

丁士宁. 基于 CNKI 数据库的 Faster R-CNN 模型应用可视化分析[J]. 智能计算机与应用, 2024, 14(4): 113-117. DOI: 10.20169/j.issn.2095-2163.240416

基于 CNKI 数据库的 Faster R-CNN 模型应用可视化分析

丁士宁

(信阳农林学院 信息工程学院, 河南 信阳 464000)

摘要: 为了解 Faster R-CNN 模型在国内的研究现状和研究热点, 基于在 CNKI 数据库中检索的 248 篇核心期刊文献, 使用 Excel 和 VOSviewer 软件对年度发文量、作者、关键词进行分析并进行可视化。结果表明: 国内关于 Faster R-CNN 模型的研究最早出现在 2017 年, 且在 2017-2021 年度发文量逐年递增, 2021 年年度发文量达到 73 篇的峰值, 并在 2022 年回落到 43 篇; 超过 95% 的作者仅发表 1 篇文献, 发表文献不少于 3 篇的作者仅 9 位, 且作者之间的合作松散; Faster R-CNN 模型的应用领域主要集中于交通标志检测、车辆检测、杂草识别、遥感图像等方面, 模型的改进方法包括 Soft NMS、注意力机制、K-means 聚类等。

关键词: Faster R-CNN 模型; Excel; VOSviewer; 核心期刊; 可视化

中图分类号: TP391.9

文献标志码: A

文章编号: 2095-2163(2024)04-0113-05

Visualization analysis of Faster R-CNN model research based on CNKI database

DING Shining

(Xinyang Agriculture and Forestry University, School of Information Engineering, Xinyang 464000, Henan, China)

Abstract: In order to understand the research status and hotspots of the Faster R-CNN model in China, based on 248 core journal articles retrieved from the CNKI database, Excel and VOSviewer software were used to analyze and visualize the annual publication volume, authors, and keywords. The results show that research on the Faster R-CNN model in China first appeared in 2017, and the annual publication volume increased year by year from 2017 to 2021. The annual publication volume reached a peak of 73 articles in 2021, and dropped back to 43 articles in 2022. More than 95% of authors have published only one article, and only 9 authors have published no less than 3 articles. And the cooperation between the authors is loose. The application fields of the Faster R-CNN model mainly focus on traffic sign detection, vehicle detection, weed identification, and remote sensing images. The improvement methods of the model include Soft NMS, attention mechanism, K-means clustering, etc.

Key words: Faster R-CNN model; Excel; VOSviewer; core journals; visualization

0 引言

目标检测是计算机视觉领域内的核心问题之一, 其目的是确定图像内感兴趣物体的类别和位置。基于深度学习的目标检测方法根据检测阶段, 可以划分为单阶段(One-stage)检测模型和双阶段(Two-stage)检测模型^[1]。单阶段模型只需一步即可生成预测目标的种类和位置, 其中包括 Retinanet^[2]、SSD^[3]、YOLO 系列^[4]等模型。双阶段模型将目标检测分为两步, 首先生成大量的候选框, 再预测目标的种类和位置, 其中包括 Faster R-CNN^[5]、Mask R-CNN^[6]等模型。

Faster R-CNN 作为经典的二阶段检测模型, 通

过区域候选网络(Region Proposal Network, RPN)和 R-CNN 联合训练, 实现高效准确的目标检测。Faster R-CNN 模型现已在各个方面得到应用, 如医学病灶检测、钢板表面缺陷检测、人体手部位姿估计、农业病虫害识别、交通标志检测等。宋雯琦等^[7]将 Faster R-CNN 模型应用到肺炎检测上, 并在原模型的基础上, 在特征提取主干网络、区域候选框选择策略方面提出改进, 最后实验验证模型的有效性。李雪露等^[8]将改进的 Faster R-CNN 模型应用到钢板表面缺陷检测上, 提升了缺陷的检测精度, 缺陷定位更准确。郑涵等^[9]提出了基于改进 Faster R-CNN 的手部位姿估计方法, 解决了手部检测中的自遮挡和尺度问题, 并且提升了检测准确率。宋中

