

基于 Mask R-CNN 的试管-支架系统 Data Matrix 码识别方法

刘石坚¹, 林锦嘉¹, 陈梓灿¹, 邹峥²

(1. 福建省大数据挖掘与应用技术重点实验室, 福建 福州 350118;

2. 福建师范大学 计算机与网络空间安全学院, 福建 福州 350117)

摘要: 在试管-支架自动化系统的输入图像中, Data Matrix (DM) 码呈现为多个小目标, 图像存在成像模糊、边缘干扰严重等问题, 使得传统方法难以达到良好的识别效果。为此, 提出一种基于深度学习的 Data Matrix 码识别方法 DeepDMCode, 以 Mask R-CNN 模型为基础, 通过内容差异化数据合成和同步自动化标注, 实现训练数据的增强, 提升模型的学习能力。在模型分割结果的基础上, 提出一种旋转校正方法, 确保可用标准解码库实现 DM 码的解码。以分辨率为 1 600×1 200、支架容量为 96 的数据实验表明, 由于该方法在前期码定位阶段最大程度地还原码边界信息, 准确度可达 0.92 (mIoU), 完成单张图像中所有 DM 识别的平均速度为 5.2 s, 优于 YOLO、SegNet、CenterNet 等主流工业基准算法。

关键词: 试管-支架系统; Mask R-CNN; Data Matrix 码; 人工数据合成; 实验室自动化

中图分类号: TP301

文献标志码: A

文章编号: 1672-4348(2023)04-0378-07

A Mask R-CNN based method for Data Matrix code recognition from tube-rack system

LIU Shijian¹, LIN Jinjia¹, CHEN Zican¹, ZOU Zheng²

(1. Fujian Provincial Key Laboratory of Big Data Mining and Applications, Fuzhou 350118, China;

2. College of Computer and Cyber Security, Fujian Normal University, Fuzhou 350117, China)

Abstract: The Data Matrix (DM) codes from the tube-rack automatic system can be seen as multiple small targets in the input image, which has problems such as blurred imaging and serious edge interference and makes it difficult for traditional methods to achieve good recognition results. Therefore, a Data Matrix code recognition method based on deep learning named DeepDMCode was proposed. Based on the Mask R-CNN model, the training data was enhanced and the learning ability of the model was improved through content-differentiated data synthesis and synchronous automatic annotation. Based on the segmentation results of the model, a rotation correction method was proposed to ensure that the DM code can be decoded using a standard decoding library. Experiments were carried out with images of resolution 1 600×1 200, which were captured from racks of 96 capacity. Results show that the method can restore the code boundary information to the greatest extent in the early code positioning stage, and the accuracy can reach 0.92 (mIoU). The average speed of completing all DM recognition in a single image is 5.2 s, which is better than mainstream industrial benchmark algorithms such as YOLO, SegNet, and CenterNet.

Keywords: tube-rack system; Mask R-CNN; Data Matrix code; data synthesis; laboratory automation

生物实验室常需要采集大量样本, 并保存到试管中、放置于支架槽内。为方便快速管理和分析, 装有样本的试管底部往往会印刷一种名为

Data Matrix (DM)^[1] 码的矩阵式二维码, 作为样本唯一标识。相较其他条码, DM 码具有尺寸小、存储信息量大、纠错能力和抗干扰能力强等特点, 被

收稿日期: 2023-05-09

基金项目: 福建省教育厅科研项目 (JAT210283); 福建工程院校级科研项目 (GY-Z21047, GY-Z20068)

第一作者简介: 刘石坚 (1983—), 男, 湖南长沙人, 副教授, 博士, 研究方向: 深度学习, 机器视觉。

广泛应用于集成电路、机械磨具、精密生物器械等小件物品上。深度学习技术在过去几年发展迅速,已被广泛用于各领域,将其应用于 DM 码识别和信息读取,有利于实验室自动化^[2]。

1 相关工作

DM 码的传统识别方法主要利用其所具有 L 形状几何边界的特点来实现。例如,王等人^[3]采用 Hough 变换^[4]来实现 DM 码的检测定位。然而在实际应用中,输入数据中往往存在多个条码,且条码周围通常存在较多干扰。虽然许多方法^[5~6]试图解决上述问题,但在面对复杂情况时,方法的稳定性仍然难以保证。相较而言,基于深度学习的方法因不依赖人工特征,具有更好的准确性和泛化能力,成为目前的研究趋势。

