

摘要

基于动态 Bagging 集成策略的心肺影像多目标检测轻量化方法

生物学图像处理的主要目标对象是各种不同图像格式的医学影像，目前临床上所采用的生物学图像类型主要是 X 射线图像 (X-CT)。在目前的影像医学检查中，大多需要透过二维 CT 图像才能看到病变体，这通常要通过医师的检查才能诊断。由于中国逐步进入老龄化社会，医师工作繁重的情况十分明显，繁杂的工作可能影响检测的准确率。因此，设计一种计算机辅助医学影像检测系统应用于疾病检测前景广阔，对医生诊断病情具有指导意义。

目前，已有的医学图像检测方法只能识别单一种类疾病，无法对多种类疾病检测，并且现有的计算机视觉算法并非轻量化，检测过程也非常考验硬件资源。近年来，深度学习模型在准确率和轻量化方面都有出色表现，同时便捷的新算法也为医学图像检测提供了新途径、新角度。

针对目前医学图像检测存在的不足，本文提出了基于 Bagging 集成学习的深度学习方法用于医学图像检测，主要工作如下：

在集成学习框架方面，提出了一种解决各类样本数量不平衡的动态加权 Bagging 集成学习方法 (DW-Bagging)。针对本文数据集各类疾病数量不均衡，原始 Bagging 集成学习中等权重投票方式，扩大了模型方差和泛化误差。本文提出了置信度加权投票方式的 Bagging 集成学习策略来代替等权重投票的方式，优化后的集成学习方式 DW-Bagging 降低了模型方差，改善了模型泛化性能，从而提高模型检测精度。

在提升基学习器检测性能方面，针对检测过程中出现的效率低耗时长感受野小精度低等问题。首先选择 YOLOv5 作为集成学习的基学习器，但是原始 YOLOv5 主干特征提取网络采用 C3 结构，这使得模型参数冗余，推理分析效率低下，难以应用在某些高度集成的工业生产场景。为了实现网络模型的轻量化，并在速度和精度之间取得平衡，本文采取了一种更换骨干网络的方法，将更轻量的 EfficientNetLite 网络引入到模型中。这种替换方法使得模型具有更少的参数，提高了检测速度，并在不降低检测精度的情况下，实现了更高效的目标检

测任务；为了提升模型的感受野，本文提出了一种使用内容感知手段完成特征重组的特征金字塔网络构型（CARAFE-FPN），解决特征融合过程中目标尺度不一致而丢失小目标信息的问题；最后是损失函数的设计，为了提高检测精度并加速收敛，使用边界框损失函数 SIOU Loss，添加 Angle 惩罚成本有效地减少了损失的总自由度，极大地提升了训练过程中收敛速度。

本文在胸部影像 CT 数据集上进行实验，对数据进行转换格式、预处理和构造新数据集。本文设置多组边界框损失函数进行对比实验，选择性能最佳的 SIOU 损失函数。从弱学习器和集成学习框架两方面进行改进，选择 mAP 和 F1 分数作为基学习器精度评价指标，FPS 和 Params 作为弱学习器轻量化评价指标，Accuracy 作为集成学习模型的整体性能。与已有的检测方法相比，本文在动态集成框的基础上增加了轻量化的特征提取和上采样模块。实验结果表明，本文方法在弱学习器精度方面 mAP 达到 44.1%，F1 分数达到 45.8%，轻量化方面 FPS 为 64.3 f/s ，Params 为 14.4MB，动态加权模型集成后准确率到达 81.69%，所以本文方法在多目标胸部医疗图像检测上更有优势。

关键词：

多目标检测，深度学习，集成学习，轻量化，动态加权

Abstract

A Lightweight Method for Multi-target Detection in Cardiopulmonary Imaging Based on Dynamic Bagging Integration Strategy

The main target objects of biomedical image processing are medical images in various image formats. Currently, the type of biomedical images used clinically is mainly X-ray images (X-CT). In the current imaging medical examination, most of the lesions can only be seen through two-dimensional CT images, which usually can only be diagnosed through a doctor's examination. As China gradually enters into an aging society, it is obvious that doctors have a heavy workload, which may affect the accuracy of detection. Therefore, the design of a computer-aided medical imaging detection system has broad prospects for application in disease detection, and it has guiding significance for doctors to diagnose diseases.

