

# 深度学习运用于胸腔X射线影像研究的文献计量学分析



黄夏璇<sup>1,2</sup>, 陈咏梅<sup>3</sup>, 袁师其<sup>1,2</sup>, 黄 韬<sup>2</sup>, 何宁霞<sup>2</sup>, 吕 军<sup>2,4</sup>

1. 暨南大学附属第一医院神经内科 (广州 510630)
2. 暨南大学附属第一医院临床研究部 (广州 510630)
3. 暨南大学学报编辑部 (广州 510632)
4. 广东省中医药信息化重点实验室 (广州 510632)

**【摘要】目的** 基于 SCIE 和 PubMed 数据库了解深度学习在胸腔 X 射线影像相关领域研究文献发表情况。**方法** 检索 2017 年 1 月 1 日至 2021 年 12 月 31 日收录于 SCIE 和 PubMed 的关于胸腔 X 射线影像的文献, 针对发文量、出版机构、出版期刊、引文、作者及关键词等信息进行统计分析。**结果** 共纳入文献 440 篇, 深度学习在胸腔 X 射线影像研究相关文献发文量呈逐年增长趋势。美国发文量最多, 总被引频次为 4 409 次, 篇均被引频次为 12.32 次, 美国的 *IEEE Access* 期刊发文量最多, 达 29 篇。发文量排名第一的出版商是德国的 Springer Nature, 为 83 篇。核心作者 7 位, 发文最多的有 10 篇, 主要关键词频次出现最多的为 COVID-19。**结论** SCIE 和 PubMed 收录的关于深度学习在胸腔 X 射线影像相关领域的文献整体呈逐年上升趋势, 基本以英文文献增长为主, 核心作者群尚未形成, 尚未出现引文量和发文量均丰富的领军人物, 高影响力文献数量有限。

**【关键词】** 深度学习; 胸腔 X 射线影像; SCIE; PubMed; 文献计量学分析; 新型冠状病毒肺炎

## Bibliometric analysis of deep learning in chest X-ray imaging research

Xia-Xuan HUANG<sup>1,2</sup>, Yong-Mei CHEN<sup>3</sup>, Shi-Qi YUAN<sup>1,2</sup>, Tao HUANG<sup>2</sup>, Ning-Xia HE<sup>2</sup>, Jun LYU<sup>2,4</sup>

1. Department of Neurology, The First Affiliated Hospital of Jinan University, Guangzhou 510630, China
2. Department of Clinical Research, The First Affiliated Hospital of Jinan University, Guangzhou 510630, China
3. Editorial Department of Journal of Jinan University, Guangzhou 510632, China
4. Guangdong Provincial Key Laboratory of Traditional Chinese Medicine Informatization, Guangzhou 510632, China

Corresponding author: Jun LYU, Email: lyujun2020@jnu.edu.cn

**【Abstract】Objective** To investigate the development of SCIE and PubMed deep learning literature on chest X-ray imaging. **Methods** The literature on chest X-ray images published in SCIE and PubMed from January 1, 2017 to December 31, 2021 was searched, and

DOI: [10.12173/j.issn.1004-5511.202201031](https://doi.org/10.12173/j.issn.1004-5511.202201031)

基金项目: 广东省科技计划项目 (2021B1212040007)

通信作者: 吕军, 博士, 研究员, 博士研究生导师, Email: lyujun2020@jnu.edu.cn

<https://yxxz.whuzhmedj.com/>

the number of articles, publishing institutions, journals, citations, authors and keywords were statistically analyzed. **Results** A total of 440 papers were included, and the number of papers presented an annual growth trend. The country with the largest number of papers was the United States, with a total citation frequency of 4 409 times and an average citation frequency of 12.32 times. The *IEEE Access* in the United States published the most articles, reaching 29 articles. The number one publisher is Germany Springer Nature with 83 articles. There are 7 core authors, 10 of which have published the most papers, and the most frequently cited keywords in the research content are COVID-19. **Conclusion** The literature on deep learning in the field of chest X-ray imaging collected in SCIE and PubMed shows an overall upward trend year by year, mainly in English. However, a core author group has not yet been formed, and there is no clear leader with prolific citations and publications, and the number of high-impact publications is still limited.

