

# 深度学习在混合现实车间巡检中的应用<sup>①</sup>



刘云江<sup>1</sup>, 关 慧<sup>1,2</sup>, 王鸿亮<sup>3,4</sup>, 王继娜<sup>5</sup>

<sup>1</sup>沈阳化工大学 计算机科学与技术学院, 沈阳 110142)

<sup>2</sup>(辽宁省化工过程工业智能化技术重点实验室, 沈阳 110142)

<sup>3</sup>(中国科学院 沈阳计算技术研究所, 沈阳 110168)

<sup>4</sup>(中国科学院大学, 北京 100049)

<sup>5</sup>(辽宁省先进装备制造业基地建设工程中心, 沈阳 110001)

通信作者: 关 慧, E-mail: h.guan@syuct.edu.cn

**摘 要:** 传统的车间巡检方法主要依靠人工检查及记录结果, 过程繁琐且无法实时共享全过程. 为了提高工作效率, 将深度学习应用到混合现实车间巡检中. 采用了深度学习与混合现实技术相结合的方式, 利用 ResNet 网络对车间设备进行分类识别, 完成分类识别后, 再利用 HoloLens 的空间感知能力定位到该设备进行确认, 最后显示该设备的基本信息、运行状态和报警等信息. 经实验验证, 与传统的车间巡检方法相比, 具有较高识别率的 ResNet 有效地过滤了噪声, 提高了 HoloLens 的利用率和识别率, 同时也提高了巡检人员的工作效率.

**关键词:** 混合现实; ResNet; 深度学习; 车间巡检; HoloLens; 卷积神经网络

引用格式: 刘云江, 关慧, 王鸿亮, 王继娜. 深度学习在混合现实车间巡检中的应用. 计算机系统应用, 2022, 31(5): 118-123. <http://www.c-s-a.org.cn/1003-3254/8515.html>

## Application of Deep Learning in Mixed Reality Workshop Inspection

LIU Yun-Jiang<sup>1</sup>, GUAN Hui<sup>1,2</sup>, WANG Hong-Liang<sup>3,4</sup>, WANG Ji-Na<sup>5</sup>

<sup>1</sup>(College of Computer Science and Technology, Shenyang University of Chemical Technology, Shenyang 110142, China)

<sup>2</sup>(Key Laboratory of Industrial Intelligence Technology on Chemical Process of Liaoning Province, Shenyang 110142, China)

<sup>3</sup>(Shenyang Institute of Computing Technology, Chinese Academy of Sciences, Shenyang 110168, China)

<sup>4</sup>(University of Chinese Academy of Sciences, Beijing 100049, China)

<sup>5</sup>(Liaoning Advanced Equipment Manufacturing Base Construction Engineering Center, Shenyang 110001, China)

**Abstract:** The traditional method of workshop inspection mainly relies on manual check and recording, which is cumbersome and cannot be shared in real time. For higher work efficiency, deep learning is applied to mixed reality workshop inspection. It is combined with mixed reality technology, and the ResNet network is used to classify and identify workshop equipment. After classification and identification, HoloLens' spatial perception ability is leveraged to locate and confirm the equipment. Finally, equipment basic information, operating status, and alarms are displayed. The experimental results show that compared with traditional workshop inspection methods, ResNet, with a high identification rate, can effectively filter noises, improve the utilization rate and identification rate of HoloLens, and consequently improve the work efficiency of inspection personnel.

**Key words:** mixed reality; ResNet; deep learning; workshop inspection; HoloLens; convolutional neural network (CNN)

随着智能制造与信息技术的融合发展, 智能制造正在成为制造业发展的主要趋势, 巡检是智能制造发

展的重要一环, 是车间设备正常运行的重要保障<sup>[1]</sup>. 混合现实技术作为信息技术的一种, 可以有效地辅助企

① 基金项目: 辽宁省教育厅 2021 年度科学研究经费项目 (LJKZ0434); 沈阳市重大科技成果转化专项 (20-203-5-40)

收稿时间: 2021-08-03; 修改时间: 2021-08-31; 采用时间: 2021-10-09; csa 在线出版时间: 2022-04-11

业提高生产效率,降低运维成本,提升员工专业技能水平,加速产业智能升级<sup>[2]</sup>.混合现实技术(MR)是虚拟现实(VR)技术的进一步发展,它不同于虚拟现实技术的虚拟化信息的单独存在,也不同于增强现实(AR)技术的现实信息与虚拟信息的简单叠加<sup>[3]</sup>,而在于物理环境与数字世界融合在一起,使其在视觉和听觉上与所处世界融为一体.