1.1 目标识别策略

基于深度学习的方法在实现条码定位时可分为目标识别策略和实例分割策略。

目标识别策略的思路是先通过 YOLO^[7]、Faster R-CNN^[8]等目标识别方法找到输入图片中包含条码的局部矩形区域(即粗定位),再结合上述传统方法完成条码的最终定位。例如,易等人^[9]提出一种级联多尺度特征融合网络快速检测算法实现对识别物流包裹的条形码。谢等人^[10]基于 Faster R-CNN 框架提出一种香烟条码识别方法。针对复杂背景下的多条码检测问题,文献[11][12]提出不同的基于改进 YOLO 模型的条码检测方法。

1.2 实例分割策略

目标识别策略的局限性在于需要两个阶段来实现条码的定位,其深度学习的优势仅体现在第一阶段。实例分割策略则直接基于条码的实例分割结果进行定位,更简洁高效。例如,Zharkov 等人^[13]使用基于实例分割的深度学习方法识别一维条形码。胡等人^[14]提出基于 SegNet 的 DM 码定位方法。分割策略的关键问题在于构建准确的分割模型,来获取可靠的 DM 码边界。

本课题提出的 DeepDMCode 方法与已有深度学习方法的主要差别在于处理对象和应用领域的不同。已有相关研究多处理一维条形码和 QR 码识别,面向多 DM 码的文献较少,讨论试管-支架系统中 DM 码识别的文献更加寥寥无几。其次,本课题虽采用基于 Mask R-CNN^[15]模型的实例分

割路线,但针对提升准确性所提出的数据合成方法、标注方法及旋转矫正方法是普适性的策略,与特定分割模型的选取无关。

2 研究内容及挑战

课题以包含 96 个试管槽的一套固定支架为环境,在支架底部放置相机获取拍摄图像,以自动识别如图 1 所示图像中所有试管 DM 码为研究目标。

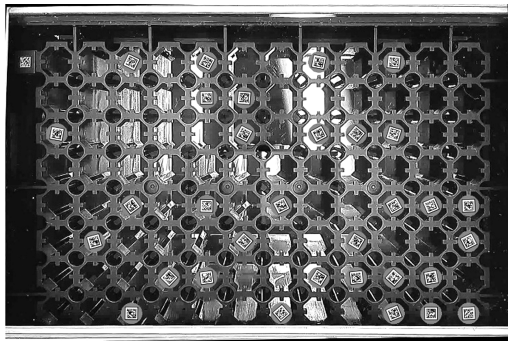


图 1 试管支架底部成像图片

Fig.1 Image captured at the bottom of a tube rack

DM 码识别具体可分为码定位(第 3 章)和矫正解码(第 4 章)两部分,即首先找到 DM 码所在位置,然后将其旋转至标准朝向进行解码。其中码定位是识别的前提基础,旋转矫正使得通过标准解码库实现解码成为可能。

本课题面临以下挑战:

(1)DM 码为小目标。单个试管 DM 码约占整副图像面积的 1/1 200,如图 1 所示,定位和识别时容易对误差敏感。

(2)位置和旋转角度分布不均。如图 1 所示,单个 DM 码的位置和方向是随机分布的,这种差异和变化增加码识别难度。

(3)目标边界模糊。相对如图 2(a)所示的标准 DM 码,在实际应用中,当照片受到光照不均或噪声影响时,DM 码边界常常呈现模糊甚至消失的情况,如图 2(b)所示。

(4)存在相似性干扰。如图 2(c)所示,在试管槽孔连接处以及镂空槽孔处存在与 DM 码边界相似的特征,这类相似性干扰会降低识别区分能力。

(5)数据量不足。深度学习方法依赖大规模数据进行模型训练,但试管支架图片数据来源有限,且没有公开的数据集作为补充。

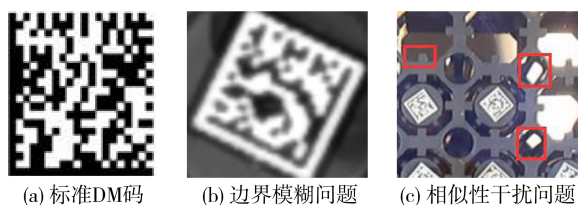


图 2 DM 码展示

Fig.2 Demonstration of data matrix codes

本课题以支架试管图片作为输入,将单张图片中多个 DM 码作为检测对象,在无法获取大量学习样本的前提下,以减少繁琐的小目标手工标注和提高识别精度为目标,实现多目标码的识别。

