

摘要

基于 YOLOv5 改进的肺结节检测算法

电子计算机断层扫描影像为灰度图,使得影像中肺结节的表征信息有限,同时肺结节的形态多样,以不规则的类圆形为主,且大小不一,不利于检测网络的特征提取,从而导致肺结节检测网络模型出现误检率高、定位不精确等问题。针对此问题,本文利用 YOLOv5 目标检测网络及 Vision Transformer 模型,提出了 YOLO-Transformer (YOLO-TF) 目标检测网络,且在 YOLO-TF 网络的基础上,进一步精简参数设计了 YOLO-TransformerLite (YOLO-TFLite) 网络。

当前,在目标检测领域中,存在各种不同类型的网络结构。在进行网络创新优化前,为了挑选出适用于肺结节检测任务的网络,本文进行了基础算法的选型实验。通过对比目前双阶段检测网络及单阶段检测网络中性能最优的检测网络,对各网络的检测精度及检测速度进行综合对比分析,结合实际应用需求及各网络的测试数据,最终选择 YOLOv5 网络作为基线网络。

为提高网络的检测精度,本文在 YOLOv5 网络中加入 Vision Transformer 模型,优化设计了 YOLO-TF 网络,提出了级联并行 Transformer 模块,用于对全局特征的提取,以及 Deform 卷积模块,用于对细节局部特征的提取。同时设计了双维度编解码特征提取网络,相比于原特征提取网络具有更丰富的特征信息,采用特征叠加层,将局部特征与全局特征进行合并,得到多维度特征信息的特征图;此外,进一步的设计了特征重组网络,以对特征提取网络的输出进行二次特征处理,通过对不同网络层次的特征图重组、再分组的操作,均衡了特征图的特征信息量。因此,YOLO-TF 网络相比于 YOLOv5,能够提取更加丰富的特征信息,并且特征重组网络可对特征进行二次提权,提升有效特征的权值,去除冗余的背景特征、综合各网络层次的特征信息,提升各尺度检测器的检测精度。

为优化网络的参数设计,降低参数复杂度和计算复杂度,本文通过采用轻量化卷积结构、重参数化设计对 YOLOv5 进行轻量化设计。提出了 YOLO-TFLite 网络,进一步精简模块参数,设计了 DGhost 轻量化卷积模块,并对 DGhost 卷积模块进行重参数化;本文通过修改 YOLO-TF 特征重组网络的结构,在其聚合多维度特征与重组操作后,仅输出单尺度特征供检测网络使用;对于检测网络的轻量化采用 Anchor Free 机制,生成圆形检测框。在该环节中,取消了非极大值抑制(Non

Maximum Suppression, NMS) 计算, 且通过中心点判断目标的区域也可降低检测框数量的冗余。YOLO-TFLite 网络相比于 YOLOv5 及 YOLO-TF 网络, 主要采用了超轻量化的网络模块进行搭建, 并仅使用单个 Anchor Free 机制的检测器进行目标检测任务, 上述优化操作均使得 YOLO-TFLite 的网络参数量远低于 YOLOv5 及 YOLO-TF。

在对 YOLO-TF 及 YOLO-TFLite 的性能评估实验中, 本文采用了 LUNA16 数据集, 共包含有肺部 CT 影像 888 张, 按照 8:2 的比例, 将所有图片随机分为训练集与测试集。文中进行了消融实验和对比实验, 实验结果表明: YOLO-TF 的平均检测精度均值达 0.918, 相比于 YOLOv5 提升了 0.056; YOLO-TFLite 的网络参数量为 6.6MB, 仅为 YOLOv5 的七分之一, 并且 YOLO-TFLite 的平均检测精度均值相比于 YOLOv5 仅下降了 0.011, 仍保持有较优的检测精度。

为了能更好的应用所设计的模型, 文中依据 YOLO-TF 及 YOLO-TFLite 模型设计了肺结节检测系统, 通过可视化交互界面提升两个模型的易用性, 并增加了检测报告的生成、导出、模型更新等功能。

关键词:

肺结节, 深度学习, 目标检测, Transformer, 轻量化, YOLOv5

Abstract

Improved lung nodule detection algorithm based on YOLOV5

The electron computed tomography images are grayscale images, which make the characterization information of pulmonary nodules in the images limited, and the morphology of pulmonary nodules is diverse, mainly irregular and round-like, and the size varies, which is not conducive to the feature extraction of the detection network, thus leading to the problems of high false detection rate and imprecise localization in the lung nodule detection network model. To address this problem, this paper proposes the YOLO-Transformer (YOLO-TF) target detection network using the YOLOv5 target detection network and Vision Transformer model, and further streamlines the parameters on the basis of the YOLO-TF network to design the YOLO-TransformerLite (YOLO- TFLite) network.

