

基于YOLOv2的微创手术机器人肺结核病灶图像识别研究

朱尔汉, 郭语, 范哲铭, 张新扬, 陆新宇, 徐帅, 王锋锋

金陵科技学院机电工程学院, 江苏 南京

收稿日期: 2023年11月16日; 录用日期: 2024年2月18日; 发布日期: 2024年2月23日

摘要

肺结核是全球公共卫生问题, 每年数百万人受到其影响, 其中包括成千上万的死亡病例。尽管存在有效的治疗方法, 但及时诊断和早期干预仍然是防止病灶扩大的关键, 其中外科手术治疗成为了一种有效手段。本文提出了一种基于YOLOv2算法的微创手术机器人肺结核病灶识别技术。该技术结合了微创手术机器人、内窥镜、K210视觉识别模块以及YOLOv2识别算法, 以提高肺结核病灶实时监测的识别准确度和处理效率。本文构建了训练数据集, 使用YOLOv2算法进行深度网络训练, 并评估了性能指标, 包括准确率、召回率、精确度和F1分数。实验结果表明, 该技术具有高准确度和高效率, 有望改进肺结核的检测与治疗手段。

关键词

肺结核, YOLOv2算法, K210, 图像识别

Research on Image Recognition of Tuberculosis Lesions by Minimally Invasive Surgical Robot Based on YOLOv2

Erhan Zhu, Yu Guo, Zheming Fan, Xinyang Zhang, Xinyu Lu, Shuai Xu, Fengfeng Wang

College of Mechanical and Electrical Engineering, Jinling Institute of Technology, Nanjing Jiangsu

Received: Nov. 16th, 2023; accepted: Feb. 18th, 2024; published: Feb. 23rd, 2024

Abstract

Tuberculosis has been a major challenge in the global public health sector, affecting millions of

文章引用: 朱尔汉, 郭语, 范哲铭, 张新扬, 陆新宇, 徐帅, 王锋锋. 基于 YOLOv2 的微创手术机器人肺结核病灶图像识别研究[J]. 人工智能与机器人研究, 2024, 13(1): 49-55. DOI: 10.12677/airr.2024.131006

people annually, including tens of thousands of fatalities. Despite the existence of effective treatment methods, timely diagnosis and early intervention remain crucial in preventing the spread of lesions. Surgical intervention has emerged as an effective means of treatment. This paper introduces a minimally invasive surgical robot-based tuberculosis lesion recognition technology, which is built upon the YOLOv2 algorithm. This technology combines robot system design, endoscopy, the K210 visual recognition module, and the YOLOv2 recognition algorithm to enhance the identification accuracy and the handling efficiency for real-time recognition. The paper constructed a diverse training dataset, conducted deep network training using the YOLOv2 algorithm, and evaluated performance metrics, including accuracy, recall, precision, and F1 score. Experimental results demonstrate that this technology offers high accuracy and efficiency, with the potential to enhance tuberculosis detection and treatment.

Keywords

Tuberculosis, YOLOv2, K210, Image Recognition

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

肺结核一直是全球公共卫生领域的一大挑战。根据世界卫生组织的统计数据，每年有数百万人受到结核病的影响。近年来，耐多药肺结核的病例有增加趋势[1]。随着肺结核病耐药性的增强，通过外科手术治疗，以最小代价切除病灶、最大限度保存肺功能，是治愈肺结核、消灭传染源的有效手段[2]。

如今，具有视觉识别功能的微创手术机器人技术已经在医疗领域中取得了显著的进展。刘龔龙等[3]指出中国微创手术机器人在近十年发展迅速，专利和论文均呈现增长趋势。付宜利等[4]发现微创外科手术机器人智能化控制技术是其未来发展的趋势。苏赫[5]提出了一种借助内窥镜和医学影像的微创手术机器人图像系统设计方法与空间映射策略，但针对该领域的图像识别技术仍存在大量空白。Bogyu Park 等[6]设计了用于手术机器人的视觉识别模型，可以提供更高的手术精度、更小的创伤和更快的康复速度。然而在实际应用中，该算法在病灶的准确定位和快速识别方面仍存在部分问题，尤其是针对肺结核病灶，其检测精度无法达到手术要求。LAKHANI 等[7]借助深度学习与卷积神经网络对肺结核患者的胸部 X 射线图像进行准确识别。单连峰等[8]将三种不同的卷积神经网络(CTCNN、PET-CNN、PET/CT-CNN)应用在肺结核的识别诊断中，对 CT 图像进行识别。上述研究均不具有实时检测功能，不适用于手术过程中的病灶识别。

