

基于模型的混合现实装配验证技术研究

吴浩天, 杨振兴

(中国空空导弹研究院, 河南 洛阳 471000)

摘要: 研究基于高精度数字化模型的混合现实装配验证技术, 通过训练深度学习模型, 并部署至图形工作站和头显交互一体机中验证了混合现实装配技术的可行性。最后, 实际应用至装配过程中, 对比应用前后得出, 本研究能够提高 10% 左右的装配效率, 同时提高装配效果。

关键词: 混合现实; 深度学习; 装配验证; 仿真

中图分类号: TP391.4; TP391.9 **文献标志码:** A

1 引言

随着科技的发展和制造业的转型升级, 传统的装配技术难以满足现代产品高精度、高效率的装配需求, 特别是在复杂的机械装配领域。传统的装配方式严重依赖人力, 生产周期长, 产品质量参差不齐。基于新的科技手段, 寻求新型的装配技术克服此类问题是当前制造业特别是军工制造业亟需解决的问题^[1-2]。国防科工局和中航工业集团在智能制造技术 5 级模型中要求精密制造业主动拥抱未来, 建设自动化、数字化、智能化的高柔性产线。

近年来, 虚拟现实技术^[3]的崛起为装配技术的发展提供了新的思路和方法。虚拟现实技术是通过创建虚拟三维环境, 使用户能够沉浸在虚拟世界中与虚拟物体进行实时交互的技术。这种方式不仅提高了用户的参与度和操作体验, 而且为装配过程的模拟、优化和验证提供了基础。

人工智能技术^[4]作为现阶段的通用技术, 已广泛应用于传统目标检测和缺陷检测中, 如无人机任务目标实时检测与跟踪^[5]、无损探伤缺陷智能辅助判决^[6]、医疗

影像病灶识别^[7]等。端到端的单阶段目标检测神经网络架构也被应用至嵌入式端, 实现了低算力平台的工业化部署与应用。

基于模型的混合现实 (Mixed Reality, MR) 装配验证技术旨在通过构建并训练精确的深度学习模型, 基于数字化产品模型进行识别和定位, 实现虚拟装配和真实装配的无缝对接, 从而提高装配效率、降低生产成本、保证装配质量。搭建基于图像识别的混合现实装置硬件平台, 包括图形工作站和混合现实头显交互一体机 (以下简称“MR 设备”), 并部署训练完成的深度学习算法模型。通过 MR 设备的单目相机实时采集现实中的装备目标, 计算相关目标的空间坐标位置, 为三维虚拟影像的展现提供位置信息, 验证复杂场景下的虚拟影像映射技术。

2 基于模型的混合现实装配验证技术

基于模型的混合现实装配技术利用相机获取产线信息, 在服务器端处理检测目标, 同时生成指令。MR 设备根据指令识别和追踪目标。

近几年,目标检测随卷积神经网络发展迅速,通过自适应学习图像特征,对手工标签需求降低,且能够保证目标检测精度。本文将以单阶段目标测量方法为主要依据,根据锚框(Anchor Box)尺寸已不符合目标数据的现状,采用K均值聚类算法(K-means Clustering Algorithm,以下简称“K-means 算法”)聚类新锚框尺寸,目的是将输入包含 N 个标注点的数据集分为 K 个聚类,将数据集中所有点划分到距离各个聚类点最近的聚类中心中,以此约束锚框尺寸;接着引入了广义交并比的损失函数以实现更为精确的拟合;并引入特征金字塔结构以提高计算的精度。

坐姿示意图如图1所示,本文使用深度学习模型实现作业目标的相对位姿识别。使用双目相机对待采集目标进行纵向的距离识别,以识别出相关目标在同一画面内的距离信息,即图1中 z 轴坐标信息;运用深度学习算法来对目标进行投影平面内的定位,即 x 、 y 轴坐标。

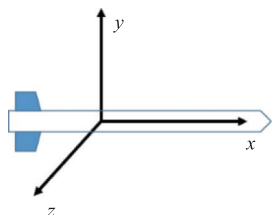


图1 坐标示意图

获取到作业目标的位姿信息后,通过MR设备自带的RGB相机实时采集视野影像信息。影像信息传输至高性能计算端,通过已部署的卷积模型识别相关特征,解算目标位置信息,将目标位置信息传出至本地状态文件,包括目标中心点坐标(X_{rgb} , Y_{rgb})、目标像素尺寸(宽度、高度)。再通过标定好的位置信息计算焦距值,引入实时变化的目标像素宽度值,经计算可以得到目标的距离信息,将目标的平面坐标经比例尺变化后的真实坐标与目标的距离信息拼凑成目标点的三维坐标信息。将该信息传输至MR设备,即可进行虚拟影像投射。目标检测系统的工作原理示意图如图2所示。