Currently, various types of network structures exist in the field of target detection. Before carrying out network innovation optimization, in order to select the network suitable for the lung nodule detection task, the selection experiments of the underlying algorithm are conducted in this paper. By comparing the best performing detection networks among the current two-stage detection networks and single-stage detection networks, a comprehensive comparison and analysis of the detection accuracy and detection speed of each network was performed, and the YOLOv5 network was finally selected as the baseline network by combining the actual application requirements and the test data of each network.

To improve the detection accuracy of the network, this paper optimizes the design of YOLO-TF network by adding the Vision Transformer model to the YOLOv5 network, and proposes a cascaded parallel Transformer module for the extraction of global features and a Deform convolution module for the extraction of detailed local features. In addition, a dual-dimensional codec feature extraction network is designed, which has richer feature information than the original feature extraction network, and a feature overlay layer is used to merge local features with global features to obtain a feature map with multidimensional feature information. In addition, the feature

reorganization network is further designed to perform secondary feature processing on the output of the feature extraction network, and the amount of feature information in the feature map is balanced by reorganizing and regrouping the feature maps of different network levels. Therefore, the YOLO-TF network can extract richer feature information compared with YOLOv5, and the feature reorganization network can improve the detection accuracy of detectors at each scale by second weighting the features, increasing the weights of effective features, removing redundant background features, and synthesizing the feature information at each network level.

In order to optimize the parameter design of the network and reduce the parameter complexity and computational complexity, this paper lightens YOLOv5 by adopting a lightweight convolutional structure and reparameterization design. The YOLO-TFLite network is proposed to further streamline the module parameters, and the DGhost lightweight convolution module is designed, and the DGhost convolution module is reparameterized; this paper modifies the structure of the YOLO-TF feature reorganization network by outputting only single-scale features for the detection network after its aggregation of multidimensional features and reorganization operations; for the lightweighting of the detection network using the Anchor Free mechanism is used to generate circular detection frames. In this link, Non Maximum Suppression (NMS) calculation is eliminated, and the redundancy of the number of detection frames can also be reduced by judging the region of the target through the center point. Compared with YOLOv5 and YOLO-TF networks, the YOLO-TFLite network is mainly built with ultra-lightweight network modules and uses only a single Anchor Free mechanism detector for the target detection task, all these optimizations make the number of network parameters of YOLO-TFLite much lower than YOLOv5 and YOLO-TF.

In the performance evaluation experiments of YOLO-TF and YOLO-TFLite, the LUNA16 dataset with 888 lung CT images was used in this paper, and all images were randomly divided into training and test sets according to the ratio of 8:2. Ablation experiments and comparison experiments were conducted in the paper, and the experimental results showed that the mean detection accuracy of YOLO-TF

reached 0.918, which was improved by 0.056 compared with YOLOv5; the number of network parameters of YOLO-TFLite was 6.6 MB, which was only one-seventh of YOLOv5. The average detection accuracy of YOLO-TFLite is only 0.011 lower than that of YOLOv5, which still maintains a better detection accuracy.

In order to better apply the designed models, a lung nodule detection system based on the YOLO-TF and YOLO-TFLite models is designed in this paper, which enhances the ease of use of the two models through a visual interactive interface and adds the functions of test report generation, export, and model update.