在大型生产车间的巡检中,邱俊豪等设计的设备点检与巡检管理系统利用Web系统和小程序解决了现存点检流程中主要依靠纸质记录各类信息的低效率问题<sup>[4]</sup>.在图书馆设备运维方面,孙臻等利用混合现实设备辅助运维人员处理故障,提高了设备运维效率<sup>[3]</sup>.在智能车间巡检方面,刘敏洋等提出了一套工业智能运维系统,通过对混合现实技术的应用达到工厂智能化管理的目的<sup>[2]</sup>.相较于传统的车间巡检方法而言,从Web系统和小程序到混合现实技术的应用已经有了长足的进步,针对于混合现实技术的应用而言,李玄基在基于HoloLens增强现实的关键技术研究与实现的目标空间定位实验中,设备的真实空间位置同目标空间定位计算出的位置进行对比发现存在着一定的空间误差范围<sup>[5]</sup>,此时,若加上设备的类别信息则可以提高识别不同设备的能力.为了提高设备的识别率,本文将深度学习应用到混合现实车间巡检中.传统的浅层次结构模型在处理图像数据方面表现较差,特征提取难以满足要求,而深度学习在弥补了这一缺陷,提取图像深层次的结构特征方面更具有优势<sup>[6]</sup>.本文采用了卷积神经网络进行设备的分类识别,完成分类识别后再通过位置信息确认设备,提高了HoloLens利用率和识别率.

混合现实技术在医学、工业制造和军事等领域已经得到了广泛应用,作为混合现实技术实现的载体,其代表产品为Magic Leap公司的Magic Leap One、微软公司的HoloLens<sup>[7-10]</sup>.本文采用了微软公司于2019年2月25日最新发布的HoloLens2设备进行实验.

## 1 混合现实的车间巡检架构

### 1.1 混合现实车间巡检

首先,边缘设备的基本信息、运行状态等数据通过网卡收集,然后通过控制系统,最终存储到服务器上的MySQL数据库中;其次,用于进行深度学习训练的数据集存储在服务器上,训练采用深度学习框架PyTorch搭建ResNet50网络,将训练完成后的权重文件保存在

服务器上,并使用PyTorchServe搭建Web服务,提供设备分类识别功能;巡检人员头戴HoloLens在车间巡检到某边缘设备,通过HoloLens的摄像机功能发送该设备图像到服务器,服务器通过设备分类识别功能验证后,再次发送由HoloLens的空间定位功能获取的位置信息,并与数据库中的设备位置信息进行比较,成功则返回该设备的在数据库中的信息;最后,HoloLens显示三维信息并通过网络同步到显示器上.

### 1.2 整体架构

本系统共包含了5层结构,即边缘设备层、数据接入层、数据存储层、应用实现层和应用交互层,系统整体架构图如图1所示.

边缘设备层主要包括数控机床、数控系统及工业机器人等设备,通过部署的网卡采集数据,是实现车间巡检的基础;数据接入层对采集到的数据进行管理与整合,通过控制系统将数据存入数据存储层的MySQL数据库中;应用实现层通过混合现实开发包(MRTK)开发HoloLens应用,在服务器端部署PyTorchServe服务,实现设备分类识别功能;应用交互层提供了交互设备,包括HoloLens以及同步HoloLens交互界面的显示器,这些应用平台可以辅助巡检人员实时掌握车间运行状态,维护车间正常运行.

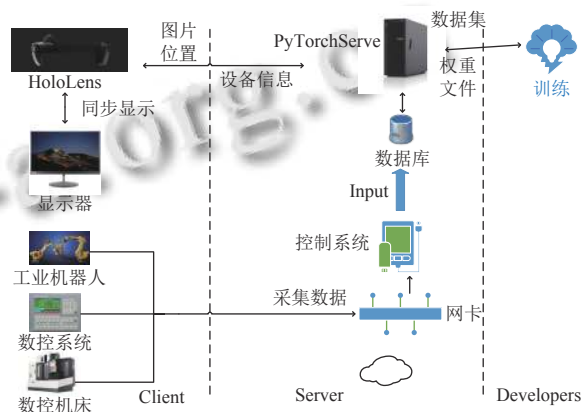


图1 系统整体架构图

### 1.3 巡检功能

空间定位功能:在车间入口处启动并佩戴HoloLens,以此时的位置为世界坐标原点<sup>[5]</sup>,实际的车间设备的位置是相对于原点的;当车间巡检人员佩戴HoloLens移动到目标设备前,可以通过空间定位功能计算出该目标设备的相对位置,通过比对数据库中的位置信息判断出当前设备.

