,找到面积最大的连续外部轮廓,结果如图 6(b) 所示。



(a) 进行一次膨胀运算处理



(b) 对膨胀运算后的图像进行腐蚀运算处理

图 5 形态学运算图像处理实例

Fig. 5 Example of morphological operation image processing

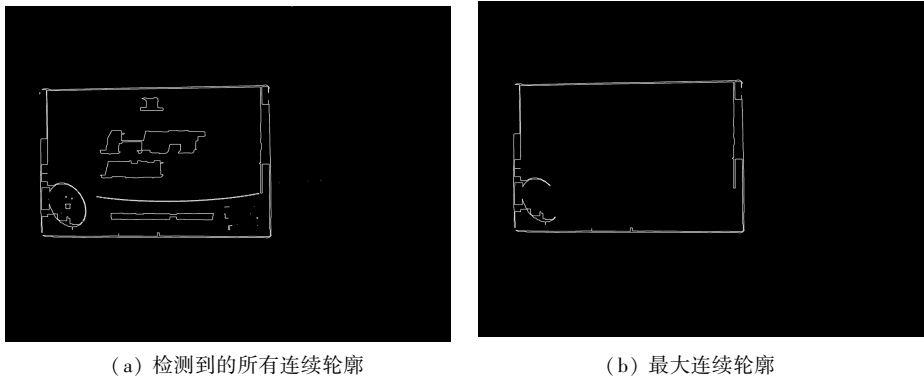


图 6 连续轮廓检测实例

Fig.6 Example of continuous contour detection

2.2.6 矩形轮廓拟合及 ROI 提取

连续区域通常代表着场景中的某个物体,在表示和定位图像区域的方法中,边界框可能是最简洁的。它的定义是:能完整包含该形状的最小垂直矩形。调用 `cv::rect`

函数得到最大连续轮廓的矩形边界框,结果如图 7(a)所示;根据矩形边界框从原始图像中提取出感兴趣的区域,结果如图 7(b)所示。

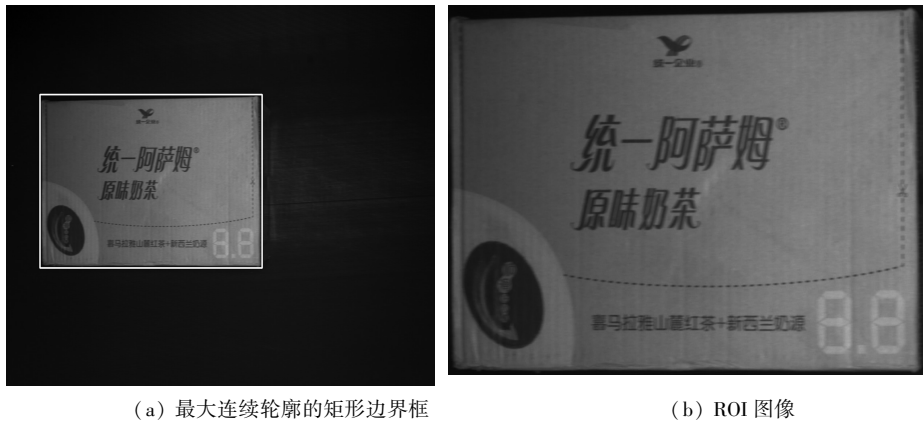


图 7 矩形边界框及 ROI 图像处理实例

Fig.7 Example of rectangular boundary box and ROI image processing

3 训练卷积神经网络模型

对传统的图像分类来说,特征提取不是一件简单的事情,要通过人工的方式设计有效的特征集合,需要很多的时间和精力。深度学习解决的核心问题之一就是自动地将简单的特征组合成更加复杂的特征,并使用这些组合特征解决问题。

如图 8 所示,深度学习可以自动从数据中学习更加复杂的特征表达方式,使得最后一步权重学习变得更加简单

且有效。训练卷积网络模型的实质,就是输入原始图像,经过一系列前向运算,返回每种类别的概率,将置信度最大的类别作为预测结果,以达到识别的目的。

采用模块式开发方式,将代码拆成 3 个程序:第 1 个是 `save_tfrecord.py`,它定义了从指定目录下读取图像数据,按比例生成训练集和验证集;第 2 个是 `inference.py`,它定义了前向传播的过程以及神经网络中的参数;第 3 个是 `train.py`,它定义了神经网络的训练过程。针对不同的应用场景,通常只需要修改 `inference.py` 文件即可。