**【Keywords】** Deep learning; Chest X-ray; SCIE; PubMed; Bibliometric analysis; COVID-19

近年来,深度学习的计算机视觉技术逐渐成为人工智能领域的研究热点<sup>[1]</sup>。作为机器学习的领域之一,深度学习利用深度神经网络模拟人脑进行分析学习,获得特征表达<sup>[2]</sup>。2017年以来,深度学习逐渐与临床影像结合,图像分类检测系统在临床医学影像方面开始得到有效实践<sup>[3-4]</sup>。胸腔X射线影像是肺部和心血管疾病诊断和治疗的重要依据,但大量的数据和潜在的病理特征信息大幅降低了诊断的效率,而深度学习的出现及其与人工智能的结合有助于对医学图像进行特征提取和自动化分类,为提高诊断效率提供了发展方向<sup>[5]</sup>。本文拟对SCIE和PubMed收录的关于深度学习在胸腔X射线影像方面的研究进行梳理归纳并分析总结,以更好地了解国内外深度学习结合临床影像的研究热点及发展趋势,为进一步推动深度学习及其临床应用提供参考。

## 1 资料与方法

### 1.1 纳入与排除标准

纳入标准:①SCIE和PubMed数据库中已发表的与深度学习在胸腔X射线影像有关的文献<sup>[6-9]</sup>;②发表时间为2017年1月1日至2021年12月31日。

排除标准:①研究主题仅为深度学习而与胸腔X射线影像无关的文献;②文献载体为图书、报纸、成果;③重复研究。

### 1.2 文献检索与数据来源

检索2017年1月1日至2021年12月31日

收录于SCIE、PubMed数据库,与深度学习在胸腔X射线影像研究相关的文献,主要检索词为:Deep learning、Machine Learning、Transfer learning、chest X-rays\*、CXR\*。以PubMed为例,检索策略见框1。

```
#1 Deep learning[Title/Abstract] OR Machine Learning[Title/Abstract] OR Transfer learning[Title/Abstract]
#2 (chest X-rays*[Title/Abstract] OR CXR*[Title/Abstract])
AND ("2017/01/01"[Date - Publication] : "2021/12/30"[Date - Publication])
#3 #1 AND #2
```

#### 框1 PubMed检索策略

#### Box 1. Search strategy in PubMed

### 1.3 文献筛选与资料提取

利用EndNote X7软件对检出的文献进行分类整理,去除重复文献。对检出文献研究年份、国家(地区)、研究机构、研究方向、作者、主题分类、来源出版物、高被引文献及论文总被引频次等信息进行提取。采用普赖斯曲线方程 $F(t)=a \cdot e^{bt}$ ,揭示文献随时间增长的规律<sup>[10-12]</sup>。

### 1.4 统计学分析

采用文献计量学方法进行分析,利用VOSviewer 1.6.17软件对高频关键词进行核心作者发文量分析和热点分析,生成可视化网络图<sup>[13]</sup>。对于核心作者关系图,设定作者最少发文数为2,并计算与其他作者之间合作连接的总强度,选择总连接强度最大的作者,颜色范围从蓝色至黄色,

作者之间的合作程度越紧密则颜色越接近黄色。在关键词关系网络图中，每一个节点代表一个关键词，节点的圆圈直径越大则关键词出现的频次越高，不同关键词之间连接线越粗，表示两者之间的关系越紧密。通过 Bibliometric<sup>[14]</sup> 软件分析文献高频词，推测该领域研究热点，并基于所有文献分布的国家、作者，以及主要关键词构建交互式的桑基图，包含权重关系流向，线条的粗细代表关联的紧密程度，线条的颜色代表不同的流量大小。

## 2 结果

### 2.1 一般情况

检索获得相关文献 717 篇，其中 SCIE 362 篇、PubMed 355 篇，根据纳排标准剔除不符合条件及重复发表文献，最终纳入文献 440 篇。文献类型以学术论文为主，共 414 篇（94.09%），其余为综述（17 篇，3.86%）、会议摘要（8 篇，1.82%）、社评（1 篇，0.23%），提示学术论文是胸腔 X 射线影像深度学习研究成果的主要产出形式。

### 2.2 文献年载文量分布

对纳入文献发表年份和文献累计量分布情况进行分析，结果显示 2017—2021 年发文量呈逐年增长的趋势，2020—2021 年论文数量增长幅度最大，近 5 年文献累积总量随年份呈指数增长，反映了 SCIE 和 PubMed 收录的与胸腔 X 射线影像有关的深度学习文献的总体发展趋势良好（图 1）。