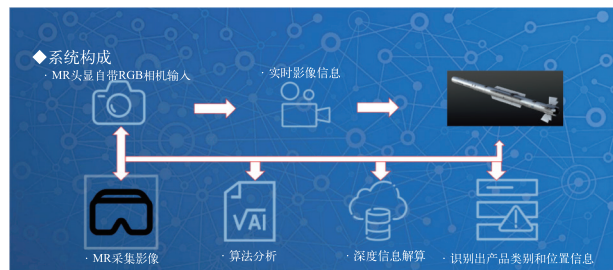


图2 目标检测系统的工作原理示意图

3 数据样本采集与标注

装备在生产维护过程中,摆放位置和角度不尽相同,且装备表面由金属制成,喷漆后表面纹理不明显,存在不同位置相同效果等特点。此外,影响因素还包括环境光照,导致检测失误与定位偏差,本文使用卷积神经网络学习装备表面特征减少上述因素对结果造成的影响。

本文中使用头戴式MR设备作为整个机器视觉装备目标检测系统的载体,通过采用单目RGB相机,直接采集装备目标外部信息,同时,该相机具备一定的实时场景描述能力。

为了准确识别出装配对象,实现对特定场景下的目标装备对象实时进行准确识别,采集的样本数据分布如图3所示。采样数量2 954张,标注框集中在视野中心位置,样本分布为宽高比大者居多,且为长条形,整体样本采样效果,符合实际应用场景。

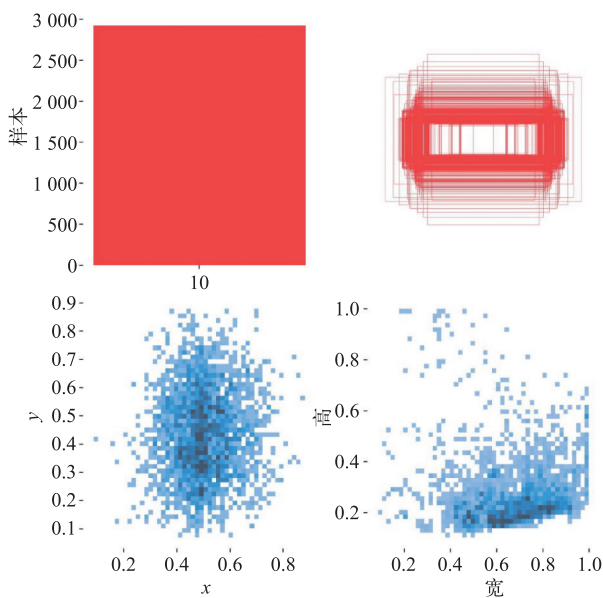


图3 数据集训练分布图

4 深度学习模型设计

4.1 目标识别与检测网络设计

目标检测技术^[8]旨在模仿人类视觉系统的功能,通过算法和模型识别和定位图像或视频中的具体物体。本质是对视觉信息进行解析,使计算机能够理解图像中的内容。

本文采用Darknet多层卷积神经网络^[9](如图4所示)作为特征提取模块来对MR设备采集的影像信息

进行处理和目标对象识别。由于实时性要求，本文采用单阶段端到端的目标检测网络作为基线网络，实现对独立目标的实时追踪与检测，基础模型框架如图 4a 所示。

对基线网络进行迁移训练，本文选用单阶段目标检测网络模型作为基线网络，兼顾实时性与准确性。此外由于装备目标喷漆、表面纹理特征不明显和反光等问题，本文将目标探测问题转变为逻辑回归问题，先分割输入图像的所有栅格，将各个栅格分别处理，最后用非极大值抑制技术选择正确的预选框。结果信息包括了目标的中心点位置、高度和宽度等。而在相机视角下，目标检测的概率为 $P_{\text{confidence}}$ ：

$$P_{\text{confidence}} = P_{\text{object}} \times I_{\text{ou}} \quad (1)$$

式中， P_{object} 为布尔值表征边界框中是否包含检测物体。

在基本的图像特征提取方面，通过输入像素自动转换，将原始输入像素转化至适配本文模型网络的输入像素值。整个网络的原理如图 4b 所示，结构如图 4c 所示。本文参考残差网络，在不同层间配置直联链路，不影响速度的同时提取更多图像特征。