At present, the existing medical image detection methods can only identify a single type of disease, but cannot detect multiple types of diseases, and the existing computer vision algorithms are not lightweight, and the detection process is also a test of hardware resources. In recent years, deep learning models have performed well in terms of accuracy and lightweight, and convenient new algorithms have also provided new ways and perspectives for medical image detection.

Aiming at the shortcomings of current medical image detection, this paper proposes a deep learning method based on Bagging ensemble learning for medical image detection. The main work is as follows:

In terms of the ensemble learning framework, a dynamic weighted Bagging ensemble learning method (DW-Bagging) to solve the imbalance of the number of various samples is proposed. In view of the imbalance in the number of various diseases in the data set in this paper, the original Bagging ensemble learning medium-weight voting method expands the model variance and generalization error. This paper proposes a confidence-weighted voting Bagging integrated learning strategy to replace equal-weight voting. The optimized integrated learning method DW-Bagging reduces model variance, improves model generalization performance, and improves model detection accuracy.

In terms of improving the detection performance of the base learner, it is aimed at

the problems of low efficiency, time-consuming, long receptive field, small receptive field and low precision in the detection process. Firstly, YOLOv5 is selected as the base learner for integrated learning, but the original YOLOv5 backbone feature extraction network adopts the C3 structure, which makes the model parameters redundant, reasoning and analysis efficiency is low, and it is difficult to apply to some highly integrated industrial production scenarios. In order to realize the lightweight of the network model and strike a balance between speed and accuracy, we adopted a method of replacing the backbone network and introducing the lighter EfficientNetLite network into the model. This replacement method makes the model have fewer parameters, improves the detection speed, and achieves a more efficient target detection task without reducing the detection accuracy; in order to improve the receptive field of the model, this paper proposes a content The feature pyramid network configuration (CARAFE-FPN) that completes feature recombination by means of perception solves the problem of missing small target information due to inconsistent target scales in the feature fusion process; finally, the design of the loss function, in order to improve detection accuracy and accelerate convergence, use boundary The box loss function SIoU Loss, adding the Angle penalty cost effectively reduces the total degree of freedom of the loss, which greatly helps the training convergence process and effect.

This paper conducts experiments on the chest image CT data set, converts the data format, preprocesses and constructs a new data set. In this paper, multiple sets of bounding box loss functions are set up for comparative experiments, and the SIoU loss function with the best performance is selected. Improvements are made from two aspects of the weak learner and the integrated learning framework. mAP and F1 score are selected as the accuracy evaluation indicators of the base learner, FPS and Params are used as the lightweight evaluation indicators of the weak learner, and Accuracy is used as the overall performance of the integrated learning model. Compared with the existing detection methods, this paper adds lightweight feature extraction and upsampling modules on the basis of dynamic integration boxes. The experimental results show that the mAP of the method in this paper reaches 44.1% in terms of weak learner accuracy, the F1 score reaches 45.8%, the FPS in terms of lightweight is 64.3 f/s , the Params is 14.4MB, and the accuracy rate reaches 81.69% after the dynamic weighted model is integrated. Therefore, the method in this paper is It has more advantages in the detection of multi-target chest medical images.