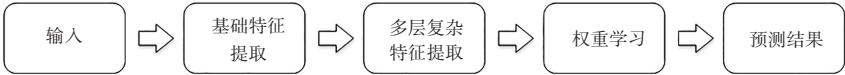


图 8 深度学习流程

Fig.8 Deep learning process

3.1 生成数据集

在训练卷积网络时,每一个数据来源被抽象成一个“数据集”,以数据集为基本对象,可以方便地进行 batching、随机打乱等操作。在图像相关任务中,输入数据通常以 `tfrecord` 形式存储。`tfrecord` 格式可以统一不同的原始数据格式,并更加有效地管理不同的属性,当训练数

据量较大时,还可以将数据分成多个 `tfrecord` 文件来提高处理效率。

一般来说,获取的图像大小是不固定的,但神经网络输入节点的个数是固定的,所以,在将图像的像素作为输入提供给神经网络之前,需要先将图像的大小统一。为了减少模型训练所需时间,提高模型识别效率,通常采用

227×227、100×100、32×32 像素等尺寸作为输入图像大小,通过 tf. image. resize_images 函数将获得的图像像素压缩为指定大小,然后保存至数据集文件中。

深度学习模型分类识别的准确度与训练样本数量有极大的关联性,在大多数情况下,在训练数据达到一定数量之前,越多的训练数据可以使分类结果更加准确。可通过相机自动采集大量图像数据,也可通过旋转、偏移、添加噪声等方式扩充样本数量^[10],如获取 ROI 图像后,每隔 10° 旋转 1 次图像,样本数量即可扩充 36 倍。

3.2 输入数据

TensorFlow 提供了一套多线程输入数据的框架,如图 9 所示。

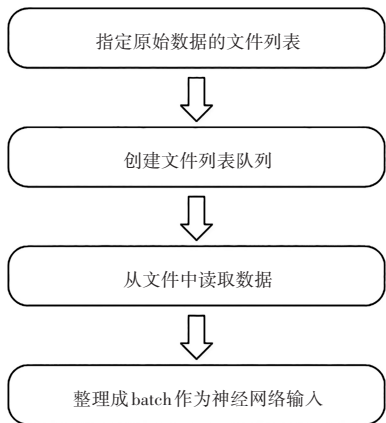


图 9 输入数据处理流程图

Fig. 9 Flow of input data processing

通过 tf. train. string_input_producer 函数来有效管理输入的 tfrecord 文件列表,读取 tfrecord 文件,通过 tf. train. batch 函数将多个输入样例整理成训练数据 batch,然后再提供给神经网络的输入层。

3.3 构建卷积神经网络

CNN 是一种深度学习的神经网络模型,通过前向传播计算输出值,反向传播调整连接权值。CNN 的基本架构通常包含卷积层、池化层和全连接层。对原始输入多次进行卷积操作和池化操作,得到越来越复杂的特征图,针对这个特征计算输入图片的卷积值,其中数值最大的区域与特征匹配,得到合适的特征,并将其输出到下一层结构中^[11]。采用 11 层卷积神经网络结构,包括 4 层卷积层、4 层池化层和 3 层全连接层,参数配置如表 1 所示。采用 ReLU 激活函数代替常规的 sigmoid 函数,可以加快网络训练和收敛,有助于提高网络的分类性能;采用最大池化而不是平均池化,可以避免平均池化层的模糊问题^[12]。

3.4 训练网络模型

在自动化物流输送线上采集 13 种类别的图像,这 13 种类别为冰红茶、卤香牛肉面、安慕希、早餐核桃牛奶、水蜜桃、百岁山、脉动(椰汁菠萝)、脉动(青柠)、营养快线(组合装)、藤椒牛肉面、阿萨姆、香菇鸡蛋面以及香辣牛肉面,每种类别的图像约 50 张。考虑到样本数量直接影响图像分类识别的精度,通过旋转、添加随机噪声等方法,将正常图片样本扩充到 18 720 张数据集,按 0.8 的比例划分训练集和验证集,其中,训练集 14 976 张,验证集 3 744 张。