本模型检测功能依赖特征金字塔网络。不同层的特征图构成特征金字塔网络不同的级联，以此合并上采样的特征，并借助卷积网络将修正检测装备目标图片，以不同大小的特征图对应检测目标，提升检测性能，获取目标分类信息和实际位置。

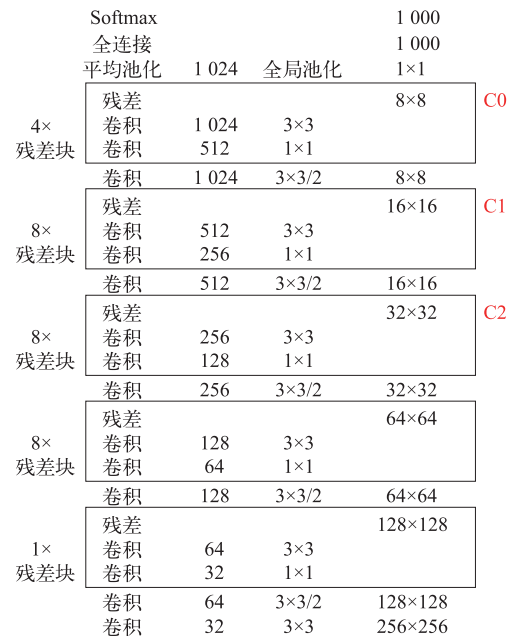
4.2 K-means 算法更新训练锚框尺度

针对设计网络锚框尺度，原始锚框尺寸与目标尺寸存在差异，网络训练效率和检测准确率较低。为上述问题，本文采用 K-means 算法确定锚框方案。聚类标准使用欧式距离：

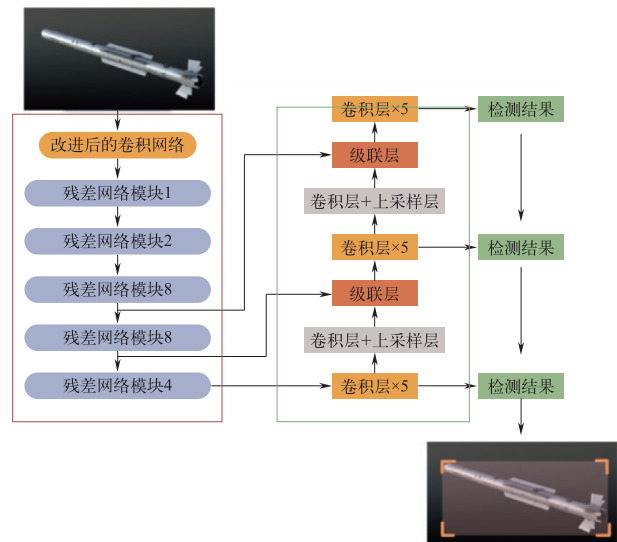
$$f(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2)$$

训练卷积神经网络训练时，使用池化层方法降低计算成本，提高模型鲁棒性。本文引入空间金字塔池化算法解决网络训练过程中输入图像尺寸唯一的现象，引入该方法后图像输入尺寸不受限制，可一次性计算输入图像中不同大小的特征区域，池化后输出融合特征，实现更高效特征提取，原理如图 5 所示。

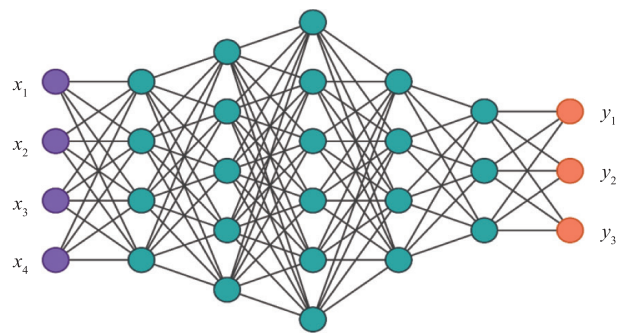
空间金字塔池化方法降低了训练时由于固定输入图像尺寸的裁剪、缩放等操作带来的特征丢失概率。将空间金字塔池化模块安排在改进后的卷积网络特征提取之后，对 3 种不同尺度的目标识别均产生影响。