本文旨在探索将 YOLOv2 算法引入视觉识别机器人系统，以提高肺结核病灶的识别精度和效率的方法。本文所采用的 YOLOv2 算法在 VOC2007 数据集上，保持检测速度为 40 帧/s 的情况下 mAP 能够达到 78.6% [9]，能够在保证精度的同时优化检测速率。作为一种基于回归分析的单阶段检测算法，YOLOv2 算法以其实时性、多类别支持和边界框回归的特点而著称[10]。通过将这一算法与微创手术机器人系统整合，有望为用户提供更准确、更快速的病灶识别功能，从而改进肺结核的治疗手段和方式。

2. 机器人系统结构与工作原理

2.1. 总体结构与工作原理

根据机器人的设计需求，我们将结构分成了三大模块，包括：机械臂操作模块、图像识别模块和单

片机模块, 见图 1。

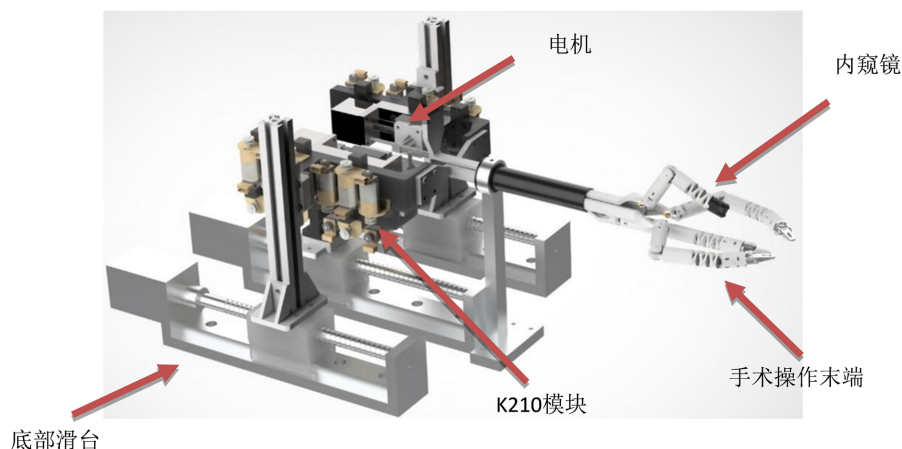


Figure 1. Minimally invasive surgical robot
图 1. 微创手术机器人

机械臂操作模块由操作末端部分, 电机模块部分以及底部滑台部分组成。主要负责将内窥镜伸入病人体内拍摄、协助医生进行切除操作等。内窥镜很适合在人体内部进行非侵入性的观察、检查和治疗。它由一个细长、灵活的数字传感器组成, 可以通过自然腔道(如口腔、喉咙、食道、胃、肠道等)或通过小切口(例如肚脐切口)插入体内。内窥镜可以搭载摄像头、光源和其他传感器, 将内部图像传送到显示屏上, 以供使用者实时观察。

STM32 单片机模块提供了强大的处理能力和计算资源, 能够有效处理机器人控制和通信任务。结合 K210 图像识别模块, STM32 单片机可以协调并执行机械臂的运动。STM32 单片机能够在低功耗模式下运行, 特别是在病灶识别这种需要长时间运行的应用中, 低功耗设计可以减少功耗需求, 从而提高设备的可用性和可靠性, 同时不会影响图像传输的速度。

当图像识别任务开始时, 机械臂会驱动电机和底部滑台将携带有内窥镜的操作末端送至正确位置, 内窥镜将拍摄到的数据通过信息传输模块实时传送给 K210 模块。K210 模块在完成运算后将数据发送给上位机, 由上位机将数据传输至电脑, 电脑在完成信息整合和加工后将结果输出在屏幕上, 使用者根据输出的结果控制单片机操作机械臂, 见图 2。

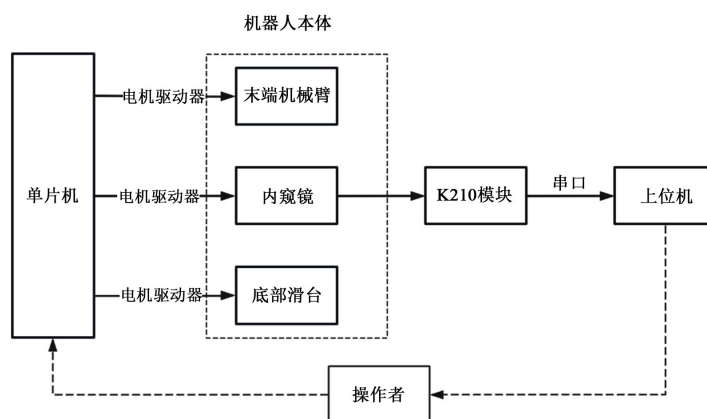


Figure 2. Structure diagram of minimally invasive surgical robot
图 2. 微创手术机器人结构图

2.2. K210 主控模块设计

K210 包含 RISC-V 双核 CPU 的低功耗高性能的芯片，其核心功能是机器视觉与听觉，包含用于计算卷积人工神经网络的神经网络处理器(KPU)，其中内置了卷积、批归一化、激活、池化运算单元等。通过将 K210 模块与内窥镜系统整合，可以实现对肺结核病灶的高效处理和识别。