3 码字区域精定位

在 DM 码字区域定位环节,首先针对训练样本数量有限的问题,采用多层随机方式来生成人工合成数据,在保证多样性的前提下,最大程度地扩展数据规模;其次,为解决样本数量增加所导致的标注工作量剧增问题,使用同步式策略在合成数据时同步产生标注,提高标注效率。

3.1 多重随机的内容差异化数据合成

数据增强是一种广泛应用于深度学习中的技术,旨在提升模型性能^[16]。通过对现有训练数据进行各种变换和扩充,有效增加样本的多样性,从而提供更丰富、更全面的训练数据集。传统数据增强方法通常通过图片翻转、调整明暗等来扩充数据,此类调整属于浅层调整,没有对关键特征即 DM 码内容进行变换,导致数据只是在数量上增大,而不能在内容上进行实质性的扩充,无法真正为深度识别提供多样性。

为了在增加图片集规模的同时保证图片集的差异化和丰富性,本课题提出通过多层随机构建内容差异的数据增强方案,从光照、位置、内容、方向多层次地模拟真实场景,提高识别的鲁棒性,降低识别时对位置和数量因素的敏感性。

3.1.1 合成数据背景设计

合成数据可分为含 DM 码的试管底部前景数据和其余含支架的背景数据。可以分别对前景和背景数据进行设计,再将两者叠加起来生成新的数据。

当支架未满载时,光线会穿过未放置试管的孔洞区域,导致成像光照不均。为模拟和复现这种真实光照环境,在背景设计时,对顶部光源打开状态下的全空支架进行成像,获取背景数据(即

图 1 中没有试管 DM 码时的状态)。

3.1.2 内容差异化的前景设计

差异化前景试管码构建是数据合成的重点部分,本课题对前景从局部到整体进行分层构建,即先构建单个试管码单元;后将多个码单元组成码目标群。因此前景设计根据这种层次分为以下几个阶段。

(1)构建试管码单元阶段。为保证编码内容多样性,随机生 8 位随机字符 $r(C)$ 作为内容,其中 r 均代表随机生成器, C 代表字符,结合 libdmtx 库编码生成 DM 码图像 g ,该图像与圆形试管底部叠加,形成试管码单元图片。考虑到实现场景中试管摆放方向完全随机,因此需要对试管码单元施加一个随机旋转角度 $r(\alpha)$ 如公式(1),进行旋转得到试管码单元 m 。

$$\begin{cases} g = \text{libdmtx}(r(C)) \\ \begin{bmatrix} m(x) \\ m(y) \end{bmatrix} = \begin{bmatrix} \cos r(\alpha) & \sin r(\alpha) \\ -\sin r(\alpha) & \cos r(\alpha) \end{bmatrix} \begin{bmatrix} g(x) \\ g(y) \end{bmatrix} \end{cases} \quad (1)$$

(2)构建码目标群阶段。目标群设计主要包含目标总数的设计和各个目标在支架的空间位置设计。为模拟实际操作中支架上试管数量和位置会因实验任务不同而存在较大差别,为此针对数量和位置设计如下:

首先,为了模拟不同目标总数差异性,假定支架试管总数设置为 $r(N)$ ($1 \leq r(N) \leq N_0$),值域为 1 到 N_0 之间, N_0 为支架能承载的最大试管数,试管总数用随机生成器生成,而 k 个试管对应的目标单元 m_j 组成的集合 $M\{m_1, m_2, \dots, m_j, \dots, m_k \mid k \leq r(N)\}$ 形成一个目标群,其中单元总数 k 不超过实际试管总数 $r(N)$;其次,为确保目标位置随机,考虑到支架大小统一,槽孔排列规则,视觉上类似矩阵,因此为每个目标单元 $m_k(x_i, y_j)$ 生成一个随机空间位置 $r(i, j)$, i 和 j 分别代表 m_k 所在行序号和列序号。这样目标码群内的每个目标单元可以随机获取一个二维行列位置,分配到具体的槽孔位置。