a) Darknet 基础模型框架



b) 网络原理



c) 模型示意

图 4 Darknet 神经网络架构

通过目标检测模型的推理，输出 MR 设备中 RGB 相机视野下的目标中心坐标大小和目标尺寸大小，并存储在本地的状态文件中。具体的目标检测推理过程如图 6 所示。

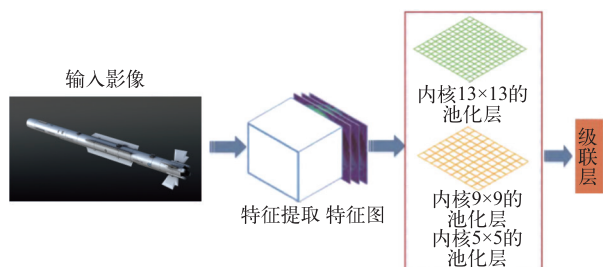


图 5 空间金字塔特征池化原理

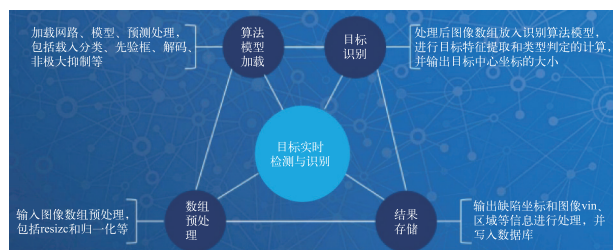


图 6 目标检测模型推理过程示意图

5 模型训练与部署

训练数据时使用 COCO 数据集^[10]，由于在基线网络的基础上，需要进行迁移学习才能够获得稳定的目标检测模型，本项目中的目标检测模型训练为有监督式的学习过程。数据使用软件自动标注，通过大量影像标注，可得到相应带标签的图片集合，图片标签包含目标在图片中的位置和类别信息。通过加载标注数据文件，来对基线模型进行训练，经过数次迭代和调优，训练损失函数误差值逼近 0.3 以下，即可完成相关训练，损失函数的设计如下：

$$L_{\text{box}} = \lambda_{\text{coord}} \sum_{i=1}^{S^2} \sum_{j=0}^B I_{i,j}^{\text{obj}} \cdot I_{i,j}^{\text{box}} \quad (3)$$

$$L_{\text{cls}} = \lambda_{\text{class}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{i,j}^{\text{obj}} \cdot I_{i,j}^{\text{cls}} \quad (4)$$

$$L_{\text{obj}} = \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{i,j}^{\text{noobj}} \cdot I_{i,j}^{\text{obj}} + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{i,j}^{\text{obj}} \cdot I_{i,j}^{\text{obj}} \quad (5)$$