Key word:

Multi-target detection, deep learning, integrated learning, lightweight, dynamic weighting

目 录

第 1 章 绪论.....	1
1.1 研究背景和意义	1
1.2 国内外研究现状	2
1.2.1 集成学习研究现状	2
1.2.2 目标检测算法研究现状	5
1.2.3 轻量化目标检测算法研究现状	7
1.3 主要研究内容	7
1.4 本文组织结构	8
1.5 本章小结.....	9
第 2 章 集成学习理论基础	10
2.1 集成学习弱学习器的生成方法	10
2.1.1 Boosting 算法.....	10
2.1.2 Bagging 算法.....	11
2.1.3 Stacking 算法.....	12
2.2 集成学习结合策略	12
2.2.1 绝对多数投票法	13
2.2.2 相对多数投票法	13
2.3 本章小结.....	14
第 3 章 目标检测理论基础	15

3.1	二阶段目标检测方法	15
3.1.1	RCNN	15
3.1.2	Fast RCNN	16
3.1.3	Faster CNN	17
3.2	一阶段目标检测算法	18
3.2.1	SSD	19
3.2.2	YOLO	21
3.3	本章小结.....	30
第4章	基于改进 Bagging 集成学习的医学图像检测算法	32
4.1	算法流程.....	32
4.2	嵌入轻量化模块	32
4.2.1	引入 YOLOv5-EfficientNetLite 网络	32
4.2.2	构建 YOLOv5-CARAFE 特征上采样	38
4.3	设计 SIOU 边界回归损失函数	41
4.3.1	IoU	41
4.3.2	GIOU	42
4.3.3	DIOU	43
4.3.4	CIoU	44
4.3.5	SIOU	45
4.4	改进 Bagging 集成学习策略	47
4.5	本章小结.....	48

第 5 章 实验设计及结果分析	49
5.1 数据集介绍及处理	49
5.1.1 数据集介绍	49
5.1.2 数据集处理与构建	49
5.2 实验环境及参数设置	52
5.3 性能指标评价	53
5.4 实验结果及分析	55
5.4.1 目标检测算法对比实验	55
5.4.2 损失函数对比实验	55
5.4.3 轻量化模型对比实验	56
5.4.4 集成学习对比实验	57
5.5 本章总结.....	59
第 6 章 总结和展望	60
6.1 总结.....	60
6.2 研究展望.....	61
参考文献.....	62
作者介绍及在学期间所取得的科研成果	69
致谢.....	70

第1章 绪论

1.1 研究背景和意义

计算机视觉是一种利用计算机技术和算法处理和分析图像或视频数据的领域。通过模拟人类视觉系统的工作原理，计算机视觉能够实现对目标的自动识别、分类、检测、跟踪、分割等任务，并能够在很大程度上弥补人类视觉系统的局限性。计算机视觉技术以其高精度、高速度、低成本、便于信息集成等多重优点，成为医学辅助、智慧交通和航空航天等领域的重要支撑，能够替代人工进行高强度、连续、精准的作业，显著提高工作效率和质量。

在计算机视觉领域，有三个基本的视觉检测任务：第一类是图像分类^[1]，旨在给定的图像进行语义分类，即判断该图像中包含哪些对象。第二类是目标检测^{[2][3]}，在图像中定位并识别特定对象。第三类是语义分割^[4]，主要针对将图像分成不同的区域，并对每个区域进行语义分类。这些问题都是计算机视觉中的核心问题，也是人脸识别^{[6][7][8]}、行人检测^{[9][10][11]}、视频分析^{[12][13]}和标志检测^{[14][15][16]}等计算机视觉应用的基础。

近年来，基于人工智能分析和医疗诊断交叉应用越来越普及，借助其超越人类记忆力、学习能力和运算速度的特点，成为医疗健康行业的重要助手和支持。智能医学影像是指在医学影像诊断中应用人工智能技术的一种具体形式，这种技术可以有效提高医生的诊断准确性和效率，为患者提供更好的医疗服务。我国是人口大国，人口老龄化居高不下，医疗水平呈现地区差异化。当前，新冠病毒仍在持续变异，我国医疗资源分配存在不平衡，一些地区医疗床位和重症医疗床位供给不足，平均每千人口仅有 6.7 张普通床位和不到 0.4 张 ICU 床位，无法满足全民健康服务的需求，并且“白肺”症状的持续增多。面对当前医疗领域存在的问题，导致医生任务繁重，医学图像处理辅助医生诊断病情具有重要地位，随着医学影像数据的不断增长，智能移动设备的发展趋势已经由云计算向边缘计算和移动计算转变，以提高响应速度和节省存储成本。因此，轻量化的深度卷积神经网络被广泛应用于移动设备上，其设计目的是为了在低硬件条件下实现高性能的影像分析与诊断^[17]。