3.1.3 背景叠加与空间密度模糊

对支架背景和前景进行叠加,即将码单元图片在背景图片中进行逐一对齐,对齐的两个对象为:码单元图片中心点,和被分配槽孔的中心点。假设第 1 行第 1 列试管槽中心点(支架左上角位置)为 (x_1, y_1) ,则其第 i ($1 < i \leq n$) 行、第 j ($1 < j \leq m$) 列中心点位置可通过公式(2)所示方式确定,其中 n 和 m 分别是试管槽的最大行、列数, Δl 和

Δh 分别为两相邻试管槽中心点之间的宽度差和高度差。

$$\begin{cases} x_i = x_1 + (i-1) * \Delta l \\ y_j = y_1 + (j-1) * \Delta h \end{cases} \quad (2)$$

为模拟真实光照不均造成的图像模糊,通过统计目标群在支架中的分布位置,计算目标群放置的空间密度,来设置相应的模糊度。

(1)构建 n 行 m 列槽孔矩阵 G 。矩阵元素代表每个槽孔是否放置了试管,元素值为 1 代表有试管,为 0 则无试管。

(2)统计槽孔矩阵内的试管密度。针对 G 施加卷积操作 $\otimes$,设置卷积核 $h(t \times t)$ 大小为 t ,用卷积核在 G 中逐行逐列滑动并统计固定区域内放置试管总数,获得大小为 $(n-t+1, m-t+1)$ 大小的矩阵结果。

(3)计算试管空间密度。将结果展开成一维向量即 F 向量,分别计算其均值 $\text{mean}(F)$ 和标准差 $\text{std}(F)$ 如公式(3)所示,将标准差除以均值便可以得到试管密度 ρ 。 ρ 为大于 0 小于 1 的值,数值越大,代表试管摆放越不均匀,应施加的模糊度越高,反之数值越小,代表光照越均匀,应施加的模糊度越低。

(4)施加模糊效果。使用高斯低通滤波,假设滤波的截止频率范围为 (D_0, D_1) ,截止频率是影响模糊关键因子,截止频率越高,图片越清晰,因此结合试管密度,截止频率应设置为 d 如公式(3),这样的设计根据实际目标群分布,设置对应的模糊度。

$$\begin{cases} \rho = \frac{\text{std}(F(G_{n \times m} \otimes h(t \times t)))}{\text{mean}(F(G_{n \times m} \otimes h(t \times t)))} \\ d = (1-\rho) \times (D_1 - D_0) + D_0 \end{cases} \quad (3)$$

3.2 数据同步自动化标注

通过多重随机合成方案能有效扩充训练数据,但随之引发了新增数据的标注工作量过大的问题。使用人工分割方形码字区域(即掩膜)作为标注,耗时费力又容易出现差错。由于合成图片集的过程,和相关随机系数都是已知,这代表决定码字区域的大小、位置、方向都可知,因此可以重用合成流程,共享相关参数,来生成最终的标注,为此提出一种同步自动标注方法,生成对应的标注数据。如图 3 所示,该流程衍生自合成过程,可分为 3 部分:构建掩膜单元前景、构建掩膜群前景、叠加纯色背景。

(1)构建掩膜单元。当前 DM 码单元生成并旋转后,其码字区域的尺寸以及旋转角度 $r(\alpha)$ 便确定。同步生成大小、角度一致的方块区域,即掩膜单元图片。单元图片中像素值设为随机标签值 $l(1 \leq l \leq 255)$,其余区域设置为 0。

(2)构建掩膜群。为保证标注目标与合成图片中的目标一致,根据当前图片合成时的试管总数 $r(N)$,以及码目标群分配的槽孔行列位置 $r(i, j)$,同步生成同数目和分布的掩膜群。

(3)背景叠加。共享合成背景尺寸,但与合成背景不同,创建全 0 像素图片即黑色掩膜背景,而非支架背景图片,按照公式(2)准确计算槽孔位置,将掩膜图片叠加至黑色背景上。