基金项目: 信阳农林学院青年教师科研基金项目(QN2021057)。

作者简介: 丁士宁(1993-), 男, 硕士, 助教, 主要研究方向: 模式识别与人工智能。Email: 1728252083@qq.com

收稿日期: 2023-09-11

山等^[10]将在 Faster R-CNN 模型的基础上提出改进并应用到柑橘病虫害,实验验证该方法具有较好的识别速度和鲁棒性。李厚杰等^[11]提出了一种基于伪样本正则化 Faster R-CNN 模型,实现了各种交通标志的检测,并有效降低了过拟合的风险。

本文首先在中国知网(China National Knowledge Infrastructure, CNKI)上检索将模型 Faster R-CNN 作为研究对象的论文,以此检测结果作为基础,采用 Excel 和 VOSviewer 工具,分别对年度发文量、作者和关键词进行分析,探讨国内该模型的研究现状和研究热点。

1 数据来源与研究方法

本文以 CNKI 作为检索平台,只探索以 Faster R-CNN 模型为主要研究对象的期刊论文,将该模型作为对比模型的研究文章不在考虑之内,因此以篇名“Faster R-CNN”作为检索条件。仅检索期刊文献,期刊来源类别勾选“SCI 来源期刊”、“EI 来源期刊”、“北大核心”、“CSSCI”、“CSCD”和“AMI”,数据检索日期为 2023 年 9 月 8 日,共检索到 272 条结果。将检索结果以 EndNode 格式导出。然后按照如下规则进行数据筛选:

(1)根据论文名称、摘要、关键字,筛除明显与主题无关的文献;

(2)筛除英文文献,本文仅对中文文献进行分析;

(3)筛除不包含发表日期的文献。

经过筛选后,保留文献 248 篇。

本文使用 Excel 对 Faster R-CNN 模型的年度发文量进行分析,通过折线图直观展示该模型的研究文献数量随年度变化情况;使用荷兰莱顿大学研发的文献可视化工具 VOSviewer 软件^[12]对作者、关键词进行分析,生成作者合作网络和关键词聚类图谱,探讨该模型的研究情况。

2 结果与分析

2.1 年度发文量分析

年度发文量体现了该研究是否为热点^[13]。Faster R-CNN 模型的年度发文量曲线如图 1 所示。国内关于 Faster R-CNN 模型研究始于 2017 年度,发文量仅为 2 篇。之后,随着年限的增加发文量也迅速增加,2021 年达到峰值,该年度的发文量达到 73 篇。2022 年的年度发文量有所回落,该年度的发文量达到 43 篇,比上一年减少 30 篇。在 2019 年-2022 年,每年的文献发表数据均超过了 40 篇。经

分析认为,关于 Faster R-CNN 模型的研究在近几年维持在较高的热度,具有广阔的研究前景。

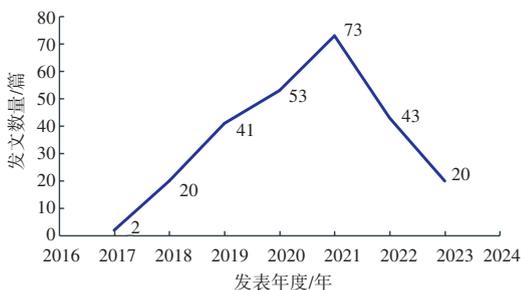


图 1 年度发文量曲线

Fig. 1 Annual document volume curve

2.2 发文作者分析

对文献作者进行统计分析,可以探究该领域研究的核心作者和合作强度^[14]。本文检索出的所有文献共包含作者 998 人,其中出现频次为 1 次的作者有 952 人,占比 95.39%;出现频次在 2 次的作者有 37 人,占比 3.71%;出现频次在 3 次的作者有 8 人,占比 0.80%;出现频次在 4 次的作者仅有 1 人,占比 0.10%。明显看出,出现频次为 1 的作者数量明显远多于其他出现频次的作者数量。其中出现频次在 3 次及以上的作者有 9 人,见表 1。

表 1 出现频次≥3 的作者

Table 1 Authors with a frequency no less than 3

序号	作者姓名	出现频次
1	赵振兵	4
2	周建平	3
3	鱼鹏飞	3
4	许燕	3
5	王纪武	3
6	罗海保	3
7	李伟	3
8	樊湘鹏	3
9	翟永杰	3

使用 VOSviewer 软件绘制的作者合作网络图谱如图 2 所示(图中仅展示出现频次在 2 次及以上的作者)。在图中,节点代表作者,节点之间的连线代表作者之间的合作。从图中未发现有重大影响力的核心作者,且作者之间的合作比较分散,仅有部分小团体,如:王凤随、刘芙蓉、陈金刚、王启胜组成的小团体,以及周建平、樊湘鹏、温德圣、许燕组成的小团体等,且团体之间的合作极少。