Keywords:

Lung Nodules, Deep Learning, Object Detection, Transformer, Lightweighting, YOLOv5

目 录

摘 要	I
Abstract	III
第一章 绪论	1
1.1 研究背景及意义	1
1.2 国内外研究现状	2
1.3 本文主要研究内容	4
1.4 论文结构	5
第二章 相关技术背景	7
2.1 卷积神经网络	7
2.2 激活函数	11
2.3 Transformer 简介	12
2.4 常见的目标检测算法	14
2.4.1 R-CNN	15
2.4.2 Faster R-CNN	16
2.4.3 YOLO	17
2.5 本章小结	17
第三章 基于深度学习的肺结节检测算法选型实验	18
3.1 引言	18
3.2 数据集及实验环境说明	18
3.3 模型性能评估指标说明	19
3.4 算法选型实验及对比分析	21

3.5 本章小结	23
第四章 高精度 YOLO-TF 检测网络的设计实现	24
4.1 引言	24
4.2 双维度编解码特征提取网络设计	24
4.2.1 Vision Transformer 相关理论	24
4.2.2 级联并行 Transformer 模块设计	27
4.2.3 DeformCSP 卷积模块设计	30
4.2.4 双维度编解码特征提取网络	33
4.3 特征重组网络设计	34
4.4 YOLO-TF 网络结构设计	37
4.5 实验及数据分析	39
4.5.1 消融实验	39
4.5.2 对比实验	45
4.6 本章小结	47
第五章 轻量化 YOLO-TFLite 检测网络的设计实现	48
5.1 引言	48
5.2 轻量化编解码特征提取网络设计	48
5.2.1 Ghost 卷积相关理论	48
5.2.2 Rep-DGhost 模块设计	50
5.2.3 Rep-DGhostNet 编解码网络设计	53
5.3 轻量化瓶颈网络设计	55
5.4 基于 Anchor Free 机制的检测器设计	56

5.5 实验及数据分析	59
5.5.1 消融实验	59
5.5.2 对比实验	65
5.6 本章小结	67
第六章 肺结节检测系统的设计与实现	69
6.1 引言	69
6.2 系统功能需求分析设计	69
6.3 系统结构设计	70
6.4 模块工作流程设计	71
6.5 肺结节检测系统实操流程	73
6.6 本章小结	75
第七章 总结与展望	76
7.1 总结	76
7.2 展望	77
参考文献	79
致 谢	86
作者简介及在学期间所取得的科研成果	87

第一章 绪论

1.1 研究背景及意义

在全球，肺癌是与癌症有关的最重要死亡原因，肺还是人体中最大的脏器，更是经常与外界接触的脏器，和人的呼吸息息相关。2022 年全球癌症数据统计^[1]显示，肺癌在美国癌症死亡中占有重要地位，每天有 350 万人因为肺癌死亡。中国癌症新增病例数是美国的两倍，但是死亡病例数却是美国的 5 倍之多。在中国最常见的癌症就是肺癌，我国每年新发肺癌病例 78.7 万，因为肺癌死亡的人数超过 63 万，女性肺癌的病例也在逐年升高。受新冠肺炎影响，很多人通过接受了 CT 扫描来排除感染。肺癌是由肺结节发展而来的，因此，对于肺结节的筛查，也变成了一项判断肺癌的关键检查项。由于肺部器官与外界相通的属性，常见的肺部疾病除了肺癌，还有肺结核，气胸，肺部阴影，肺脓肿等。

对于肺部结节检查，通常采用计算机断层扫描，也就是常说的 CT 技术进行筛查，CT 技术就是用 X 射线束扫描人体器官某一部分的某一厚度层面。由于人体的组织厚度不同，造成对 X 射线束的吸收和透过率有差别，因此可以采用灵敏度极高的仪器对人体进行测量，从而获得器官被检查部位的断面或者立体影像，也就是 CT 影像，来发现器官检查部分是否病变。当身体检查部分发生病变后，密度与正常组织的密度不相同，可以利用 CT 影像来发现病变情况。除 CT 成像技术外，常用的医学成像技术还包括正电子发生电脑断层扫描技术 PET，核磁共振成像 MRI。

然而，CT 成像也含有一定的缺陷，自从使用 CT 成像技术后，临床上虽然从 CT 影像来进行病变部位的识别与判断提高了效率，但是对于医生来说也是一项具有挑战性和耗时耗力的任务。现在人工智能和深度学习^[2]的兴起，让自动化进行肺部结节检测成为了一项具有探索价值的任务。早期的肺结节筛查需要人工标注大量的标记来进行判断，需要耗费很长的时间，且检测的时间由于图片的增多也大大增加，随着深度学习和人工智能的发展，人们直接根据 CT 影像自动学习肺结节的相关特征，从而进行自动化结节检测这项任务具有了非常大的可能性。

