

doi: 10.3969/j.issn.1003-3106.2020.10.014

引用格式: 乔成, 周磊, 卢玉斌, 等. 基于 CNN 的车载终端自动化测试系统设计与实现 [J]. 无线电工程, 2020, 50(10): 880-886.
[QIAO Cheng, ZHOU Lei, LU Yubin, et al. Design and Implementation of Automated Test of Vehicular Terminal Based on Convolutional Neural Network [J]. Radio Engineering, 2020, 50(10): 880-886.]

基于 CNN 的车载终端自动化测试系统设计与实现

乔成, 周磊, 卢玉斌, 叶军

(扬州航盛科技有限公司, 江苏 扬州 225000)

摘要: 随着人工智能技术飞速发展, 机器学习对于提升自动化测试效率, 改善电子产品质量起着至关重要的作用。针对传统车载终端(车机)测试方法效率低下, 人工成本高的问题, 提出了一种利用 ResNet50_Sub_Pixel_DSNT(RSPD) + 透视变换矫正+YOLOv4 的测试方法。该方法利用 RSPD 实现对车机屏幕 4 个关键角点的定位, 并通过透视变换算法对定位目标区域进行截取矫正, 采用 YOLOv4 模型对矫正后的图片进行相关功能控件位置识别, 驱动机器臂点击识别区域目标。测试结果表明, RSPD + 透视变换矫正+YOLOv4 的测试方法, 能够准确控制机器臂完成车机屏幕功能控件的点击测试操作, 其准确度达到 97.3%。

关键词: 机器学习; 关键点检测; 目标识别; 卷积神经网络

中图分类号: TP181

文献标志码: A

开放科学(资源服务)标识码(OSID):



文章编号: 1003-3106(2020)10-0880-07

Design and Implementation of Automated Test of Vehicular Terminal Based on Convolutional Neural Network

QIAO Cheng, ZHOU Lei, LU Yubin, YE Jun

(Yangzhou Hangsheng Technology Co., Ltd., Yangzhou 225000, China)

Abstract: With the rapid development of AI technology, machine-learning plays a vital role in improving the efficiency of automated test and the quality of electronic products. A test method using ResNet50_Sub_Pixel_DSNT(RSPD) + perspective conversion correction + YOLOv4 is proposed to solve the problems of low efficiency and high labor cost in traditional test method of vehicular terminal. RSPD is used to locate the four key angular points of screen, then the target area is captured and corrected based on the perspective conversion algorithm. Finally, YOLOv4 model is used to locate the associated functional controls based on the corrected picture, driving the robot-arm to click on the target controls. The results show that the accuracy is up to 97.3%. The method based on RSPD + perspective conversion correction + YOLOv4 can accurately instruct the robot-arm to perform the test operation by clicking on the functional controls on the vehicle screen.

Key words: machine learning; key point detection; target recognition; convolutional neural network

0 引言

由于汽车智能化程度越来越高, 车载终端设备的设计和制造变得越来越复杂, 产品质量把控难度增加。近些年深度学习技术的飞速发展, 特别是图像技术的广泛应用, 为自动化测试技术指明了方向, 成为众多学者、工程师研究的重点。其中, 在精密仪表设备制造领域, 武华敏^[1]等人将 CNN 神经网络应用到光学设备透镜的生产中, 实现对透镜十字像整体特征的识别与定位, 提升了透镜在图像质量较差状态下得到准确的十字线中心点能力; 在医学影像领域, BoNet^[2]等人利用 CNN 模型构建了一个端到端

的自动骨龄评估网络, 该网络首先提取手掌骨 X 光片的关键骨头区域, 随后预测各个骨头区域成熟度, 最后利用线性回归的方法预测骨头的年龄, 其预测年龄的误差均值在 0.8 岁左右, 表现明显优于传统的 Tanner-Whitehouse^[3]骨龄测试方法; Ali Narin^[4]等人利用 3 种不同的 CNN 模型(ResNet 50、Inceptionv3e、Inception-ResNetV2), 用于检测新冠病毒肺炎感染者的胸部 X 射线, 结果显示使用 ResNet 50 模型构建的分类器模型能够达到 98% 的准确度, 该方法能够极大地提升新冠病毒检测的精度和效率; 在自动化测试领域, H. Baumgartl^[5]利用 CNN 识别铝