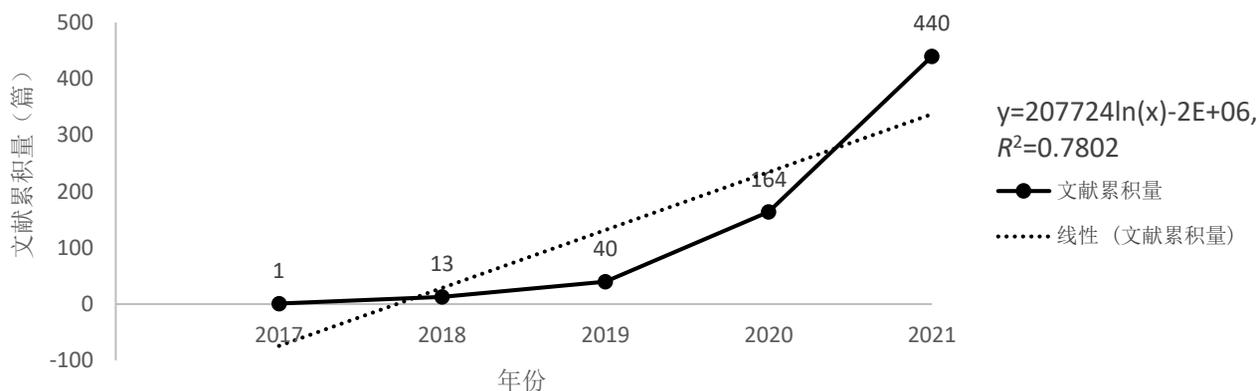


图1 年累积文献发文量趋势

Figure 1. The trend of annual cumulative publications

### 2.3 文献来源国家/地区分布

纳入文献共涉及 47 个国家和地区，通过 Bibliometric 软件获得文献发表总量排名前 15 的国家 / 地区分布图（图 2）。美国文献总量排名第一，占比约 14.77%；中国位列其后，占比约 14.09%；印度排名第三，占比约 10.00%。大部分都是与本国作者合作论文为主。

### 2.4 期刊载文量分布

图 3 展示了主要出版商的发文情况。德国的 Springer Nature 排名第一（83 篇），其收录了大量关于 SARS-CoV-2 与 COVID-19 在胸腔 X 射线上深度学习文章；其次为荷兰的 Elsevier（74 篇）、美国的 IEEE（48 篇）。发文量排名前十的期刊主要来自美国（4 家）和英国（3 家），对于深度学习的研究在国际上具有一定的贡献力和领先优势，发文量最多的期刊是美国的 IEEE Access，共 29 篇，期刊排名情况详见表 1。

### 2.5 文献作者分布

纳入文献的主要作者共 2 438 位，表 2 展示了发文量最多的 7 位作者的相关信息。发文量最大的是美国的 Rajaraman S（10 篇），被引频次最多的是美国的 Duong Tim（8 848 次），来自中国的作者 Wang J 发文 5 篇、被引频次为 455。根据普赖斯定律计算得出核心作者最少发文量为 2.37 篇，即发表 2 篇及以上论文的作者入选核心作者候选人。核心作者发文总量相对较少，尚未出现引文量和发文量丰富的领军人物<sup>[15]</sup>，详见图 4。