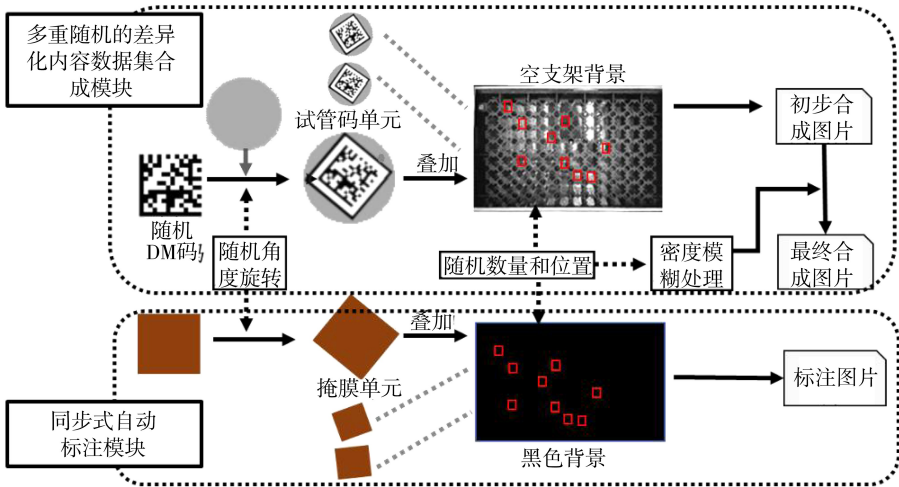


图 3 同步合成与标注方法示意图

Fig.3 Demonstration of simultaneous synthesis and annotation method

图 4 展示了使用 DeepDMCode 法实现的一组合成图像及其标注数据样例。

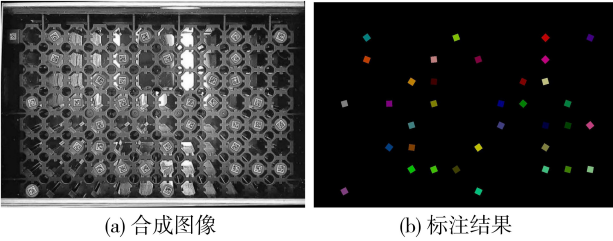


图 4 使用同步合同与标注方法所生成的合成图像及其同步式标注结果展示

Fig.4 Demonstration of the synthesis and annotation result by the proposed method

4 旋转矫正及解码

旋转矫正的目标是将图 5(a)所示的 DM 码修正为图 5(b)所示状态,以便调用标准解码库进行编码信息的读取。

具体来说,图 5(a)中彩色掩膜所示为上述方法所得分割结果,在掩膜图片中将 y 值最小点作为角点 $a:(x_a,y_b=y_{\min})$,以 a 为原点构建坐标系,此时 DM 码与 X 轴的夹角为 θ 可以按照公式(4)计算获得。根据得到的角度,以逆时针方向对码单元进行旋转,即可得到矫正后结果。最后使用第三方解码库,对其进行规范解码便可提取最终字符内容。

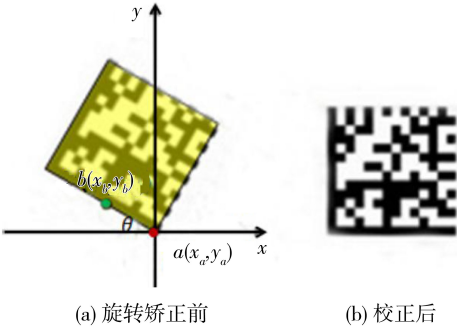


图 5 旋转矫正方法示意图

Fig.5 Demonstration of the rotation correction method

$$\theta=\begin{cases} \arctan\left(\frac{y_b-y_a}{x_b-x_a}\right), & x_b\neq x_a \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

5 实验及结果

实验均运行于一台显卡为 NVIDIA GeForce RTX 3060 (12 GB),CPU 为 Intel Core i7-11700 (2.50 GHz)的台式计算机上。输入数据分辨率为

1 600×1 200,训练数据集包括两部分:第一类是原始采集图片集,由 uEye UI-3013XC series 摄像机拍摄,共 516 张,使用手工标注;第二类是通过 2.1 节方法合成的新增数据,共 3 000 张,使用 2.2 节方法进行自动标注。验证集和测试集均为上述第一类数据,三者比例约为 8:1:1。