K210 对于大多数机器学习算法拥有很强的图像处理能力。它可以与 YOLOv2 算法结合，在嵌入式环境中处理复杂的图像处理任务，在保证低功耗的同时实现病灶检测。另外，K210 模块可以轻松集成到嵌入式环境系统中，通过 USB 与上位机进行数据传输和通信。这有助于实现病灶图像的实时分析和传输。

3. YOLO 算法原理与病灶识别模型设计

3.1. YOLO 算法原理

YOLOv2 算法之所以能够在检测物体算法中脱颖而出，是因为它提供了准确的结果、标签和 IoU 值。该算法在处理小物体方面表现更好[11]，并且比以前可用的版本响应速度更快。单阶段架构只需要一个网络来预测边界框和类别概率，这种设计使 YOLOv2 的运算更加高效和快速。为了识别人和物体，YOLO 只是通过神经网络进行一次前向传播。算法在单次运行中对物体进行预测。根据目标确认模块的实时性要求，算法会基于 YOLOv2 进行深度网络训练[12]、构造训练数据集[13]，并对该训练集进行聚类选择锚框。对于网络结构，该算法将底层深度特征的细节纹理信息与高层深度特征的语义信息相结合，使目标检测更加准确。

YOLOv2 相较于其前代算法和其他算法有许多改善。该算法提出了 Darknet-19 的网络结构。该结构删除了传统网络的最后一个 FC 层，代之以一个 1×1 的卷积，进一步使用全局平均池化来减少参数，从而明显改善运算速度并防止过拟合。用户可以任意指定锚框的数量，在 YOLOv2 中将中间特征图和最终特征图结合使用，将前卷积层的高分辨率特征图添加到后卷积层的低分辨率特征图中。高分辨率特征图隐含了小物体上的数据，提高了小物体检测性能。

YOLOv2 在最后一个卷积层输出后，在卷积或池化之后，激活函数之前，会对每个数据输出进行标准化，进行批量标准化 BN。数据集中的每一张图片会被分成 13×13 的网格。YOLOv2 预测每个网格 5 个锚框[14]，通过这 5 个锚框(anchor box)来预测 5 个先验框(bounding box)。不同于 YOLOv1 直接将 t_o 作为边框置信度，YOLOv2 会对其进行激活函数(Sigmoid)变换，得到 $\sigma(t_o)$ 作为其边框置信度。

$$b_x = \sigma(t_x) + c_x \quad (1)$$

$$b_y = \sigma(t_y) + c_y \quad (2)$$

$$b_w = p_w e^{t_w} \quad (3)$$

$$b_h = p_h e^{t_h} \quad (4)$$

$$\Pr(\text{object}) * IoU(b, \text{object}) = \sigma(t_o) \quad (5)$$

上述公式中 c_x 与 c_y 是所得图像的左上角的横竖坐标。 p_w 与 p_h 表示先验框的大小。由于先验框与预测框位置不同，为了提高精度，我们还需要计算两者的坐标误差，但是在训练前期，使预测框学习到的先验框的形状和位置只会是在迭代的第 1 到第 12800 次之间进行计算。

3.2. 病灶识别模型与程序设计

利用 YOLO 算法，我们对病灶模型进行了训练。首先，我们收集并整理了大量肺结核病灶图像数据集[15]，其中包含各种不同角度、光照条件和大小的图像。YOLOv2 支持多分辨率输入，通过适当设置输

入图像的分辨率来计算负载和检测精度[16]。然后使用准备好的数据集，指定模型的参数、数据集路径、类别信息，运行训练脚本来训练 YOLOv2 模型。一旦模型训练完成，即可使用它进行推理，即在新的图像上运行检测。将希望检测的图像输入模型，模型会返回检测到的物体的边界框和类别信息。根据病灶识别模型，我们可以通过编译代码来筛选出病灶并展示检测结果。