收稿日期: 2020-04-28

铸件 X 射线图像中的缺陷零件,该方法结合“转移学习”的方法,能够达到 90% 以上的检测精度; Miao He^[6] 等人利用(LAMSTAR) 神经网络来诊断轴承故障,该方法将历史大数据中提取特征对电力驱动器轴承的健康状态进行判定,实验结果表明, LAMSTAR 神经网络能够在不同状态下对轴承状态故障进行准确分类。

本文首次将 CNN 深度学习图像技术应用到车载终端的自动化测试当中,并且提出了一种 RSPD + 透视变换矫正+YOLOv4^[7] 的测试方法,其中 RSPD 关键点检测模型相比于 BoNet^[2] 端到端结构的 CNN 模型,能有效解决全连接层结构中空间泛化能力差,特征图信息丢失的问题。同时采用最新的 YOLOv4 目标识别模型对车机屏幕功能控件进行识别,与其他 SOTA^[8-10] 以及 YOLOv3^[11-12] 模型相比较, YOLOv4 能更有效平衡模型的预测精度与运行效率。结果表明,该方法能够有效应用于车机自动化测试过程中。

1 自动化测试系统设计

基于深度神经网络的自动化测试技术涉及关键

点检测、图片矫正、目标识别和机器臂驱动 4 部分。将摄像头实时采集的车机界面图片作为输入参数传递给测试模块,机器臂即可自动完成相关测试操作,其整个实现框架如图 1 所示。关键点检测模块利用 ResNet 结构的 CNN 提取图片特征信息,通过 Sub-Pixel^[13] 进行上采样,将深度信息转换为平面信息,利用 DSNT 进行 Heatmap 热力图和坐标点位置的转换,求取 k 个关键点位置 $p_k = (i_k, j_k)$, $k = 1, 2, 3, 4$ 分别代表 4 个关键点 x, y 轴坐标。关键点坐标输入后续图片矫正模块,结合透视变换法,进行车机界面图片截取以及矫正变换,为目标识别提供标准图片输入。目标识别模块通过 YOLOv4 识别图片内功能控件,输出控件标签名称(按钮、对话框、滑块、...)以及控件对应的矩形区域坐标 $[x_{\min}, x_{\max}, y_{\min}, y_{\max}]$ 。最后,在机器臂驱动模块中计算目标控件矩形区域的中心点坐标 $C_l = (x_l, y_l)$,其中 l 是标签类别,将获取的 C_l 坐标进行坐标空间转换映射到机器臂的坐标系中,驱动机器臂运作。