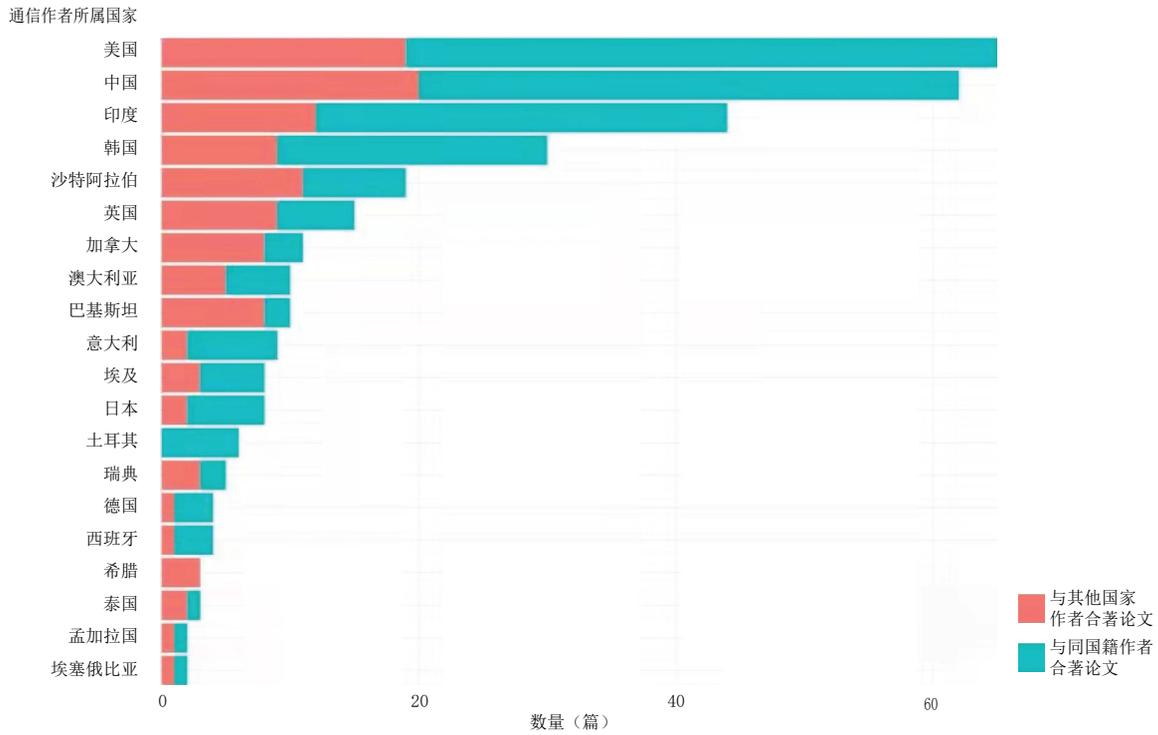


图2 纳入文献的国家/地区分布

Figure 2. The country/district distribution of included publications

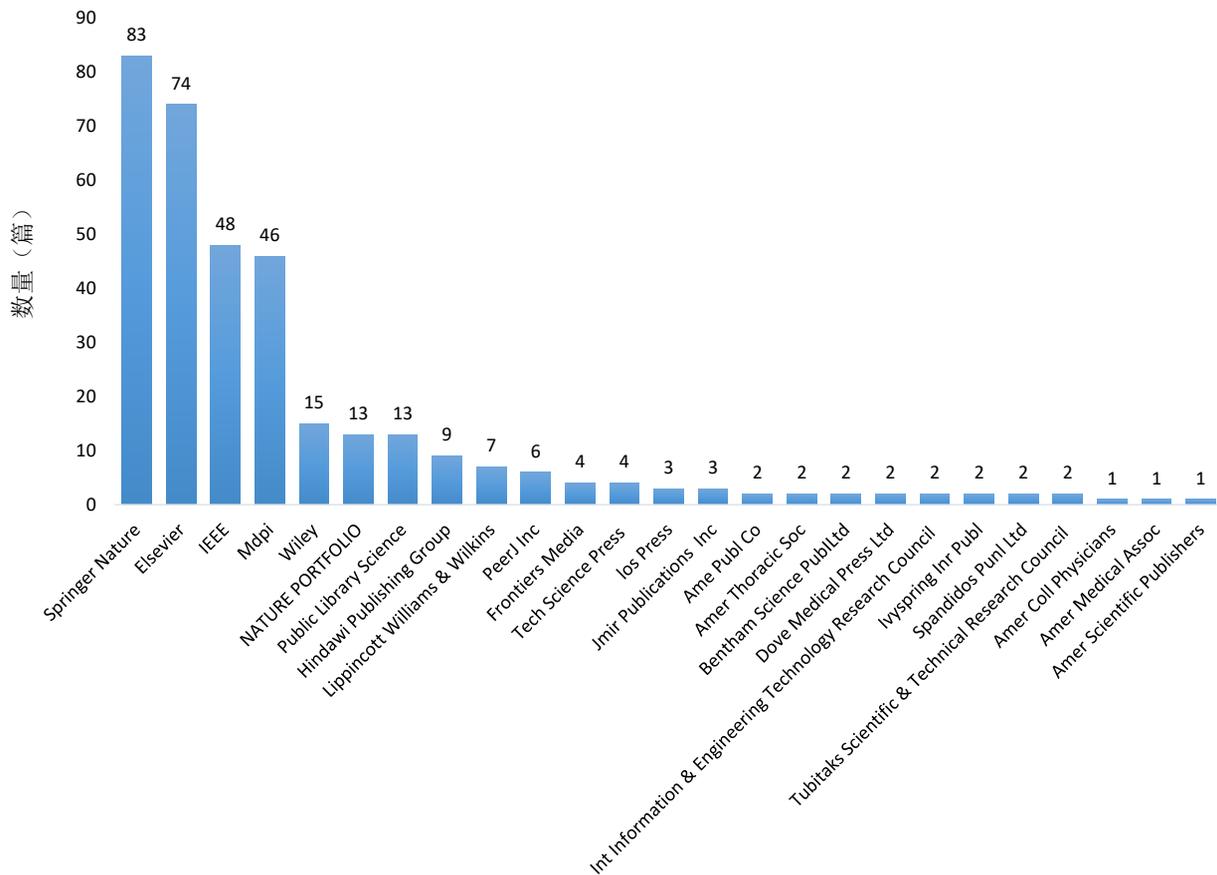


图3 出版商发文量分布

Figure 3. The publications distribution of related publishers

表1 期刊发文量前十统计  
Table1. Top ten statistics journals

刊名	国家/地区	发文量 (篇)	百分比 (%)
IEEE Access	美国	29	13.06
Scientific Reports	英国	16	7.21
Annual International Conference of the IEEE Engineering in Medicine and Biology Society	美国	16	7.21
Computers in Biology and Medicine	英国	14	6.31