$$Loss = L_{\text{box}} + L_{\text{cls}} + L_{\text{obj}} \quad (6)$$

将训练稳定后的权重文件拷贝至推理代码中，形成固定的推理模式，在系统运行过程中自动加载相关权重模型，实现对识别对象的实时检测与跟踪。训练过程如图 7 所示。

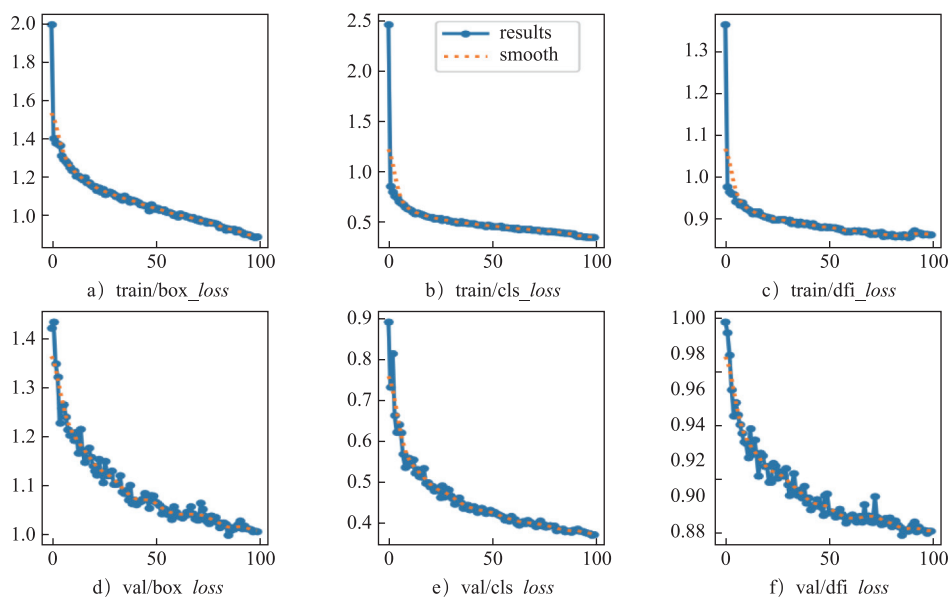


图 7 训练过程 Loss 值变化

对标注完成的样本数据进行训练，通过一定时间的训练使模型达到收敛，确保训练出的模型具备较高的识别准确度。模型训练结果如图 8 所示，定位损失（box_

loss）接近于 0，类别损失（cls_loss）逐步趋近于 0，精度（precision）和召回率（recall）较高，mAP50 和 mAP50-95 接近于 1。mAP50-95 表示在不同 IoU 阈值（从

0.5 到 0.95，步长 0.05）上的平均 mAP 。通过对模型训练时每次迭代结果的分析，可以得知该训练模型在训练过

程中的平均召回率均已达到 98% 以上，在不同精度和召回率条件下的性能较为优异。

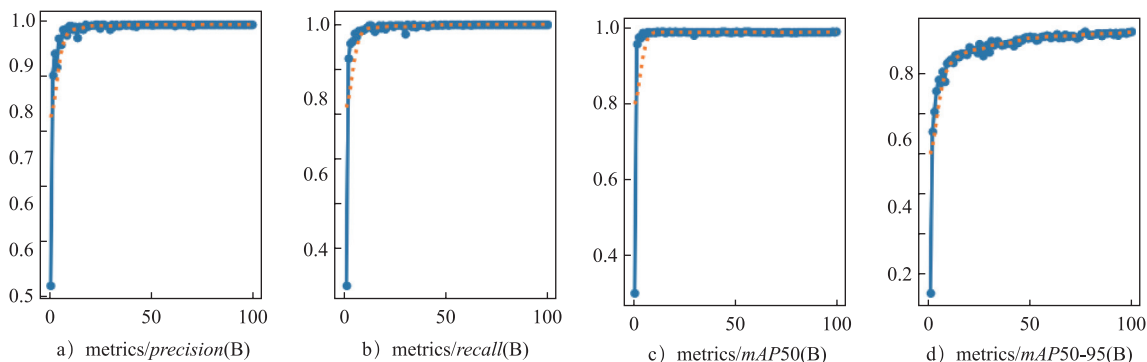


图 8 模型精度、召回率和 mAP 变化曲线

6 混合现实装配技术验证

6.1 混合现实装配技术验证

整体智能目标识别程序应用部署在高性能下位机内，通过开机自运行调取相关视觉目标检测算法，实时检测装备目标，反馈目标的位置坐标和尺寸。

本文中采用单目相机测距方法来实现对目标距离位置的测算。单目相机测距这里使用的焦距的不变性，来间接测量目标距离。焦距定义：

$$F = \frac{P_0 \times D_0}{W_0} \quad (7)$$

式中， D_0 为实际宽度测算距离， P_0 为视野全宽度的像素值， W_0 为实际宽度值。

对相机进行标定时，需准确测量出 MR 设备 RGB 相机至实际目标物体的距离 D_0 ，同时获取标定状态下的文本文件中的目标物体宽度 P_0 ，从而方便求取 F 数值。实时状态的目标物体距离相机的距离测算为