5.1 检测结果

图 6 展示为应用 DeepDMCode 进行 DM 码精定位的典型结果,其中各 DM 码实例均能够被正确识别(见图中彩色掩膜区域)。

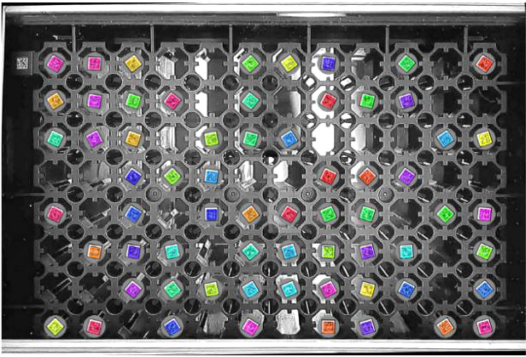


图 6 DeepDMCode 的 DM 码检测结果展示

Fig.6 Demonstration of DM code detection result by DeepDMCode

5.2 对比实验

选择目前条码识别主流的 3 种模型:YOLO^[12]、SegNet^[14]以及 CenterNet^[17]和 DeepDM-Code 进行对比。针对码识别中码定位和码内容提取两个阶段的结果进行对比,同时针对本课题数据合成模块设计消融实验,来验证其贡献。

5.2.1 检测结果对比

分别使用 3 种方法对模糊图像中的目标进行定位,选择相同局部区域展示图片结果如图 7 所示,可见,YOLO 无法识别码的方向,且存在误检,将具有相似直角和强边缘的背景错认为目标码;SegNet 检测到的区域边界较粗糙,且大于真实码区域范围;DeepDMCode 可较为准确的检测出模糊图片中对象边界,且不会受到相似性背景干扰,这得益于所提出的同步式差异性数据增强方法。

将定位结果与真实码字区域进行对照,使用 mIoU^[20]作为准确性指标,对 4 种方法进行数值评估。由于 YOLO 和 CenterNet 只适用于粗定位,无法精确反应码字区域,故用基于 L 型特征方法进行再处理,保证可比性,结果如表 1 所示。DeepDMCode 的 mIoU 值为 0.92,优于其他 3 种方法。

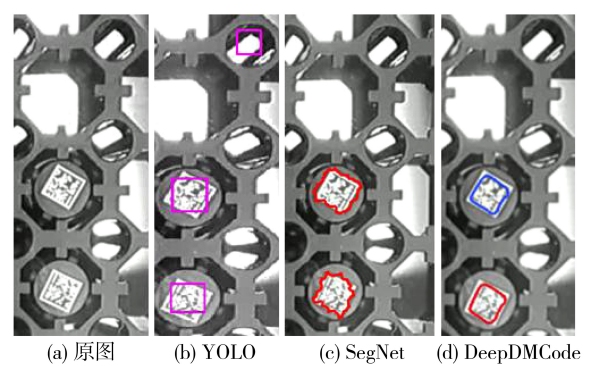


图 7 检测结果对比

Fig.7 Comparison of detection results

表 1 准确性对比	
Tab.1 Accuracy comparison	
方法	mIoU
YOLO	0.73
SegNet	0.80
CenterNet	0.74
DeepDMCode	0.92

5.2.2 DM 码内容提取结果对比

针对 3 种方法的定位结果,选择 libdmtx 进行解码,与真实结果进行对照,分别计算正确率和平均解码时间两个指标,统计结果如表 2 所示,DeepDMCode 正确率 90.8%,平均解码时间为 5.2 s;CenterNet 和 YOLO 目标检测虽然更快,但正确率较低。

5.2.3 消融实验

为验证本课题数据集成模块对码定位的贡献,使用仅改变对比度和明暗的传统数据增强方法进行 DM 码定位,与内容差异化增强后定位结果进行对比,计算其 mIoU 指标。从表 3 中可知,在传统增强数据集上运行结果的准确性较 DeepDMCode 约低 10%,这证明内容差异化合成模块

参考文献:

[1] 邓惠朋,王毅,董晓文. 汉信码与 PDF417、数据矩阵码(DM 码)、QR 码的对比[J]. 中国自动识别技术,2021(5):50-53.