结合作者出现频次的统计和作者合作网络图谱,认为作者关于 Faster R-CNN 模型的研究分散,且不存在有重大影响力的作者。

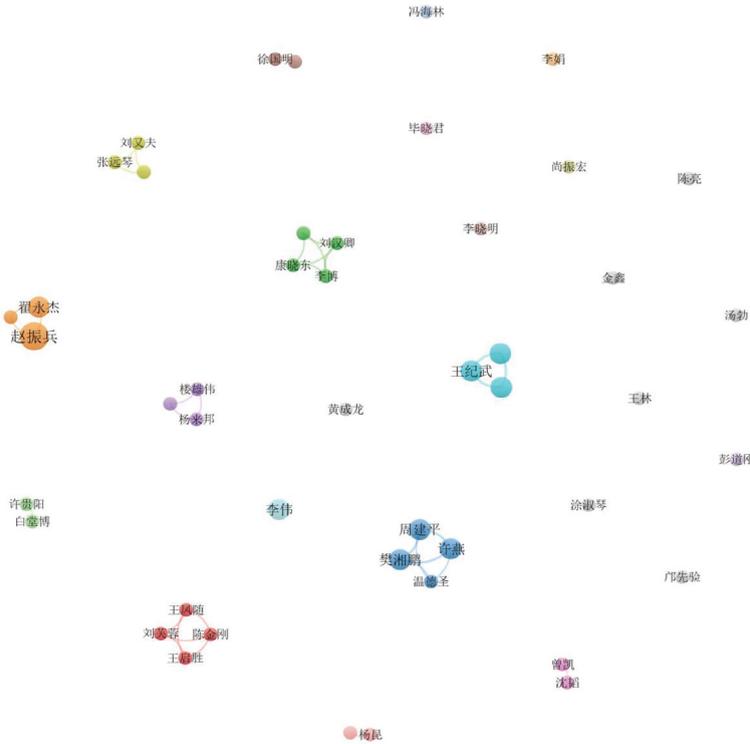


图 2 作者合作网络图谱

Fig. 2 Collaboration network map of authors

2.3 关键词分析

关键词是文章的精炼总结,对关键词进行分析可以明了该领域内的研究现状和研究热点^[15]。在分析之前,首先进行简单的同义词合并,如将 Faster R-CNN 算法、Faster R-CNN 网络、Faster R-CNN 模型等关键词统一归并到关键词 Faster R-CNN 上。同义词归并之后,出现频次前 10 的关键词见表 2。出现频次排名前 3 的关键词是 Faster R-CNN、深度学习、目标检测。Faster R-CNN 模型从属于深度学习,主要用来进行目标检测。虽然这 3 个关键词出现频次很高,但缺乏分析意义,因此在 VOSviewer 软件绘制关键词图谱时已被剔除。绘制的关键词聚类图谱如图 3 所示,在图谱上关键词的出现频次均大于等于 3。主要关键词包括卷积神经网络、特征金字塔网络、残差网络、注意力机制、K-means、roi align 等。其中,卷积神经网络、特征金字塔网络、残差网络、非极大值抑制、roi pooling 等词涉及到 Faster R-CNN 的模型结构,而注意力机制、K-means、roi align 等涉及到模型的改进。下面就这些关键词作简要说明:

(1) 卷积神经网络和残差网络。骨干特征提取

网络 Backbone 是 Faster R-CNN 模型的基础结构之一,其主要作用是提取输入图像的特征。Backbone 一般选择某种卷积神经网络,最常用的是残差网络 ResNet50/101^[16]。

(2) 特征金字塔网络。特征金字塔网络 (Feature Pyramid Networks, FPN) 对 Backbone 输出的单尺度或多尺度特征图像进行特征分治和特征融合,从而提高模型的检测效果^[17-18]。由于特征金字塔网络强大的性能和模块化的设计,该结构已广泛应用到 Faster R-CNN 网络。

(3) ROI Pooling 和 ROI Align。ROI Pooling 层保证任意大小的特征图均可以变成指定维度的输出,但其中的两次取整操作会带来误差,从而影响检测精度,而模型 Mask R-CNN^[6]中 ROI Align 结构克服了这一问题,用 ROI Align 结构取代 ROI Pooling 层可以作为 Faster R-CNN 网络的一个改进。