$$D = \frac{F \times W_0}{P} = Z_{\text{Real}} \quad (8)$$

式中， D 为实际目标距离， Z_{Real} 为转换至实际坐标系下的 Z 轴数值， P 为目标检测视野宽度像素值。

通过像素比例尺 ($f_0 = W_0/P_0$ ，即每个像素占实际坐标系的数值关系) 来推理出 RGB 相机中的图像与实际坐标的关系表述，可以容易得出

$$X_{\text{Real}} = X \cdot f_0 \quad (9)$$

$$Y_{\text{Real}} = Y \cdot f_0 \quad (10)$$

将上述两项与 Z 轴解算数值进行拼凑，可得到最终

目标的实际三维坐标值 (X_{Real} , Y_{Real} , Z_{Real})。

在开展目标测距之前，需要对焦距 F 进行测算和预制，这里称之为距离标定。首先，装备目标需水平安放至某一平面或支撑装置上，将 MR 头戴式设备携带至测试人员头部，正常开启高性能计算机，系统自动调起智能目标检测程序，实时检测装备目标；其次，站在装备目标正中心位置，开始逐步沿着对装备目标的中心垂直线移动后退，小步移动至当充满视野范围内时，停止移动；再次，记录当前有目标检测程序反馈来的目标宽度像素值和当前移动位置至装备目标中心位置实际距离长度，精确至小数点后 1 位；最后，将采集到的信息注入至公式 (7)，得到最终的焦距数值，完成距离标定。

6.2 实际数据集采样方式

为满足模型训练要求数据集，以水平放置的装备目标为主。同时选择几种不同放置方向和角度的装备以更好地体现本文所设计方法的检测性能，照片采集方式示意图如图 9 所示，详细描述见表 1。

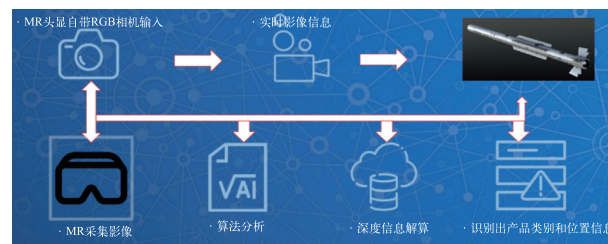


图 9 照片采集方式示意图

表 1 数据集采样方式

序号	位置	X-Y 旋转 角度	采集距离 /m						
1	水平向左	0°	1	2	3	4	5	6	7
2	水平向左	30°	1	2	3	4	5	6	7
3	水平向左	60°	1	2	3	4	5	6	7
4	水平向右	0°	1	2	3	4	5	6	7
5	水平向右	30°	1	2	3	4	5	6	7
6	水平向右	60°	1	2	3	4	5	6	7

6.3 样本采集布置方案

通过对增加数据在不同光照条件下的样本类别，选择不同类型光照方式，同时选择不同的背景形式，丰富样本层次。详细描述见表 2。

表 2 光照类别示意

序号	光照条件类别
1	前景与背景对比饱和度高
2	曝光程度高
3	弱光线条件
4	正常光照条件

采集样本时，需将相机进行标定，对目标与相机之间的距离进行设置，通过不断移动相机与目标之间的距离，分别采集不同景深的装备目标影像；对装备目标在 X-Y 水平面内进行绕 Z 轴旋转，分别依次旋转 30°、60°，制作实际操作中的倾斜样本；采用人工设置光源形式，对被采样装备目标进行人工模拟打光，从不同角度和位置，对装备目标打光，设置光照强度，从而模拟实际场景中的不同类别光照，丰富采样类别的层次度。

6.4 数据预处理及数据增强

训练数据来自于对类似目标对象进行的多角度图像采集，还包含大量的仿真目标数据，避免出现过拟合现象，导致无法检测真实目标。同时借助图像增强算法弥补数据量不足，提高模型泛化能力。本文通过旋转 90°、180°

和 270°，以及添加高斯噪声将目标图像生成 4 个附加影像。

目标检测模型的边界回归机制中，边界框约束包括中心点坐标 (x, y)、宽度 (d) 和长度 (l)。归一化处理边界框参数可降低边界框训练难度，具体回归方式如图 10 所示。

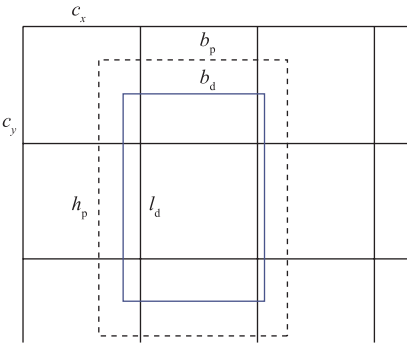


图 10 目标检测位置预测原理

边界框的预测的参数由式 (11) 计算得出。

$$\begin{cases} b_x = \sigma(t_x) + c_x \\ b_y = \sigma(t_y) + c_y \\ b_d = l_d e^{t_d} \\ b_p = h_p e^{t_p} \end{cases} \quad (11)$$