[2] WUDHIKARN R,CHAROENKWAN P,MALANG K. Deep learning in barcode recognition:a systematic literature review[J]. IEEE Access,2022,10:8049-8072.

[3] 王根源,吴祥坤,胡坤,等. DataMatrix 码的嵌入式识别算法[J]. 吉林大学学报(信息科学版),2016,34(3):419-426.

[4] RALEVIC NM,DRAZIC S,OBRADOVIC R. The Hough transformation of rectangle[C]//2008 6th International Symposium on Intelligent Systems and Informatics. Subotica,Serbia: IEEE,2008:1-4.

的嵌入,能很好地改善条码定位结果,进而提高后续码字内容提取的精确度。

表 2 DM 码解码结果对比		
Tab.2 Comparison of DM code decoding results		
方法	正确率/%	平均解码时间/s
YOLO	70.1	4.4
SegNet	78.2	8.3
CenterNet	71.3	3.5
DeepDMCode	90.8	5.2

表 3 消融实验结果	
Tab.3 Ablation results	
方法	mIoU
传统增强+Mask R-CNN	0.81
内容差异化增强+Mask R-CNN	0.92

6 结语

以试管-支架系统作为研究对象,基于分割策略对输入图像中的所有 DM 码进行定位和旋转矫正,实现编码信息的读取。本研究的主要贡献包括:(1)提出一种多重随机的内容差异化数据合成方法 DeepDMCode。(2)部分重用合成流程,提出一种同步式的自动标注方法,共享随机参数,提高标注效率。(3)利用分割结果,提出一种旋转矫正方法,使得通过标准解码库对 DM 码编码信息进行提取成为可能。实验表明,该方法能有效克服图片分辨率低、目标数量和分布不确定、光照条件动态变化等因素所造成的识别困难,减少不必要的人工干预。后续工作将专注于提升方法的速度,以实现 DM 码的实时读取。

- [5] 李东,郭浩铭,田劲东,等. 改进的 DataMatrix 码 L 形边精确定位方法[J]. 深圳大学学报(理工版),2018,35(2):151-157.
- [6] 关博熠. 复杂环境下 Data Matrix 码的定位与倾斜校正[D]. 哈尔滨:哈尔滨理工大学,2018.
- [7] 邹承明,薛榕刚. 融合 GIoU 和 Focal loss 的 YOLOv3 目标检测算法[J]. 计算机工程与应用,2020,56(24):214-222.
- [8] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal net-works[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6):1137-1149.
- [9] 易帆,李功燕,许绍云. 基于多尺度特征的条形码快速检测算法[J]. 计算机工程与设计,2020,41(1):220-225.
- [10] 谢志峰,吴佳萍,章曙涵,等. 基于深度神经网络的烟码智能识别方法[J]. 计算机辅助设计与图形学学报,2019,31(1):111-117.
- [11] 周勉,刘秉瀚. 基于深度学习的多条形码检测与实现[J]. 信息技术与网络安全,2018,37(9):65-69.
- [12] HANSEN D K, NASROLLAHI K, RASMUSEN C B, et al. Real-time barcode detection and classification using deep learning[C]//Proceedings of the 9th International Joint Conference on Computational Intelligence. Funchal, Madeira, Portugal: SCITEPRESS - Science and Technology Publications, 2017:321-327.
- [13] ZHARKOV A, ZAGAYNOV I. Universal barcode detector via semantic segmentation[C]//2019 International Conference on Document Analysis and Recognition (ICDAR). Sydney, NSW, Australia: IEEE, 2020:837-843.
- [14] 胡钊,赵江铭,张小琛,等. 利用 SegNet 的非标准 DM 码的区域提取[J]. 机械设计与制造,2021(7):62-65.
- [15] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017:2980-2988.
- [16] ZHANG JH, MIN X K, JIA J, et al. Fine localization and distortion resistant detection of multi-class barcode in complex environments[J]. Multimedia Tools and Applications, 2021, 80(11):16153-16172.
- [17] ZHOU X Y, WANG D Q, KRÄHENBÜHL P. Objects as points[EB/OL]. [2019-04-16]:arXiv:1904.07850. <https://arxiv.org/abs/1904.07850>.

(责任编辑:方素华)