然而，现如今有了一些医学图像处理的方法，但仍然存在一些问题，如检测速度慢、只能单目标疾病检测和检测精度低等，目前的方法还有很大的改进提升空间。本文提出了基于 Bagging 集成学习的轻量化多目标检测方法，对其专业性问题进行探讨分析，针对医疗图像数据集的特点，对检测框架进行适应性创新改进，为诊断医生提供辅助指导。

1.2 国内外研究现状

1.2.1 集成学习研究现状

在机器学习和模式识别领域，模式分类和回归问题是最重要的挑战之一。模式分类问题通常是指将输入数据分为不同的离散类别，而模式回归问题则是指将输入数据映射到一个连续的输出空间中。集成学习在提高学习模型的泛化能力方面具有先天的有利条件，能够高效地解决实际应用问题。因此，人们所关注的核心问题是如何让学习系统具有更好的泛化能力。所以，近年来对集成学习理论算法的创新一直是机器学习的热门领域^[18]。

1979 年，Dasarathy^[19]提出集成系统（Ensemble System）是探究如何训练和组合多个学习器获得最佳性能来完成学习任务的思想，并使用线性分类器和最近邻居（NN）分类器组成的复合系统来揭示原理。通常情况下，多个有机组合的基学习器性能比单个基学习器效果更优。并且该思想产生的算法在众多场景中得到广泛的应用，例如 Hansen 等人^[20]在 1990 年设计了一种基于神经网络的集成模型，通过这种模型可以提高分类神经网络的性能，减少方差和泛化误差，同时也能够提高神经网络的泛化能力。Schapire 等人在同一年给出了 Boosting 算法的设想，该猜想表明可以通过组合多个弱学习器来合并一个强学习器。但是，不足之处在于需要预先知道子学习器误差的范围，这使得该算法在实际问题中难以使用。为了解决这个问题，Freund 和 Schapire 于 1995 年提出了 AdaBoost 算法^[21]，这一算法在机器学习方面也广受重视，成为 Boosting 家族中最有特色的算法。在此以后，集成学习在不断探索中得到迅猛发展，出现了许多优秀的新思路新方法。例如，Jacobs 等^[22]于 1991 年提出的混合专家模型，此模型的主要思想是使用分而治之的策略并行学习，并将每个“专家”学习的结果综合起来获得一个整体的结论。Wolpert 等人^[23]在 1992 年提出了一种方案，旨在通过评

估泛化器对于学习集的偏差，从而最小化一个或多个泛化器的泛化错误率。这项研究成果为解决泛化能力问题提供了新思路，并在机器学习领域产生了广泛影响。Breiman^[24]于1996年提出了Bagging算法，它是一种通过多个基学习器并使用它们来生成获得一个强学习器的方法。集成后的预测数值结果是对各基学习器的结果的平均，预测类别则采取少数服从多数的投票机制。如果学习集被扰动，会导致构建的学习器出现显著变化。在这种情况下，使用Bagging学习可以显著提高模型的准确性。