部分代码设计：

#准备图像

dataset = LoadImages (data_path) # data_path 进行检测的图像文件夹路径

for img_path, img, im0s, _ in dataset:

img = torch.from_numpy (img).to (model.device)

#进行推理

pred = model (img) [0] # 进行推理

pred = non_max_suppression (pred, conf_thres = 0.5, iou_thres = 0.4) [0] #过滤结果

#处理检测结果

if pred is not None and len (pred):

det = pred [0] # 获取第一个检测结果

det[:, :4] = scale_coords (img.shape [2:], det[:, :4], im0s.shape).round () #重新缩放边界框

for * xyxy, conf, cls_conf, cls in det:

label = class_names [int (cls)]

#数据集中的类别名

#在图像上绘制边界框

c1, c2 = (int (xyxy [0]), int (xyxy [1])), (int (xyxy [2]), int (xyxy [3]))

4. 性能评估与结果分析

4.1. 训练数据集

本文已经指出了数据集在 YOLOv2 算法模型中的作用，而在实际训练中，需要构建包含不同尺寸、形状和位置的肺结核病灶图像的训练数据集，并对数据集里的每张图像进行详细标注，其中包括病灶的坐标、类别信息以及其他必要的标识。这一过程包含了锚框的选择，以确保模型能够准确地预测病灶的位置和尺寸。利用构建好的训练数据集，我们对 YOLOv2 算法进行了深度网络训练。模型通过学习训练数据集中的图像和标签逐渐提高了对肺结核病灶的识别能力。数据集的质量和多样性对模型的训练效果起到了至关重要的作用。

4.2. 性能评估指标

为了评估数据集和模型的质量，衡量病灶识别的精度，本文将用作性能评估的图片分成了不同数量的十组，手动统计了图中的病灶数量，并交由训练完成的模型进行识别。我们对检测结果采用了以下数据用于评估性能：

准确率(Accuracy)与召回率(Recall) [17]: 准确率表示模型正确分类的样本比例，召回率即真正例与所有实际正例的比例。这两项数据可以用来衡量模型正确识别病灶的能力。在病灶图像处理中，召回率通常更重要，因为它关注了病灶是否被漏诊。

F1 分数(F1 Score): F1 分数是精确度和召回率的调和平均，可以综合评估模型性能。

我们将性能评估指标按照每组的实际病灶数量进行排列后生成了如图 3 所示的性能指标与病灶数量

关系。

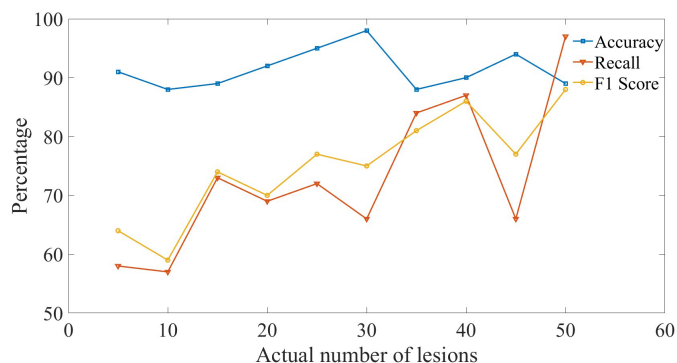


Figure 3. Relationship between performance indicators and number of lesions
图 3. 性能指标与病灶数量关系

从图中可以看出, 识别的准确率平均值能够超过 90%, F1 分数达到 75%。由于先前的研究在肺结核检测中的平均准确率低于 80%, 我们认为该研究对肺结核检测精度的提升效果显著。另外, 随着数据训练量的增加, 检测精度能够呈上升趋势。同时 K210 模块的响应时长基本达到 ms 级别, 满足实时性要求。相较于在该领域先前的一些研究, YOLO 算法的引入不仅使检测精度得到明显改善, 还对检测速度进行了提升。面对实际医疗环境中复杂多变的情况, 本文所展现出的及时性和准确性使已有的手术机器人视觉识别模型更加切合实际使用需求。部分测试结果如图 4 所示。