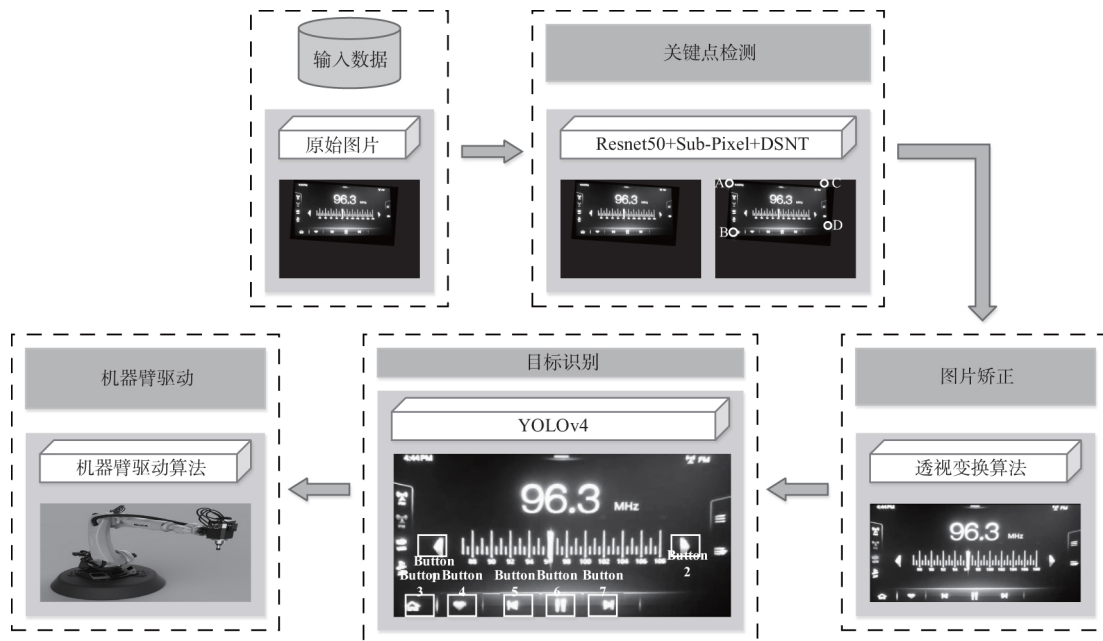


图 1 车载终端自动化测试框架

Fig.1 Automated testing framework of vehicular terminal

2 自动化测试系统模型 & 算法设计

2.1 RSPD 关键点检测模型

本文利用 ResNet-50^[12] 残差网络作为关键点检测基础网络模型结构,并对最后的 2 个全连接层进

行删除操作,在末尾加入了 5 个 Sub-Pixel^[13] 层进行特征图的上采样操作,对特征图的深度信息融合,最后加入微分空间数值转换层(DSNT)^[14] 进行关键坐标点回归操作,详细网络结构如图 2 所示。

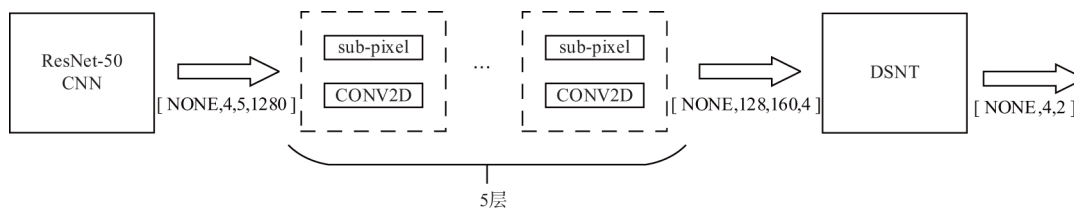


图2 RSPD 关键点检测模型网络结构

Fig2 Network structure of RSPD key point detection model

图2中 Sub-Pixel 层的实现方式为:

$$\begin{cases} I^l = PS(W_l * I^{l-1}) + b_l \\ PS(I^{l-1})_{x,y,c} = I^{l-1}_{x/r, y/r, c * r \bmod(y/r) + c * \text{mod}(x/r)} \end{cases} \quad (1)$$

式中, W_l, b_l 为 l -th 卷积层的训练权重和偏差值; W_l 为卷积过滤器, 由一个四维的张量 $n_{l-1} \times k_l \times k_l \times n_l$ 构成; n_{l-1}, n_l 分别为该卷积过滤器输入输出的通道数; k_l 为该卷积过滤器的尺寸; PS 为一种周期筛选方法, 该方法能够将特征图尺寸从 $H \times W \times r^2$ 转换为 $rH \times rW \times C$, 在 ResNet-50^[12] 输出端接入 5 个 Sub-Pixel 层能够将低分辨率的特征图转换为高分辨率, 为后面的关键点回归提供更多特征信息。在每个 Sub-Pixel 后都会执行一个二维卷积操作, 该方法主要控制特征图输出的数量, 避免其上采样速率过快, 丢失特征信息。由于本文预测的关键点为车机 4 个角坐标, 故设计最后一层卷积层结构为 $n_l \times k_l \times k_l \times 4$, 输出格式为 $H = [128, 160, 4]$ 四通道高分辨率热力图, 每个通道的热力图分别对应一个角点坐标。