近年来，计算机视觉领域有了飞速发展，在医学图像目标检测方面，YOLOv5 算法能够对于肺结节的识别提供较大的帮助。YOLOv5 在 YOLO 目标检测系列算法上进行了改进，它属于单阶段目标检测算法，文中以 YOLOv5 算法为基础，

实现了模型轻量化及检测精度增强, 鉴于肺结节体积小, 样本受限, 分布差异较大等特征, 对算法进行了改进, 为疾病的诊断提供了宝贵的肺结节检测结果, 加快了计算机辅助诊断系统的运行效率。

1.2 国内外研究现状

人工神经网络简称为神经网络, 最开始以生物学上神经网络为基础, 了解了人脑的构造及对外界反应的机制之后, 以拓扑知识结构作为理论基础, 模拟人脑的一种行为。卷积神经网络是神经网络的一大分支, 它是一种前馈神经网络^[3], 包含卷积的堆叠和计算, 具有深度结构, 是深度学习中最具代表性的一种算法。在计算机视觉中, 主要处理的是图像数据。卷积神经网络的特点更加适合图像数据, 所以它是计算机视觉常用的一种神经网络, 更是研究的热点问题之一。

在计算机视觉领域主要有四个工作: 分类, 定位, 检测, 分割。目标检测是四项工作中的一项, 其根本意义在于发现影像中所关心的对象, 识别其类别及位置信息, 由于各类物体有不同的外观、形状和姿态, 再加上成像过程中光照, 遮挡及其他因素干扰, 在计算机视觉领域中, 目标检测始终是极具挑战性的工作之一。目前目标检测技术进行的研究已经取得了许多成果, 但是还存在很多问题需要解决, 其中最主要的一个方面是如何有效地实现快速而准确地提取出运动目标。目标检测算法主要有两种, 一类是有锚点框检测方法, 一类是无锚点框检测方法。2014 年, Ross Girshick 等提出了 R-CNN^[4], 第一次将深度学习应用到目标检测, 它的主要思路是首先使用无监督的选择性搜索 (Selective Search, SS)^[5], 将输入图像中具有相似颜色直方图特征的区域进行递归合并, 产生约 2000 个候选区域, 并将其相继送入一个 CNN 特征提取网络进行高层次的特征提取, 提取后再送入一个 SVM (Support Vector Machine)^[6] 进行物体分类。该算法在 PASCAL VOC 2012^[7] 数据集上取得了 mAP53.3% 的效果, 在那时已非常好。不过这个算法也存在着明显的不足。(1) 输入的图像大小会被缩放, 导致图像失真。(2) 采用搜索选择算法提取出的候选区域, 对于每个区域都要使用卷积神经网络进行计算, 计算量大且费时。(3) 训练测试繁杂, 候选区域地提取, 特征的获取、分类与回归都是单独操作的, 中间的数据分别保存了卷积提取的特征, 需要分别存储, 占了很大的硬盘空间。针对这些弊端, 在 Ross Girshick 看来, R-CNN 特征提取的阶段限制了深度神经网络的性能, 于是提出了 Fast R-CNN^[8] 来改进 R-CNN 的

缺点，使用感兴趣区域池化（Region-of-Interest Pooling）而不是空间金字塔池化^[9]的方式，使用全连接层替代之前的 SVM 层来进行物体的分类和检测框的修正，这样可以与之前的 CNN 构成一个整体，相比于 R-CNN 大大增强了检测任务的这一特性，提高了计算效率。2015 年，Ren 等人提出了 Faster R-CNN^[10]，针对生成候选区域的方式采用 RPN（Regional Proposal Network）进行生成，然后输入到 Fast R-CNN 中，进行物体分类和回归，Faster R-CNN 首次实现实时的物体检测，具有里程碑的意义。

2016 年，Joseph Redmon 等人提出了 YOLO（you only look once）算法^[11]，一个全新的目标检测方式，其基本思想是利用端到端的卷积神经网络，直接对物体进行类别与位置预测，相比于 Fast R-CNN 等两阶段目标检测模型，YOLO 实时性更高，但是检测精度略低，是第一个一阶段目标检测方法。2020 年，YOLOv4^[12]的出现在目标检测中引起热烈讨论，也得到了 YOLOv1 作者的认可，它的优点是提出了一种实时，高精度的目标检测算法。同年，YOLOv5 也迅速提出，优化了 YOLOv4 算法，参数量更少，实现更简单，网络更加轻量化。