在近几年的集成学习研究中，众多学者设计出许多泛化性能更强和鲁棒性更高的集成学习框架。例如，Wang等人^[25]开发了一种利用数据以解决深度学习在成功应用中对大量高质量标注数据依赖问题的集成学习方法。然而，在现实应用中，大多数数据都受到标签噪声的影响，数据注释的成本高于数据采集，也更加耗费资源，这在一定程度上限制了深度学习模型在具有超大规模且标记质量优异数据集的部署。Xu等人^[26]提出了一种全新的基于不同任务的异质基学习器集成学习方法来检测不同场景下的森林火灾。Yang等人^[27]提出了一种新的数据聚类集成方法，它从原始数据集中的采取混合采样的方式生成训练子集，迭代地构建聚类集成模型。Huang等人^[28]设计了一种包括代理、奖励函数和轮盘赌推荐算法集成自适应机制（AMRE）的方法。首先，每个代理对应一个推荐算法。它还包含一个奖励值，以确定是否应该保留或替换它。其次，奖励函数被设计为根据推荐结果更新奖励值。错误的建议带来惩罚，而正确的建议带来奖励。最后，当奖励值低于给定阈值时，轮盘赌会选择另一个代理人。Jain等人^[29]提出了一种基于深度神经网络集成学习的方法，不使用单个DNN作为分类器，而是使用一个由七个独立DNN学习器组成的集合，这些DNN都会保持它们的架构和内在属性相同，但是使用不同的数据输入。为了在训练输入中引入多样性，每一个DNN将会删除七分之一的输入数据，并从剩余的样本中通过Bootstrap抽样进行补充。它在分配最终类标签之前先对预测结果执行两步置信度检查。Zhou等人^[30]提出用一种联合学习算法，来获得最终的共识聚类结果的自定进度聚类集成（SPCE）方法。利用基聚类将实例难度评价和集成学习融合到一个统一的框架中。在过去的几十年中，集成系统已被证明是提高分类算法

准确性和稳定性的有效方法。然而，如何获得多个基分类器的有效组合仍然是一个悬而未决的问题。Jiang 等人^[31]利用遗传算法，提出了一种新颖的自适应堆叠集成模型（SSEM）。与其他集成学习分类算法不同的是，SSEM 有选择地融合不同的基分类器，通过遗传算法自动选择最优的基分类器组合和基分类器的超参数，所有的机器学习方法都可以是 SSEM 的组成部分。Zhou 等人^[32]提出域自适应集成学习（DAEL）的一致性架构解决在两种设置下研究将深度训练模型从众多源域泛化到单个目标域的问题。当未标记的目标数据可用时，它是一个多源无监督域适应（UDA）问题，否则是一个域泛化（DG）问题。DAEL 的目标是协同学习分类器专家，以便在形成集成时，它们可以利用彼此的互补信息，从而更有效地应对未知的目标领域。为此，每个源域依次用作伪目标域，其自己的专家向从其他来源学习的非专家集合提供监督信号。Sikder 等人^[33]提出了一种基于决策树集成学习技术的新方法，利用眼底医疗数据集中提取的灰度值和纹理信息进行糖尿病性视网膜病变的诊断。该方法在患者失明检测数据集上进行了验证，结果表明该方法稳健可靠。这种方法可以成为大规模筛查 DR 的宝贵工具，从而降低由此引起的视力丧失率。Gultekin 等人^[34]提出一种新的动态合并集成算法，合并时运用 online 方式来更新卷积神经网络集成要素的权重。对于给定的测试实例，所提出的算法最初为每个分类器分配相等的权重，并随着迭代增加最佳 k 个的权重。近年来，基于集成学习的软件缺陷预测（SDP）受到广泛关注。然而，利用集成学习思路在改善基分类器的散度方面仍然存在一些不足。Wei 等人^[35]提出 E_RHFS 的新集成学习算法并将其应用于 SDP。在 E_RHFS 中，首先利用基于加权复杂度的 SMOTE 来管理不平衡数据；其次，利用重采样技术从平衡数据中生成多个采样集。此外，对于每个采样集，采用混合特征选择方法来选择特征，该方法结合了粒度决策熵和随机子空间方法；第三，在每个样本集上训练一个弱学习器；最后通过投票将所有弱学习器组合起来，得到最终的检测软件漏洞集成学习器。Zhu 等人^[36]提出对一个几何结构集成（GSE）学习框架来解决机器学习中不平衡数据集的分类这个巨大挑战。众所周知，已有的集成方法根据各种权重训练和组合一系列基本分类器，可能缺乏几何意义。相反，GSE 通过欧几里德度量生成超球体来界定和摆脱冗余的多数

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/285031004240011120>