目前, 主流的坐标回归方法主要采用全连接层预测法以及热力图极大值法, 但是前者对空间泛化能力造成极大的损害, 后者又由于计算热力图极大值的操作而无法进行整体微分操作, 故本文采用了 DSNT^[14] 的方法, 有效克服上述缺点, 该方法采用最小损失训练法 (MRT)^[15] 构建热力图与坐标点之间的可微分损失函数, 其简要公式为:

$$u = DSNT(H) \quad (2)$$

式中, u 为热力图 H 经过 DSNT 转换得到的坐标信息, 真实结果为 4 个角点坐标。

2.2 透视变换算法

由于在实际测试过程中, 无法保证摄像头准确定位车机设备在视频中心位置, 导致车机在图片中存在不同角度偏移情况的出现, 故本文采用透视变换^[16]方法, 对定位的 4 个角点利用透视变换, 将截取的目标区域统一转化为 240 pixel \times 320 pixel 标准格式的矩形图像。

透视变换是指利用透视中心、像点和目标点三

点共线的条件, 按透视旋律定律使透视面绕透视轴旋转某一角度, 将一个平面通过一个投影矩阵投影到指定平面上。透视变换原理如图 3 所示。

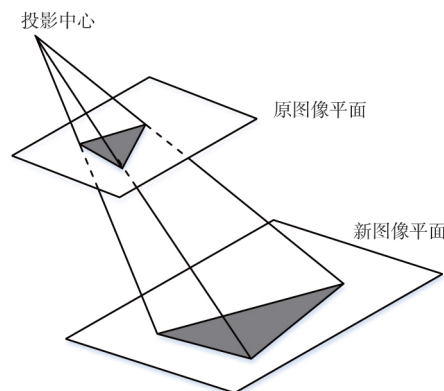


图3 透视变换原理

Fig3 Schematic diagram of perspective transformation

$$\begin{bmatrix} x' & y' & w' \end{bmatrix} = \begin{bmatrix} u & v & w \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad (3)$$

式中, u, v 是原始图片; w 等于 1。通过透视变化得到的图片坐标 x, y , 其计算方式为:

$$\begin{cases} x = \frac{x'}{w'} = \frac{a_{11} \times u + a_{21} \times v + a_{31} \times 1}{a_{13} \times u + a_{23} \times v + 1 \times 1} \\ y = \frac{y'}{w'} = \frac{a_{12} \times u + a_{22} \times v + a_{32} \times 1}{a_{13} \times u + a_{23} \times v + 1 \times 1} \end{cases} \quad (4)$$

上述变换矩阵可以分为 4 个部分, 其中

$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ 主要用于图像的缩放、转换操作,

$\begin{bmatrix} a_{31} & a_{32} \end{bmatrix}$ 部分进行平移操作, $\begin{bmatrix} a_{13} & a_{23} \end{bmatrix}$ 部分产生透视变换, 第四部分参数 a_{33} 一般默认为 1。由于透视变换矩阵有 8 个参数, 故 4 个角点坐标可以完成矩阵参数的求解需求。