表 1 卷积神经网络结构参数配置

Tab. 1 Configuration of convolutional neural network structure parameters

网络层数	内核数量	内核尺寸	步长	填充	输出尺寸
输入层	—	—	—	—	100×100×3
卷积层 conv1	32	5×5×3	1	3	102×102×32
最大池化层 max-pool1	—	2×2	2	0	51×51×32
卷积层 conv2	64	5× 5×32	1	3	53×53×64
最大池化层 max-pool2	—	2×2	2	0	26×26×64
卷积层 conv3	128	3×3×64	1	1	26×26×128
最大池化层 max-pool3	—	2×2	2	0	13×13×128
卷积层 conv4	128	3×3 ×128	1	1	13×13×128
最大池化层 max-pool4	—	2×2	2	0	6×6×128
全连接层 fc1	—	—	—	—	1 024×1
全连接层 fc2	—	—	—	—	512×1
全连接层 fc3	—	—	—	—	n_class×1

卷积神经网络的参数调试是一个难点,参数调试的好坏直接影响收敛的精度、最终模型的准确率以及网络训练所耗费的时间等。卷积神经网络的参数主要包括学习率 lr、批数据大小 batch 和模型训练次数 epochs 等^[13]。研究采用 Adam 优化器,设置 lr = 0. 1, batch = 16, epochs = 500,每隔 20 次打印一次验证集准确率和损失函数,结果如

图 10 所示。在 4G 独立显卡上采用 GPU 加速训练,训练时长仅为 1 min。

如图 10 所示,模型识别的准确率随着迭代次数的增加而增加,当迭代次数超过某个数时,模型识别的准确率不再发生变化;损失函数随着迭代次数的增加而减少,且逐渐收敛于 0,这表示训练集的预测结果和正确结果非常接近^[14-15]。

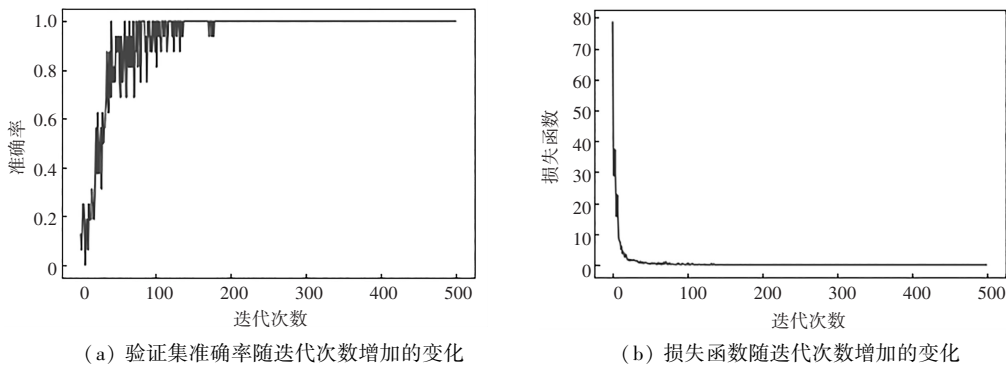


图 10 验证集准确率及损失函数随迭代次数增加的变化情况

Fig. 10 Changes in accuracy rate and loss function of the validation set as the number of iterations increases

3.5 保存离线模型

为了让训练结果可以反复调用,需要将训练得到的神经网络模型持久化。在进行离线预测时,一般输入图像数据后只关心输出结果,因此,只需要知道如何从神经网络的输入层经过前向传播计算得到输出层即可。TensorFlow 提供了 `convert_variables_to_constants` 函数,通过这个函数可以将计算图中的变量及其取值通过常量的方式保存,这样,整个 TensorFlow 计算图可以统一存放在一个文件中。当卷积神经网络结构不变时,保存的 pb 文件大小不变。轻量级 CNN 模型保存的离线 pb 模型文件,大小仅为 21.0 MB。