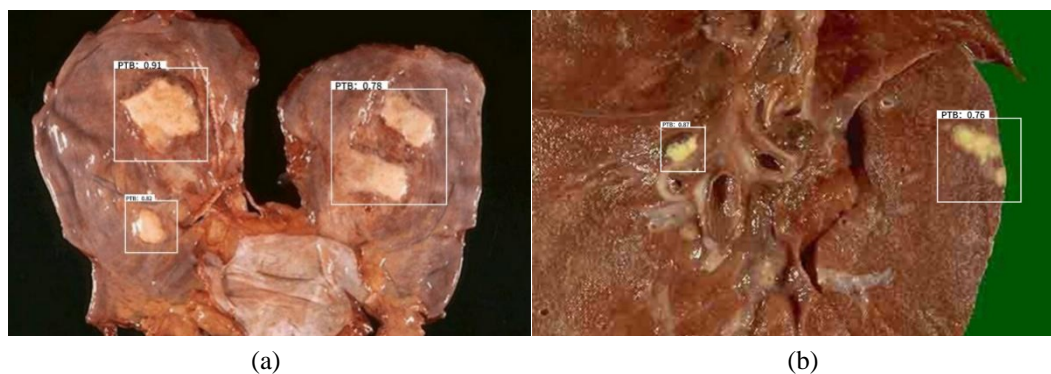


Figure 4. Test results
图 4. 测试结果

5. 结语

本文介绍了一种基于 YOLOv2 算法的手术机器人肺结核病灶识别技术, 旨在提高肺结核病灶实时检测的识别精度和效率。通过将 YOLOv2 算法与视觉识别机器人系统整合, 我们实现了更准确、更快速的病灶识别, 为改进肺结核的治疗手段和方式提供了新的途径。本文还详细介绍了机器人系统的方案设计、内窥镜的应用、K210 模块的功能以及 YOLOv2 算法的原理和病灶识别模型的设计。通过性能评估, 我们验证了该技术的有效性和实用性。

参考文献

- [1] 窦学军, 龚昌帆, 闫东杰, 等. 手术联合化疗治疗耐多药肺结核结果分析[J]. 中国防痨杂志, 2007, 29(4):

336-338.

- [2] 周胤江, 徐晓华. 肺结核手术治疗 56 例临床分析[J]. 内蒙古医学杂志, 2006, 38(4): 359.
- [3] 刘龔龙, 刘进长, 唐莉, 等. 基于论文和专利的微创手术机器人领域研发现状及前沿热点分析[J]. 科学技术与工程, 2022, 22(1): 250-258.
- [4] 付宜利, 潘博. 微创外科手术机器人技术研究进展[J]. 哈尔滨工业大学学报, 2019, 51(1): 1-15.
- [5] 苏赫. 微创手术机器人图像系统设计方法与空间映射策略研究[D]: [硕士学位论文]. 天津: 天津大学, 2017.
- [6] Park, B., Chi, H., Park, B., *et al.* (2023) Visual Modalities-Based Multimodal Fusion for Surgical Phase Recognition. *Computers in Biology and Medicine*, **166**, Article ID: 107453. <https://doi.org/10.1016/j.combiomed.2023.107453>
- [7] Lakhani, P. and Sundaram, B. (2017) Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. *Radiology*, **284**, 574-582. <https://doi.org/10.1148/radiol.2017162326>
- [8] 单连峰, 郭芳. 基于深度学习卷积神经网络的肺结核 CT 辅助诊断模型在临床中的应用价值[J]. 中国新通信, 2022, 24(24): 126-128.
- [9] 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.
- [10] 许德刚, 王露, 李凡. 深度学习的典型目标检测算法研究综述[J]. 计算机工程与应用, 2021, 57(8): 10-25.
- [11] Zhou, H., Xiao, Y., Zheng, Z., *et al.* (2022) YOLOv2-Tiny Target Detection System Based on FPGA Platform. 2022 3rd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE). Xi'an, 15-17 July 2022, 289-292. <https://doi.org/10.1109/ICBAIE56435.2022.9985817>
- [12] Ioffe, S. and Szegedy, C. (2015) Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *International Conference on Machine Learning*. Lile, 6-11 July 2015, 448-456.
- [13] Arima, K., Nagata, F., Shimizu, T., *et al.* (2023) Improvements of Detection Accuracy and Its Confidence of Defective Areas by YOLOv2 Using a Data Set Augmentation Method. *Artificial Life and Robotics*, **28**, 625-631. <https://doi.org/10.1007/s10015-023-00885-9>
- [14] 董洪义. 深度学习之 PyTorch 物体检测实战[M]. 北京: 机械工业出版社, 2020: 171.
- [15] 闫明艳, 陈根铭, 赖超, 等. 人工智能对肺结核患者病变检出及定性诊断价值研究[J]. 新发传染病电子杂志, 2018, 3(4): 214.
- [16] 栗俊杰, 毛彭军, 方骞, 等. 基于 K210 的火点检测和喷头定向控制[J]. 科学技术与工程, 2021, 21(35): 15136-15143.
- [17] 刘策, 刘小峰. 助老服务机器人多模式友好交互系统设计及实现[J]. 电子测量与仪器学报, 2019(1): 177-182.