2.3 YOLOv4 目标识别模型

在目标识别技术领域, YOLO 系列模型一直以识别精度高、效率高而被应用在众多的图像识别工程中, 本文采用的 YOLOv4^[7] 模型在前者

YOLOv3^[11]基础进行了众多改进与优化,使得其在相同执行效率下能够取得更好的识别结果。模型网

络结构如图4所示。

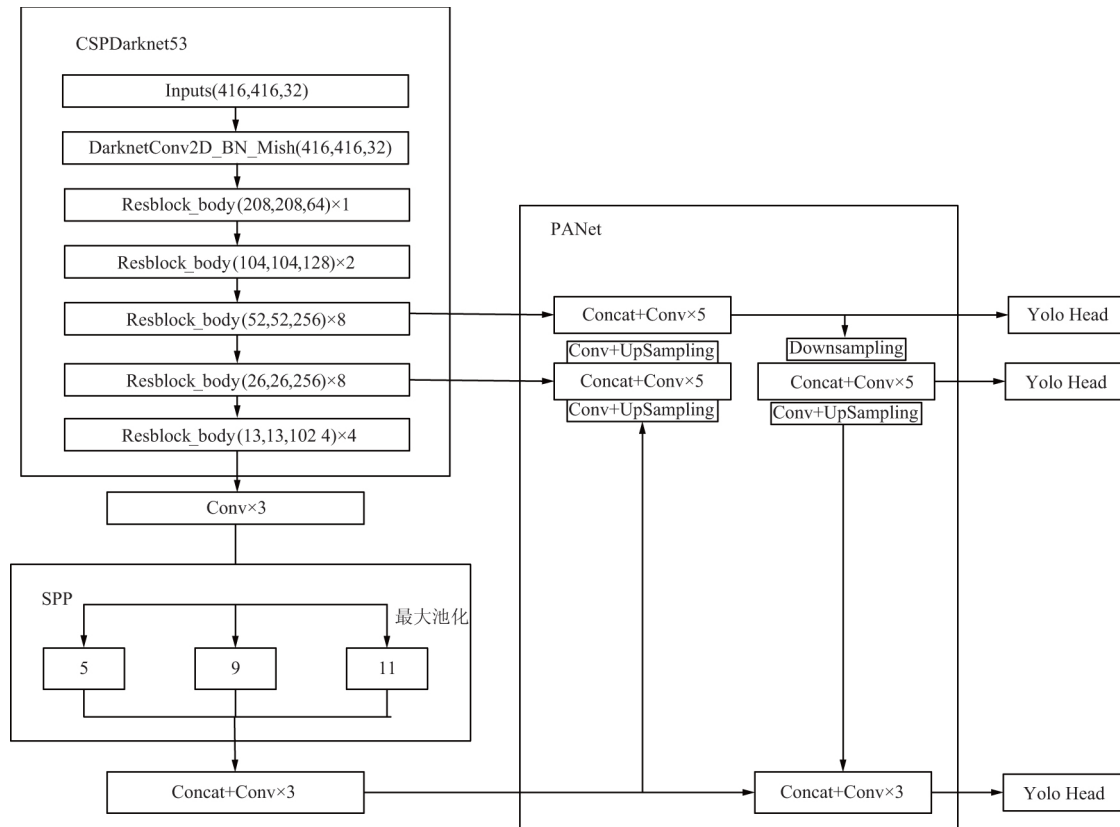


图4 YOLOv4 模型网络结构

Fig4 Model network structure of YOLOv4

由图4可以看出, YOLOv4 主要由三部分构成: CSPDarknet53(backbone) +(SPP+PANet) +YOLOv3 Head。其中,主干特征提取网络 CSPDarkNet53 相比于 YOLOv3 的 DarkNet53 网络,将卷积层的激活函数由 LeakyReLU 换成了 Mish 并且使用了 CSPnet 结构构建 resblock_body 模块,其网络结构如图5所示。

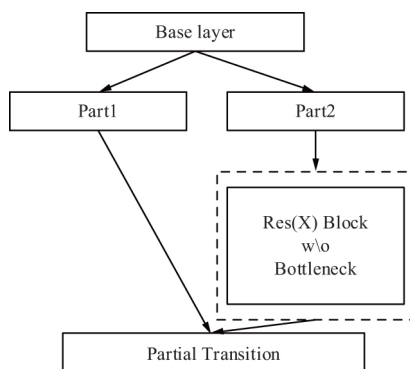


图5 SPResNet(x) 网络结构

Fig5 Network structure of SPResNet(x)

SPP^[17]网络连接在 CSPDarknet53 最后一个特征层的卷积里,在对 CSPDarknet53 的最后一个特征层进行3次卷积后,分别利用4个不同尺寸的最大池化进行处理,最大池化的核大小分别为13×13、9×9、5×5、1×1(即无处理)。SPP^[17]相比单级池化具有更强鲁棒性,并且可以在不同尺寸上提取特征。PANet^[18]网络主要对图4三个特征层的结果进行实例分割,该方法相比传统的特征金字塔的方法,增加从下到上的提取结构,能够显著提升模型在物体检测和实例分割网络上的性能。最后的网络输出结构采用 YOLOv3 Head 的结构,完成对 loss 损失函数的构建。

24 机器臂驱动算法

机器臂驱动模块主要利用电机驱动机器人手臂,触摸车机设备功能控件区域,完成相关功能的触发,其主要通过3个步骤完成相关的操作。第一步机器臂坐标区域初始化,手动控制机器臂点击车机