(4) 非极大值抑制。非极大值抑制 (Non-Maximum Suppression, NMS) 用于剔除冗余检测框,保留高质量的候选框。NMS 剔除机制严格,对有重叠物体的检测容易造成漏检,而 Soft NMS 算法引入得分惩罚机制缓解了这一问题,该方法对双阶段模

型友好,是 NMS 方法的改进^[19]。

(5)注意力机制。注意力机制可以使模型关注重要的特征,忽视不重要的或无关的特征。将注意力机制引入模型中,可以使模型将算力集中到重要特征上,提高检测精度^[20]。

(6)K-means。在 Faster R-CNN 模型中,锚框(Anchor)的参数是先验参数,需要提前设置,模型中默认的锚框参数未必适用于特定任务。因此针对特定任务采集的数据集进行聚类分析,生成适用于特定任务的锚框参数,是不错的解决办法^[21]。K-means出现频次为10,认为常用K-means聚类方法生成新的锚框。

除了这些描述模型结构和改进模型的关键词外,关键词聚类图谱还出现了交通标志检测、车辆检测、杂草识别、遥感图像等关键词,这些代表模型应

用领域的关键词出现频次不少于3,认为模型在这些领域的应用研究相对较多。

表2 出现频率前10的关键词

Table 2 Top 10 keywords with the frequency of occurrence

序号	关键词	出现频次
1	Faster R-CNN	173
2	深度学习	81
3	目标检测	72
4	卷积神经网络	45
5	特征金字塔网络	21
6	残差网络	14
7	图像处理	12
8	改进 Faster R-CNN	11
9	注意力机制	10
10	K-means	10