式中， c_x 、 c_y 分别为边界框中心点坐标距离第 1 个栅格的 x 方向和 y 方向数目； l_d 、 b_d 分别为预测边界框对应的长和宽； t_x 、 t_y 分别为预测的边界框相对于真实边界框 x 方向和 y 方向的偏移量； $\sigma(x)$ 函数为激活函数，用于偏移量的归一化； t_w 、 t_h 为预测边界框相对于真实边界框长和宽的尺度缩放； b_p 、 h_p 分别为锚框的宽和高，初始值由聚类方法得到。在基本的图像特征提取方面，本文采用含有多个卷积层的 Darknet 网络结构，在层与层之间设置了快捷链路，在保证速度的情况下，更多地提取了图像特征。

6.5 技术应用验证

基于模型的混合现实装配技术验证采用对比方法，以 10 名装配人员包括 5 名熟练操作工和 5 名一般操作工装配两种组件各 200 次，对比装配时间和经过环境试验后组件的异常率。应用技术前后效果见表 3。

表 3 应用技术前后的效果

类别		应用前平均时间 /s	应用后平均时间 /s	应用前异常率 (%)	应用后异常率 (%)
熟练工	组件 A	6.58	7.24	2	1
	组件 B	94.30	89.65	2	2
一般工	组件 A	8.25	9.76	3	1
	组件 B	112.32	105.43	2	1.5

7 结束语

本次研究通过对深度学习训练模型的测试,结合测试数据进行工作原理的探究,将某型产品工艺件作为目标,反复测试后,确定运用卷积神经网络算法模型能够实现对指定产品进行图像采集、识别和定位。测试中的训练数据来自于对类似目标对象进行的多角度图像采集,还包含大量的仿真目标数据,在数据训练的过程中,原始锚框尺寸与目标尺寸有偏差,网络训练效率低,准确率差,因此使用 K-means 算法约束锚框,规约数据点至聚类中心降低影响。

训练稳定后的权重文件形成固定的推理模式,在系统运行过程中自动加载相关权重模型,实现对识别对象的实时检测与跟踪,同时将输出目标类别与中心点坐标位置信息保存至状态文件中完成混合现实的装配过程。

最后,在实际装配时应用技术验证效果,从平均装配时间和装配后经过环境试验后组件异常率可以得出结论,本研究能够提高装配效率同时提高装配准确率,具有应用意义。IM

参考文献

- [1] 中华人民共和国工业和信息化部. 智能制造能力成熟度模型: GB/T 39116—2020[S]. 北京: 中国标准出版社, 2020.
- [2] 晋严尊, 王中一, 汤辉, 等. 智能制造技术在机载武器制造中的应用 [J]. 航空兵器, 2021, 28 (6): 1-6.
- [3] 杨青, 钟书华. 国外“虚拟现实技术发展及演化趋势”研究综述 [J]. 自然辩证法通讯, 2021, 43 (3): 97-

106.

- [4] 德勤中国, 中国科学技术信息研究所, 中国人工智能学会, 等. 制造业+人工智能创新应用发展报告 [R/OL]. (2021-10-15) [2024-05-17]. <https://www2.deloitte.com/content/dam/Deloitte/cn/Documents/energy-resources/deloitte-cn-eri-manufacturing-artificial-intelligence-innovation-application-development-report-zh-211015.pdf>.
- [5] 江波, 屈若锟, 李彦冬, 等. 基于深度学习的无人机航拍目标检测研究综述 [J]. 航空学报, 2021, 42 (4): 137-151.
- [6] 赵天伟, 林莉, 张东辉, 等. 机器学习在无损检测中的应用案例浅析 [J]. 无损探伤, 2020, 44 (3): 1-4, 11.
- [7] 伍佳莉, 李东伦, 唐泳, 等. 人工智能辅助医学影像识别技术的应用研究进展 [J]. 现代医药卫生, 2022, 38 (4): 603-607.
- [8] 帅泽群, 李军. 基于深度学习的目标检测研究 [J]. 汽车工程师, 2021 (5): 11-14.
- [9] MIREA M, WANG V, JUNG J. The not so dark side of the Darknet: a qualitative study[J]. Security Journal, 2019, 32 (2): 102-118.
- [10] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: Common Objects in Context[J/OL]. (2015-02-21) [2024-05-17]. <https://arxiv.org/pdf/1405.0312>.

收稿日期: 2024-10-23