屏幕 4 个角点,绘制机器臂内部坐标系。第二步计算 YOLOv4 识别目标区域的中心点,其实现公式为:

$$\begin{cases} x' = x_{\min} + \frac{x_{\max} - x_{\min}}{2} \\ y' = y_{\min} + \frac{y_{\max} - y_{\min}}{2} \end{cases}, \quad (5)$$

式中, (x_{\min}, y_{\min}) , (x_{\max}, y_{\max}) 为识别目标区域左上和右下角点坐标。在计算获得目标识别区域的中心点坐标 (x', y') 后,利用其在屏幕中宽高比,计算得到其在机器臂坐标中的点击位置 (x, y) ,完成第三步的工作,其公式如下:

$$\begin{cases} \frac{x'}{\text{width}'} = \frac{x}{\text{width}} \\ \frac{y'}{\text{height}'} = \frac{y}{\text{height}} \end{cases}. \quad (6)$$

在求取机器臂触控目标坐标后,发送点击信号触发点击事件的完成。机器臂触近点映射示意图 6 所示。

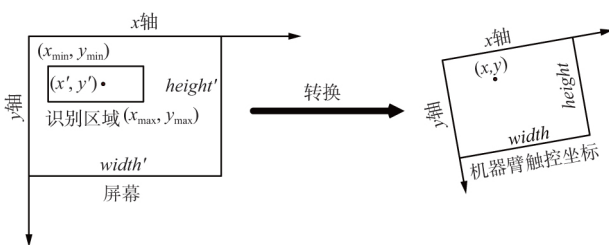


图 6 机器臂触控点映射示意

Fig6 Touch point mapping diagram of robot-arm

3 实验分析

3.1 测试数据选择

本文测试数据来源于扬州航盛测试验证科,图片分辨率为 480 pixel×640 pixel,其中 5 000 张图片用于模型的训练,1 000 张图片用于测试结果分析。采用 cvat 工具对图片进行标注,其中关键点标注采用画点法,控件标注则采用矩形轨迹跟踪法,以标准 tfrecord 格式存储标注数据。

3.2 测试结果分析

3.2.1 RSPD 检测结果分析

本文关键点检测模型主要基于谷歌 tensorflow2.0 框架构建而成,并采用 Adam^[19] 优化算法对模型的参数进行求导优化,其中超参数学习率设置为 e^{-3} ,批量数据为 30,迭代次数为 50,衰减因子为 e^{-6} ,损

失函数采用的是欧几里德距离公式。本实验的硬件设备由 CPU (E5-2630 @ 2.4 GHz),显卡 (Nvidia GTX1080 8GB)。训练结果如图 7 所示,模型测试结果均方误差如表 1 所示。由表 1 可以看到,采用 ResNet+Sub-Pixel+DSNT 网络结构训练的损失函数值明显小于其他传统的方法,其欧几里德损失函数值远小于 0.02,在测试数据阶段,也获得了平均误差 0.027 8 的优异表现,故可以说明该方法能够有效预测车机 4 个关键点角坐标的位置。

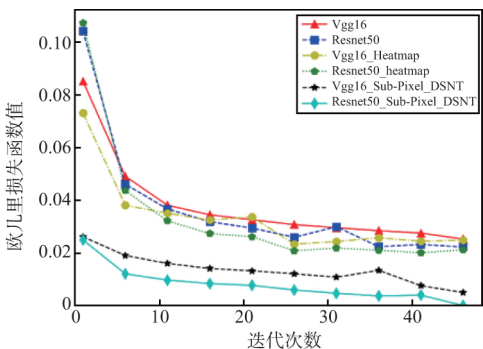


图 7 模型训练损失函数

Fig7 Loss function of model training

表 1 模型测试结果均方误差

Tab1 Root mean square error of model test results

| 模型 | RMSE |
|-------------------------|---------|
| Vgg16 | 0.130 7 |
| Resnet50 | 0.119 2 |
| Vgg16_Heatmap | 0.126 7 |
| ResNet50_Heatmap | 0.094 2 |
| Vgg16_Sub-Pixel_DSNT | 0.030 2 |
| ResNet50_Sub-Pixel_DSNT | 0.027 8 |