深度学习这一技术应用于图像目标检测中所获得的成果，为医学图像检测中的相关研究带来了启示。特别是在肺结节检测任务上，许多研究团队基于深度学习算法提出了不同的解决方案，例如：肖毅、谢珺等人提出了一种端到端的融合注意力特征的多任务肺结节检测和分割算法^[13]，在该网络中主要使用了残差注意力特征融合模块对不同尺度的不同语义特征进行融合，以增强网络的特征丰富性，提高对肺结节的检测及分割精度；李莉、乔璐等人在解决 CT 影像中体积较小的结节难以检测、错检率高等问题，基于 R-FCN 网络结构加入了特征金字塔网络，以实现对不同层次的特征进行融合^[14]，且采用 DenseNet 作为主干网络，利用其密集连接机制以实现特征信息的高效复用，提升网络检测效率；J Sweta, C Pruthviraj 等人提出了一种两阶段卷积神经网络模型^[15]，以解决肺结节检测错检率高的问题，在第一阶段中使用 UNet 网络分割出肺结节区域，在第二阶段使用二分类网络判断分割得到的区域是否为肺结节；WJ Sori, J Feng 等人针对肺结节图形态复杂性对检测精度的影响问题，提出了一种双路径卷积神经网络^[16]，该网络在预处理阶段采用残差学习去噪模型用以去除图像的噪声干扰，之后采用双阶段卷积神经网络对肺结节进行检测。综上案例说明，采用深度学习算法对肺结

节检测具有一定的可行性,但该方法也存在有一定的共性问题需要进行攻破解决。主要为肺结节根据病症的不同也会表现出不同的形态特征,且由于 CT 影片主要为灰度图,仅能够依靠边界、纹理等有限的特征对肺结节进行识别。在上述方案中,主要通过优化检测流程、优化网络结构等两大类方式提升对肺结节的检测精度,虽然具有不错的效果,但离实际使用仍具有一定的差距。

1.3 本文主要研究内容

本文主要基于 YOLOv5 网络结构对其进行改进,分别提升在肺结节检测任务中的精度与速度,全文具体主要工作有如下几点:

1. 背景调研:在该工作中首先对肺结节的基础知识进行了解,包括其检测方式、判断依据等。其次对国内外基于计算视觉与深度学习算法的肺结节检测方法进行调研学习,了解掌握当前主要检测识别技术,及在肺结节检测任务中的难点,以确定后续对网络的改进方向。
2. 基础理论及网络选型:在该部分工作中首先对神经网络与卷积神经网络的基础原理进行学习,并进一步的,对基于神经网络设计的目标检测算法的网络结构、检测原理等知识进行选型。同时选取了双阶段检测网络与单阶段检测网络中具有代表性的算法进行对比实验,以选取合适的算法进行优化,最终经对比实验,从检测精度与检测速度两方面综合考虑,本文选取了 YOLOv5。
3. YOLOv5 检测精度提升:在该项工作中,本文通过设计新的特征提取网络与瓶颈网络搭配 YOLOv5 检测器实现对肺结节的高精度检测。首先,本文选用了编解码网络作为特征提取网络的主框架,之后设计了级联并行 Transformer 模块与 DeformCSP 卷积模块作为特征提取网络的组成单元,级联并行 Transformer 模块通过对特征图的分组级联计算,提升了全局特征的代表能力,DeformCSP 卷积模块通过加入了可形变卷积,实现了更细粒度的局部特征提取,通过编解码网络,将全局特征与局部特征进行融合,以提升网络的特征提取能力;其次,设计了特征重组网络,以对双维度编解码网络中不同网络层次的特征进行聚合与重组,解决了由于网络层次不同所造成的特征差异问题,使得在后续检测中是三个 YOLOv5 检测器所输入的特征近似相等。
4. YOLOv5 检测速度提升:在该工作中,本文主要依靠轻量化 Ghost 卷积进行优化,首先将其融合可形变卷积构造了 DGhost 卷积模块,使得卷积模块在

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/845232324331011120>