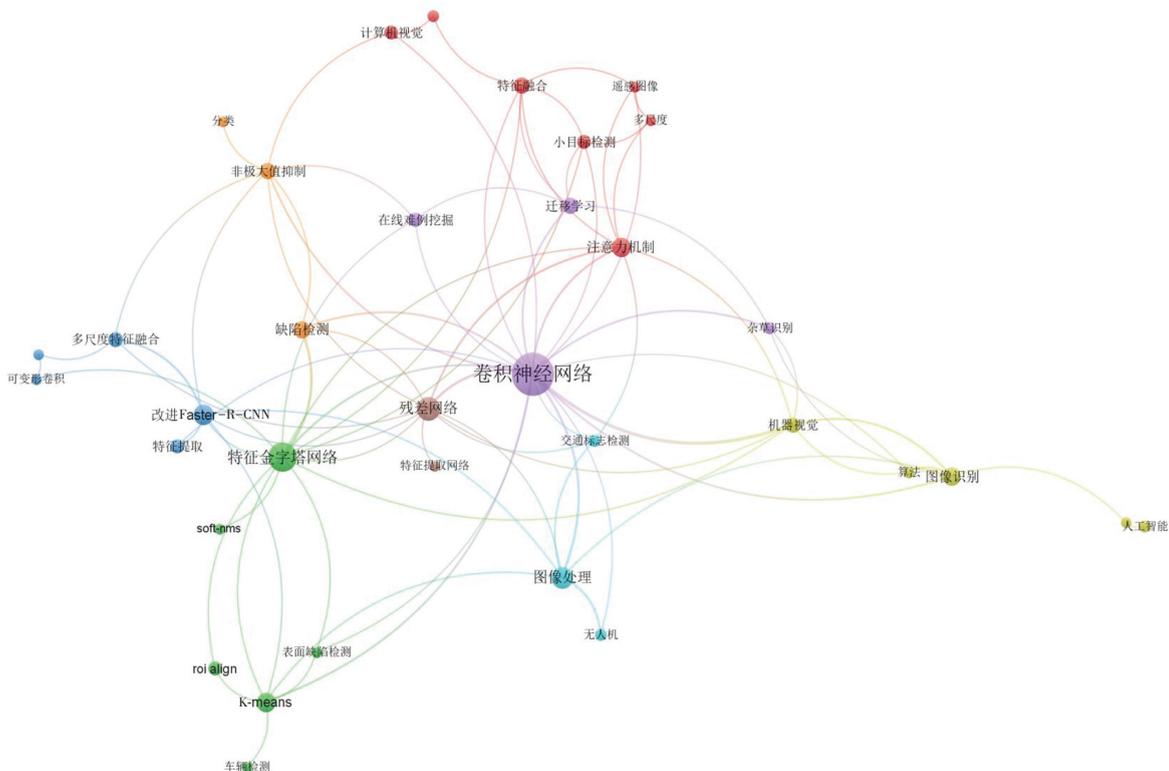


图3 关键词聚类图谱

Fig. 3 Keyword Clustering Map

3 结束语

本文在 CNKI 数据库上检索以 Faster R-CNN 为研究对象的核心期刊文献,经过数据筛选后,使用 Excel 和 VOSviewer 软件对年度发文量、作者、关键词进行分析,分析结果表明:

(1)2017 年知网上首度出现关于该模型的核心期刊文献,此后年度文献发表数量随着年限增加逐年递增,且在 2021 年达到 73 篇的峰值,并在 2022 年有小幅回落;

(2)关于该模型的研究比较分散,不存在具备核心影响力的作者;

(3)通过关键词分析,认为应用集中于交通标志检测、车辆检测、杂草识别、遥感图像等领域,在应用时,并非简单的应用原始模型,还从骨干特征提取网络、特征金字塔网络、区域候选框选择策略、锚框选择、增加注意力机制等多方面对模型提出了改进。

参考文献

- [1] 李心宇. 基于深度学习的 SAR 图像舰船目标检测[D]. 烟台:烟台大学,2023.
- [2] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2):318-327.
- [3] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multiBox detector[C]//*Proceedings of European Conference on Computer Vision*. 2016: 21-37.
- [4] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, USA: IEEE, 2016:779-788.
- [5] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6):1137-1149.
- [6] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2):386-397.
- [7] 宋雯琦,赵荣彩,姜旭,等. 基于 Faster R-CNN 的肺炎目标检测[J]. *计算机工程与设计*,2023,44(7):2087-2092.
- [8] 李雪露,杨永辉,储茂祥,等. 基于改进 Faster R-CNN 的钢板表面缺陷检测[J]. *安徽大学学报(自然科学版)*,2023,47(2):66-73.
- [9] 郑涵,田猛,赵延峰,等. 基于改进 Faster R-CNN 的手部位姿估计方法[J]. *科学技术与工程*,2023,23(3):1160-1167.
- [10] 宋中山,汪进,郑禄,等. 基于二值化的 Faster R-CNN 柑橘病虫害识别研究[J]. *中国农机化学报*,2022,43(6):150-158.
- [11] 李厚杰,王法胜,贺建军,等. 基于伪样本正则化 Faster R-CNN 的交通标志检测[J]. *吉林大学学报(工学版)*,2021,51(4):1251-1260.
- [12] VAN ECK N, WALTMAN L. Software survey: VOSviewer, a computer program for bibliometric mapping[J]. *Scientometrics*, 2010, 84(2): 523-538.
- [13] YSA B, SWAB C, GGA B. Trends of research on polycyclic aromatic hydrocarbons in food: A 20-year perspective from 1997 to 2017[J]. *Trends in Food Science and Technology*, 2019, 83: 86-98.
- [14] LI W, ZHAO Y, WANG Q, et al. Twenty years of entropy research: a bibliometric overview[J]. *Entropy*, 2019, 21(7): 694.
- [15] 林德明,陈超美,刘则渊. 共被引网络中介中心性的 Zipf-Pareto 分布研究[J]. *情报学报*, 2011, 30(1): 76-82.
- [16] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.Las Vegas, USA:IEEE, 2016: 770-778.
- [17] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*. Honolulu, USA:IEEE, 2017:936-944.
- [18] CHEN Q, WANG Y M, YANG T, et al. You only look one-level feature[C]//*Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville, USA:IEEE, 2021:13034-13043.
- [19] BODLA N, SINGH B, CHELLAPPA R, et al. Soft-NMS: Improving object detection with one line of code[C]//*Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV)*. Venice, Italy:IEEE, 2017:5562-5570.
- [20] 李哲,张慧慧,邓军勇. 基于改进 Faster R-CNN 的交通标志检测算法[J]. *液晶与显示*,2021,36(3):484-492.
- [21] 李玉,汤勃,孙伟,等. 基于改进 Faster R-CNN 的钢板表面缺陷检测[J]. *组合机床与自动化加工技术*,2022(5):113-115,119.