分类识别功能: 在巡检作业时, 由于 HoloLens 的空间定位能力存在范围误差, 而且需要检查的设备并非所有设备, 需要排除一些噪声. 因此, 为了提高 HoloLens 的识别率以及过滤这些噪声, 在进行巡检作业中, 需要通过分类识别功能确认设备类别, 提高 HoloLens 的利用率和识别率.

除了显示信息, 用户还可以通过人机交互功能对当前巡检设备添加日志信息, 并将日志信息存入数据库进行持久化处理, 如图 2 所示.

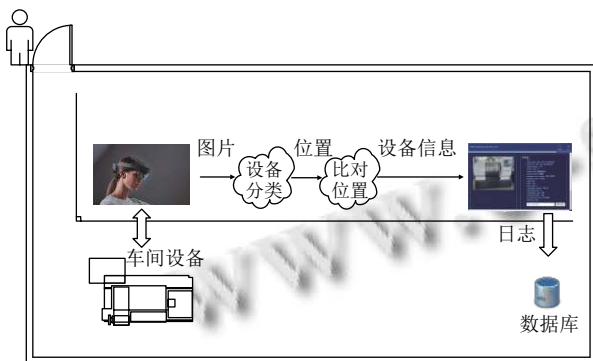


图2 巡检示意图

#### 1.4 混合现实人机交互

HoloLens 设备具有的交互方式主要包括手势识别交互、凝视交互和语音识别交互.

基于手势识别的交互方式: HoloLens 上有两个神经网络来识别手势, 一个神经网络一直处于运行中, 负责判断是否识别到了手势, 当通过初步检验后, 打开第 2 个神经网络来精确识别手势. 通过 HoloLens 朝向外部的摄像头识别到使用者的手部姿势后, 在混合现实画面中呈现人体手部骨骼运动状态的虚拟 UI<sup>[2]</sup>, 通过手部跟踪技术对虚拟对象实现交互操作, 相对于传统交互的方式而言是一种更加自然、直观的交互方式.

基于眼动追踪的凝视交互方式: HoloLens 在鼻梁部位有两个用于眼动追踪的红外相机, 以获得眼球图像数据用于眼动追踪计算, 其支持实时的眼动追踪和注视点渲染, 可以用于与虚拟物体进行交互、操控显示页面等.

基于语音识别的交互方式: HoloLens 前方上端有 3 个小孔负责收集环境音的声音, 而在护目镜的底端的麦克风则收集用户语音, 它正好在用户嘴唇的上方, 这使得在嘈杂的环境中依然可以使用语音操作. 语音识别交互技术主要通过采集用户的语音指令生成控制指

令, 处理后执行最匹配的结果.

3 种交互方式相辅相成, 首先通过注视目标, 然后再通过手势 (隔空敲击)、语音命令 (说出选择命令) 或者停留 (用户凝视着要选择的目標) 来确认选择目标进行交互, 如图 3 所示.

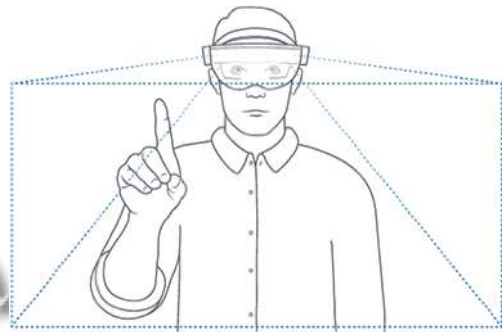


图3 人机交互图

#### 2 设备的分类识别

在车间环境下, 由于 HoloLens 的空间定位能力只在一定的范围内可以识别出设备, 所以加上类别信息进行区分可以在一定程度上提高其识别效果. 传统的浅层特征识别容易受到车间环境的影响, 特征提取能力不足, 因此, 本文利用卷积神经网络 (CNN) 的深层特征提取能力, 弥补了空间定位能力的不足, 同时减少噪声对设备的识别干扰, 提高识别能力.

CNN 是机器学习中常见的模型结构, 在图像领域有着非常广泛的应用并且取得了一定的成就<sup>[11]</sup>. 在图像分类任务上 AlexNet<sup>[12]</sup>、VGGNet<sup>[13]</sup> 和 GoogleNet<sup>[14]</sup> 均已被证明可以取得良好的识别效果. 本文采用的网络模型是微软实验室提出的 ResNet<sup>[15]</sup>, 斩获当年 ImageNet 竞赛中分类任务第一名, 并且将错误率降低到了 3.75%, 其优势在于可以利用更深层次的网络解决训练误差随网络层数的增加而增大的问题<sup>[6]</sup>.