3.2.2 YOLOv4 目标识别结果分析

本文目标识别数据是通过 CVAT 进行车机界面区域截取并利用透视变换矫正后获取,其默认标准分辨率为 240 pixel×360 pixel。其损失函数由 IOU^[20] 损失、置信度损失、分类损失三部分组成,超参数设置以及硬件配置与上文关键点检测模块相同,最终对定义的 6 个标签目标对象(按钮、对话框、群组、滑动条、复选框、输入框)测试结果如表 2 和表 3 所示。由测试结果可看出,在检测精度方面,YOLOv4 对单个目标 AP 预测结果基本都在 0.95 以上,且单帧运行之间也只有 86 ms,无论是识别精度还是效率都表现出了极好的优异性。

表 2 目标识别模型测试精度结果

Tab2 Test accuracy of different target recognition models

| 检测模型 | 按钮(AP) | 对话框(AP) | 群组(AP) | 滑动条(AP) | 复选框(AP) | 输入框(AP) | MAP |
|-------------|---------|----------|---------|----------|----------|----------|------|
| YOLOv4 | 0982 | 0972 | 0976 | 0945 | 0956 | 0969 | 0966 |
| YOLOv3 | 0933 | 0932 | 0937 | 0904 | 0910 | 0929 | 0924 |
| ATSS | 0978 | 0977 | 0979 | 0948 | 0965 | 0959 | 0967 |
| ASFF* | 0968 | 0966 | 0955 | 0953 | 0949 | 0979 | 0961 |
| CenterMask* | 0974 | 0964 | 0953 | 0965 | 0953 | 0975 | 0964 |

表 3 目标识别模型测试性能结果

Tab3 Test performance of different target recognition models

单位: ms/frame

| 检测模型 | 时间 |
|-------------|-----|
| YOLOv4 | 86 |
| YOLOv3 | 82 |
| ATSS | 153 |
| ASFF* | 117 |
| CenterMask* | 143 |

3.2.3 机器臂点击结果分析

通过前文关键点检测、图像矫正和目标检测一系列操作后,将车机屏幕识别的目标区域中心点坐标转换至机器臂定位的内部坐标系,触发机器臂“点击”动作,判断该次动作是否触发了其点击目标的相关功能。若成功触发目标控件功能,则认为该次操作正确,否则被判定为错误操作。最终测试结果如表 4 所示。

表 4 机器臂点触结果分析(1 000 次)

Tab4 Analysis of robot-arm touch results(1 000 times)

| 模型组合 | 正确点触 | 错误点触 | 准确率/% |
|---------------------------------------|------|------|-------|
| Vgg16+透视变换矫正+ YOLOv4 | 895 | 105 | 895 |
| Resnet50+透视变换矫正+YOLOv4 | 923 | 77 | 923 |
| Vgg16_Heatmap+透视变换矫正+YOLOv4 | 918 | 82 | 918 |
| Resnet50_Heatmap+透视变换矫正+ YOLOv4 | 932 | 68 | 932 |
| Vgg16_Sub-Pixel_DSNT+透视变换矫正+YOLOv4 | 965 | 35 | 965 |
| Resnet50_Sub-Pixel_DSNT+透视变换矫正+YOLOv4 | 973 | 27 | 973 |

由上表所示, RSPD + 透视变换矫正+ YOLOv4 的模型组合在 1 000 次机器臂点击测试中获得的准确率最高,达到了 97.3%,明显优于其他组合,证明其在实际生产操作中具有极高的可靠性。

4 结束语

本文基于 CNN 技术,提出了一种 RSPD + 透视变换矫正+YOLOv4 的车载终端自动化测试方法,该方法利用深度学习模型检测识别车机屏幕相关功能控件位置坐标,控制机器臂完成对控件区域的点击测试操作,其测试结果取得了 97.3% 准确率。虽然在实际应用中,该技术取得了不错的效果,但是仍然存在不足,主要问题在于当前测试环境处于理想状态下,如何提升模型的鲁棒性,使其在复杂环境下仍

能保持较高的准确率,成为接下来的研究重点。



参考文献

[1] 武华敏,杨漠雨,黄晓雪,等.基于 CNN 的十字像中心检测[J].光学仪器,2019,41(5):38-46.