4 识别结果与显示

将训练好的 pb 文件拷贝到工控机上,工控机仅需要安装 OpenCV 库,而不需要安装 TensorFlow 库。相机获取

图像后,经过图像处理得到 ROI 图像,调用 `cv::dnn::readNetFromTensorflow` 函数载入 pb 模型文件,调用 `cv::dnn::blobFromImage` 函数将图像转换为卷积神经网络的输入对象,调用 `cv::dnn::Net` 对象的 `forward` 函数将图像数据输入模型文件,经过前向传播得到输出结果。

如图 11 所示,相机获取原始图像后,提取出 ROI 区域,在原始图像上画出白色矩形边界框并显示。将 ROI 图像输入 pb 模型后,输出结果为 $[3.324\ 958\ 3, -4.946\ 034, -11.962\ 296, 7.788\ 081\ 6, 2.596\ 995\ 8, -4.809\ 036\ 3, 2.651\ 99, 7.102\ 846\ 1, -7.670\ 851\ 7, -8.638\ 577\ 5, 17.985\ 228, -5.963\ 511\ 9, -4.167\ 291\ 6]$, 经过 Softmax 运算后,取出置信度最大的为识别类别,可知识别结果为 10,对应类别为阿萨姆,置信度为 0.999 943 02。将识别结果和概率显示在界面上,并通过服务器发送给连接的上位机,通过 PLC 控制使阿萨姆流向指定的分拣口。

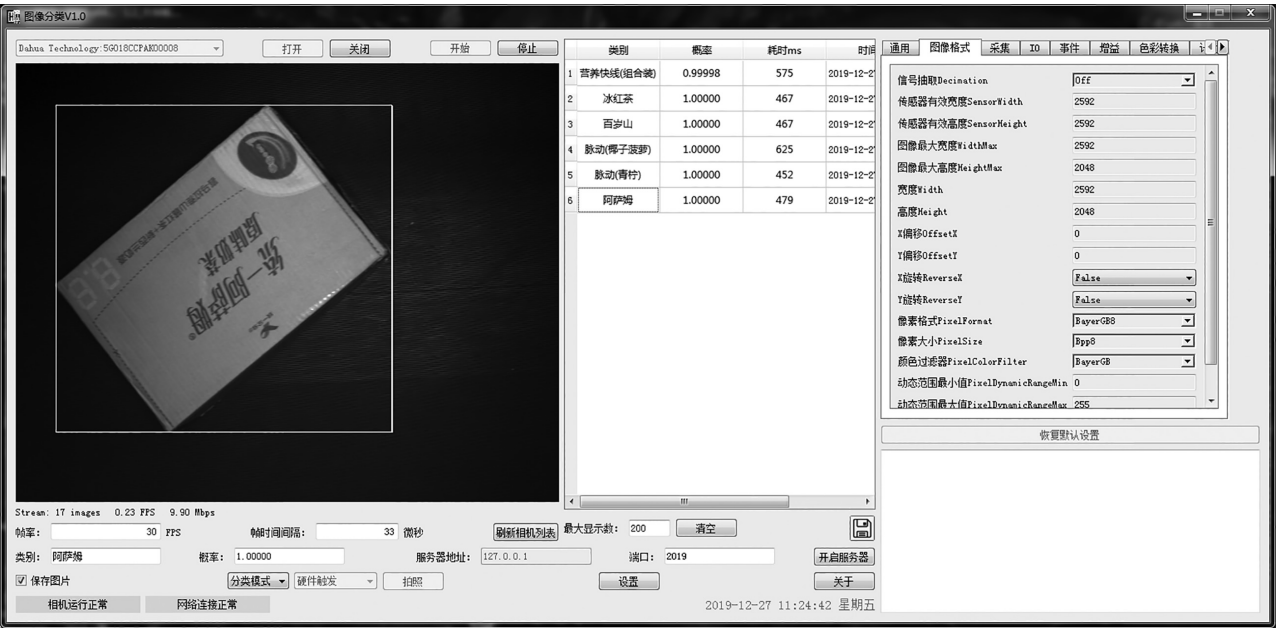


图 11 识别结果显示

Fig. 11 The identification results shows

在物流自动化输送线上对任意角度的 13 种类别的产品进行自动识别,识别次数为 300 次,均识别正确,且置信

度大于 0.99。

(下转第 183 页)