为了解决该问题, 引入了残差结构单元, 其主要功能被称为跳跃连接 (如图 4 所示). 图中  $H(x)$  为理想映射,  $F(x)$  为残差映射,  $H(x) = F(x) + x$ , 把输出变为拟合和输入的叠加, 保护了信息的完整性, 解决了加深网络层次导致的梯度消失或梯度爆炸等问题<sup>[16]</sup>.

本文利用深度学习框架 PyTorch 构建的 50 层 ResNet 网络结构模型, 训练时未使用 dropout 层而是采用了 BN 层, 并将 BN 层放在卷积层和激活层之间, 且卷积层不使用偏置. 该框架分类模型的输入为  $224 \times 224$  大

小的图片,第一个卷积层的参数是 64 个  $7 \times 7$  的卷积核, stride 为 2, padding 为 3, bias 设置为 false; 然后设置 BN 层进行标准化处理, 激活函数使用 ReLU; 最后通过最大池化层下采样处理后, 然后进行后面的 4 个卷积层操作, 其残差结构的个数分别为 3、4、6、3。通过平均池化层下采样处理后, 因需要输出 6 类, 故将最后的全连接层的神经元个数设置为 6, 最后经过 Softmax 层处理输出输入图片被分为某一类的概率。

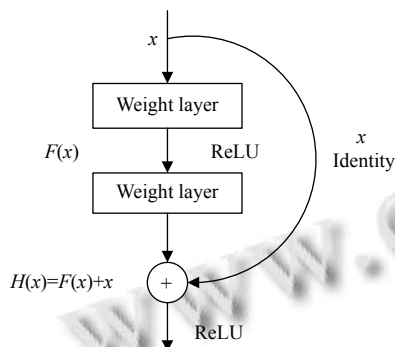


图4 残差结构单元

### 3 实验分析

#### 3.1 样本数据采集

数据来源: 由于这类工业机器人、数控机床和数控系统的图像数据广泛分布于互联网中, 其中百度的图片搜索引擎搜集的数据量较大, 符合实验要求。因此, 数据集的采集主要通过百度文库下载。

数据类别: 被采集的数据一共分为 6 类, 其中待识别的数据准备了 3 类 (工业机器人、数控机床和数控系统), 由于在采集的待识别的数据集中存在噪声特征 (如人类) 及数控系统的轮廓与其它数据 (如书籍和图表) 的轮廓特征有一定的模糊性, 于是在实验过程中加入了另外 3 类噪声数据 (人类、书籍、图表)。

采集过程: 首先利用自定义的爬虫工具根据关键字下载特定的图像数据, 然后再经过人为筛选 (若符合某类特征, 则保留数据, 否则过滤掉数据), 保留符合要求的数据。由于每种类别数据大概在 1 500 张图片左右, 采集到的图像数据经处理后按 7:3 的比例分为测试集和验证集, 最后通过在车间实际采集到的图片数据进行测试。

#### 3.2 实验环境

实验主要在服务器上进行, 操作系统是 CentOS,

深度学习框架采用 PyTorch, 设备的分类识别服务采用 PyTorchServe 部署在服务器上。服务器上的处理器是 NVIDIA GeForce RTX 3090, 24 GB 显存, 对应的 CUDA 版本号是 11.1。Python 版本号是 3.8.8。

#### 3.3 深度学习训练实验

在实验中, 用 loss 值来判断 ResNet 网络模型学习训练的成功性, 用 accuracy 值来评判识别的准确性。采集的训练集总计有 6 960 个样本, 其中待识别数据集 3 类一共 3 259 个样本, 干扰数据集 3 类一共 3 701 个样本。为了优化神经网络, 加快训练的过程, 使用了随机梯度下降 (SGD) 的优化方法, 初始时, 梯度值的学习率 (lr) 设置为 0.001, 动量 (momentum) 为 0.9, 权重衰减参数为 0.005。

实验过程: 在准确率方面, 前 200 次迭代, 训练效果显著, 准确率提升较快; 但 200 到 600 次一直处于波动状态, 第 800 次迭代后, accuracy 逐渐收敛, 稳定在 0.965 附近。训练过程的 accuracy 特征曲线如图 5 所示。实验迭代一共 1 000 次, 在 900 次基本上完成收敛, loss 值趋于 0.064 附近, 此时得到训练模型, loss 值越小, 模型拟合效果越好, 训练过程的 loss 特征曲线如图 6 所示。