[2] APAMPINATO C, PALAZZO S, GIORDANO D, et al. Deep Learning for Automated Skeletal Bone Age Assessment in X-ray Images[J]. Medical Image Analysis, 2017, 36: 14-51.

[3] MALINA R M, BEUNEN G P. Assessment of Skeletal Maturity And prediction of Adult Height (TW3 method) [J]. American Journal of Human Biology, 2002, 14 (6): 788-789.

[4] NARIN A, KAYA C, PAMUK Z. Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images

- and Deep Convolutional Neural Networks [J]. arXiv preprint arXiv: 2003.10849 2020.
- [5] BAUMGARTL H ,TOMAS J ,BUETTNER R ,et al.A Novel Deep-learning Approach for Automated Non-destructive Testing in Quality Assurance Based on Convolutional Neural Networks [C]// 13th International Conference on Advanced Computational Engineering and Experimenting. Athens: ACEX 2019: 1319–1320.
- [6] HE M ,HE D.Deep Learning Based Approach for Bearing Fault Diagnosis [J].IEEE Transactions on Industry Applications 2017 ,53(3) : 3057–3065.
- [7] BOCHKOVSKIY A ,WANG C Y ,LIAO H Y M.YOLOv4: Optimal Speed and Accuracy of Object Detection [J]. arXiv preprint arXiv: 2004.10934 2020.
- [8] ZHANG S F ,CHI C ,YAO Y Q ,et al.Bridging the Gap between Anchor-based and Anchor-free Detection via Adaptive Training Sample Selection [J]. arXiv preprint arXiv: 1912.02424 2020.
- [9] LIU S T ,HUANG D ,WANG Y H.Learning Spatial Fusion for Single-shot Object Detection [J].arXiv preprint arXiv: 1911.09516 2019.
- [10] LEE Y W ,PARK J.CenterMask: Real-time Anchor-free Instance Segmentation [J]. arXiv preprint arXiv: 1911.06667 2020.
- [11] REDMON J ,FARHADI A.Yolov3: An Incremental Improvement [J].arXiv preprint arXiv: 1804.02767 2018.
- [12] HE K M ,ZAHNG X Y ,REN S Q ,et al.Deep Residual Learning for Image Recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition.Las Vegas: IEEE 2016: 770–778.
- [13] SHI W Z ,CABALLERO J ,HUSZAR F ,et al.Real-Time Single Image and Video Super-resolution Using an Efficient Sub-pixel Convolutional Neural Network [C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition.Las Vegas: IEEE 2016: 1874–1883.
- [14] NIBALI A , HE Z , MORGAN S , et al. Numerical Coordinate Regression with Convolutional Neural Networks [J].arXiv preprint arXiv: 1801.07372 2018.
- [15] SHEN S ,CHENG Y ,HE Z ,et al.Minimum Risk Rraining for Neural Machine Rranslation [J].arXiv preprint arXiv: 1512.02433 2015.
- [16] 何援军.透视和透视投影变换——论图形变换和投影的若干问题之三 [J].计算机辅助设计与图形学学报 , 2005(4) : 734–739.
- [17] HE K ,ZHANG X ,REN S ,et al.Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition [J].IEEE Transactions on Pattern Analysis and Machine Intelligence 2015 ,37(9) : 1904–1916.
- [18] KINGMA D P ,BA J.Adam: A Method for Stochastic Optimization [J].arXiv preprint arXiv: 1412.6980 2014.
- [19] YU J H ,JIANG Y N ,WANG Z Y ,et al.UnitBox: An Advanced Object Detection Network [J]. arXiv preprint arXiv: 1608.01471 2016.

作者简介



乔成男 (1991—), 毕业于华东理工大学控制科学与工程专业, 硕士, 助理工程师。主要研究方向: 边缘计算、图像处理与模式识别。

周磊男 (1980—), 博士, 副教授。主要研究方向: 片上网络。

卢玉斌男 (1984—), 高级工程师。主要研究方向: Linux 系统架构。

叶军男 (1988—), 硕士, 工程师。主要研究方向: 图像处理与模式识别。