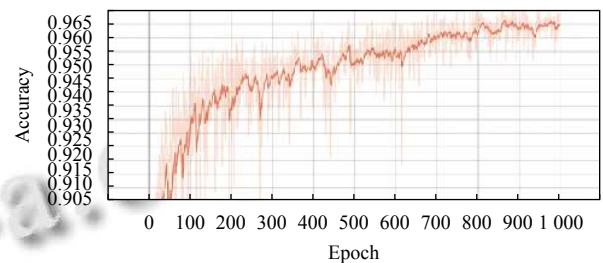


图5 训练过程的 accuracy 特征曲线图

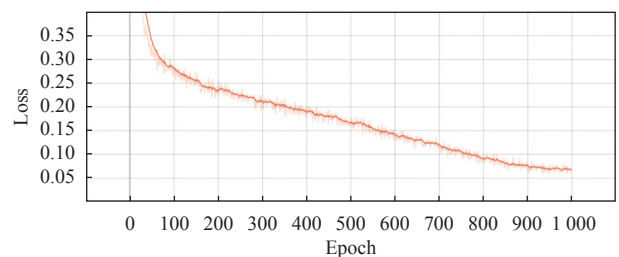


图6 训练过程的 loss 特征曲线

在实验中发现, 工业机器人样本图像中可能会出现车间工人的特征信息, 而数控系统在边缘特征上与图表和书籍 (矩形框) 样本有一定的相似度, 会形成一

定的干扰作用. 为了提高精度, 便将这些带有干扰特征的样本作为新的类别, 并建立对应的数据集, 当某类特征在图像占主要特征时, 则识别为该类别, 如图 7 所示 (class 为识别类别, prob 为识别率).

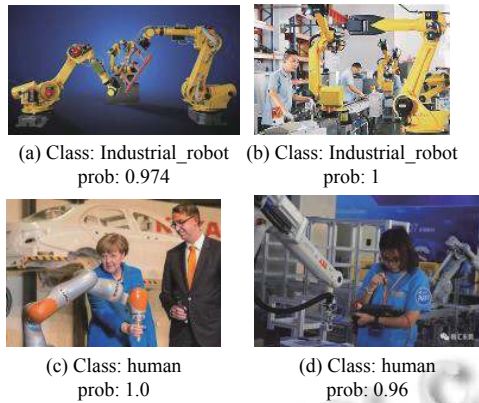


图 7 工业机器人识别

以工业机器人的识别为例, 4 张子图中包括了机器人和人类的特征信息. 若工业机器人的特征在待识别图像中占主导特征, 则识别为工业机器人; 若人类的特征信息在待识别图像中占主导特征, 则识别为人类.

### 3.4 深度学习测试实验

表 1 中, 以 (class, prob (%)) 表示预测结果, 第一列 (pic, class) 表示原图像信息, 其它列为第 epoch 次预测结果.

表 1 测试实验结果

| Epoch      | 225        | 304        | 352        | 505        | 668        |
|------------|------------|------------|------------|------------|------------|
| (pic=1, A) | (A, 82.80) | (D, 44.40) | (A, 89.81) | (A, 88.12) | (A, 82.80) |
| (pic=2, A) | (A, 96.60) | (A, 99.04) | (A, 95.32) | (A, 96.24) | (A, 96.60) |
| (pic=3, B) | (B, 81.88) | (D, 84.50) | (D, 56.57) | (B, 97.38) | (B, 81.88) |
| (pic=4, B) | (B, 99.80) | (B, 99.55) | (B, 99.16) | (B, 99.42) | (B, 99.80) |
| (pic=5, C) | (C, 98.12) | (C, 99.59) | (C, 99.97) | (C, 99.66) | (C, 98.12) |
| (pic=6, C) | (C, 89.32) | (C, 91.27) | (C, 99.37) | (C, 94.28) | (C, 89.32) |

当 epoch=304 时, 其中 pic=1 的工业机器人和 pic=3 的数控系统的图像都被识别为了书籍; 但在 epoch=352 后, pic=1 的工业机器人识别正常, 准确率逐渐上升; 而 pic=3 的数控系统在 epoch=505 后才开始识别成功, 最后达到 98.60%.

测试实验结果表明: 将噪声单独提取成一类, 在一定程度上提高了 ResNet50 的识别率, 可以区分边缘特征相似的 B 与 D (表 1 中: pic=1 代表图 8 中的第一个子图, 其余子图以此类推; epoch 值从 0 开始, 代表第一轮; 类别 A 代表 Industrial\_robot, 即工业机器人, 类别

B 代表 Numerical\_control\_system, 即数控系统, 类别 C 代表 Numerical\_machine\_tool, 即数控机床; 类别 D 代表干扰噪声 books).



图 8 测试实验图

### 3.5 混合现实巡检应用

巡检人员在佩戴 HoloLens 后, 在车间入口处启动设备, 并打开巡检应用程序. 当巡检到某机床设备时, 点击拍照, 采集到的图像会显示在虚拟 UI 上; 再点击预测, 图像被上传至服务器; 最后点击显示, 若通过服务器验证, 则显示该机床的基本信息、运行状态和报警等运维详细信息, 完成日常巡检任务. 巡检过程中, 还可以在下面的输入框中通过基于语音识别的交互方式录入日志信息, 巡检效果图如图 9 所示.



图 9 混合现实巡检效果图

## 4 结论与展望

相对于传统的车间人工巡检, 本文将深度学习应用到混合现实车间巡检中提高了巡检人员的工作效率, 利用了卷积神经网络中识别率较高的 ResNet 在一定程度上弥补了 HoloLens 空间定位功能的不足之处, 通过类别和定位信息的方式识别设备, 提高了 HoloLens

的利用率和识别率。实验结果表明, ResNet 在实际应用环境中识别率较高, 有效的过滤了噪声。但是 ResNet 所需的训练时间较长, 对于车间设备图像中自带的噪声特征有一定的模糊性, 接下来可以进行目标检测技术的应用实验, 同时进行多个车间设备的巡检, 进一步提高 HoloLens 的利用率和识别率。

### 参考文献

- 1 王帅, 郭锐锋, 董志勇, 等. 混合现实装配检测中深度学习的数据增强方法. 计算机集成制造系统, 2021, 27(3): 716-727.
- 2 刘敏洋, 李博功. 基于混合现实技术的工业智能运维系统设计. 制造技术与机床, 2020, (9): 80-82, 87.
- 3 孙臻, 齐晓晨, 孙海燕. 混合现实设备助力图书馆设备运维——微软 HoloLens 在图书馆中的应用实践. 大学图书馆情报学刊, 2021, 39(3): 120-124. [doi: 10.3969/j.issn.1006-1525.2021.03.023]
- 4 邱俊豪, 胡常伟, 彭磊, 等. 面向大型生产车间的设备点检与巡检管理系统设计. 机电工程技术, 2021, 50(3): 35-39. [doi: 10.3969/j.issn.1009-9492.2021.03.006]
- 5 李玄基. 基于 HoloLens 增强现实的关键技术研究与应用 [硕士学位论文]. 成都: 西南交通大学, 2019.
- 6 郑远攀, 李广阳, 李晔. 深度学习在图像识别中的应用研究综述. 计算机工程与应用, 2019, 55(12): 20-36. [doi: 10.3778/j.issn.1002-8331.1903-0031]
- 7 辛宁, 丁新宇, 黄可南, 等. 混合现实技术在医学中的应用. 中国胸心血管外科临床杂志, 2021, 28(5): 597-602.
- 8 沈毅, 吴罡, 周睿. 可穿戴计算设备在工业智能运维操作的辅助运用. 系统仿真学报, 2021, 33(5): 1188-1195.
- 9 潘娜, 张记炜. 混合现实技术在军事训练中的应用优势——以美军综合训练环境为例. 国防科技, 2021, 42(1): 111-117.
- 10 马双. 混合现实技术在医学教育中的应用. 中国教育技术装备, 2020, (9): 101-102, 107.
- 11 张驰, 郭媛, 黎明. 神经网络模型发展及应用综述. 计算机工程与应用, 2021, 57(11): 57-69. [doi: 10.3778/j.issn.1002-8331.2102-0256]
- 12 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Communications of the ACM, 2017, 60(6): 84-90. [doi: 10.1145/3065386]
- 13 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv: 1409.1556, 2014.
- 14 Szegedy C, Liu W, Jia YQ, et al. Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston: IEEE, 2015. 1-9.
- 15 He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016. 770-778.
- 16 王景中, 杨源, 何云华. 基于多分类和 ResNet 的不良图片识别框架. 计算机系统应用, 2018, 27(9): 100-106. [doi: 10.15888/j.cnki.csa.006517]