Plos One	美国	13	5.86
Diagnostics	瑞士	13	5.86
Journal of Digital Imaging	德国	9	4.05
Applied Sciences–Basel	瑞士	9	4.05
IEEE Journal of Biomedical and Health Informatics	美国	8	3.60
Applied Intelligence	英国	8	3.60

表2 部分核心作者发文情况  
Table 2. The publication information of core authors

核心作者	国家/地区	发文量 (篇)	被引频次
Rajaraman S	美国	10	161
Antani S	美国	7	3 576
Mittal A	印度	6	35
Sarkar R	印度	6	143
Wang J	中国	5	455
Duong T	美国	5	8 848
Kumar V	印度	5	4

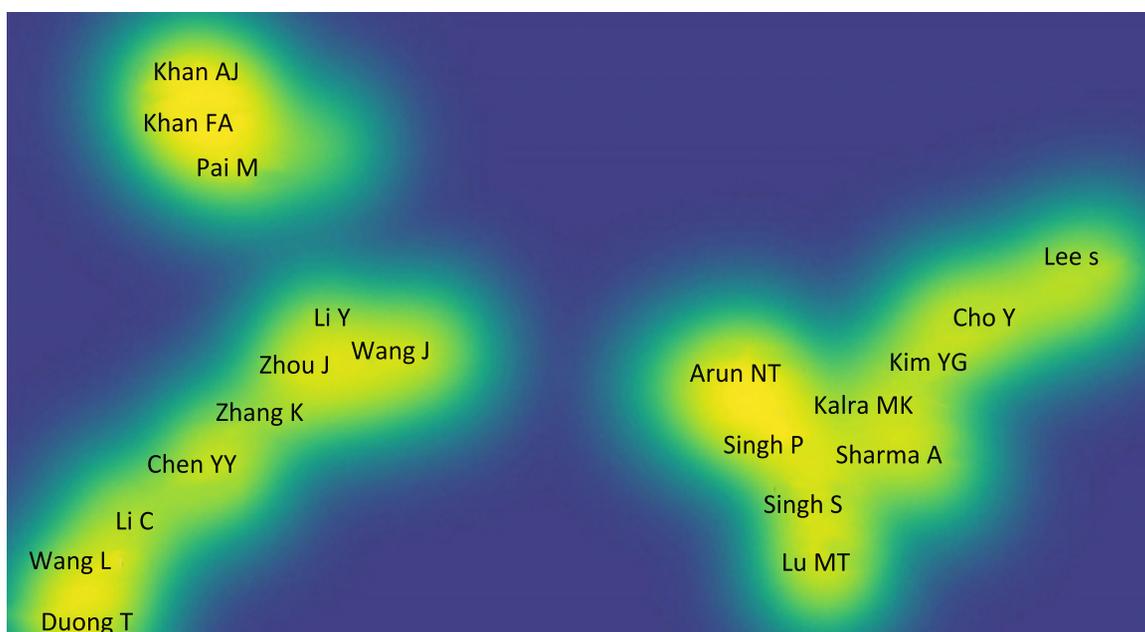


图4 核心作者关系热图  
Figure 4. The heat map of core author relationship



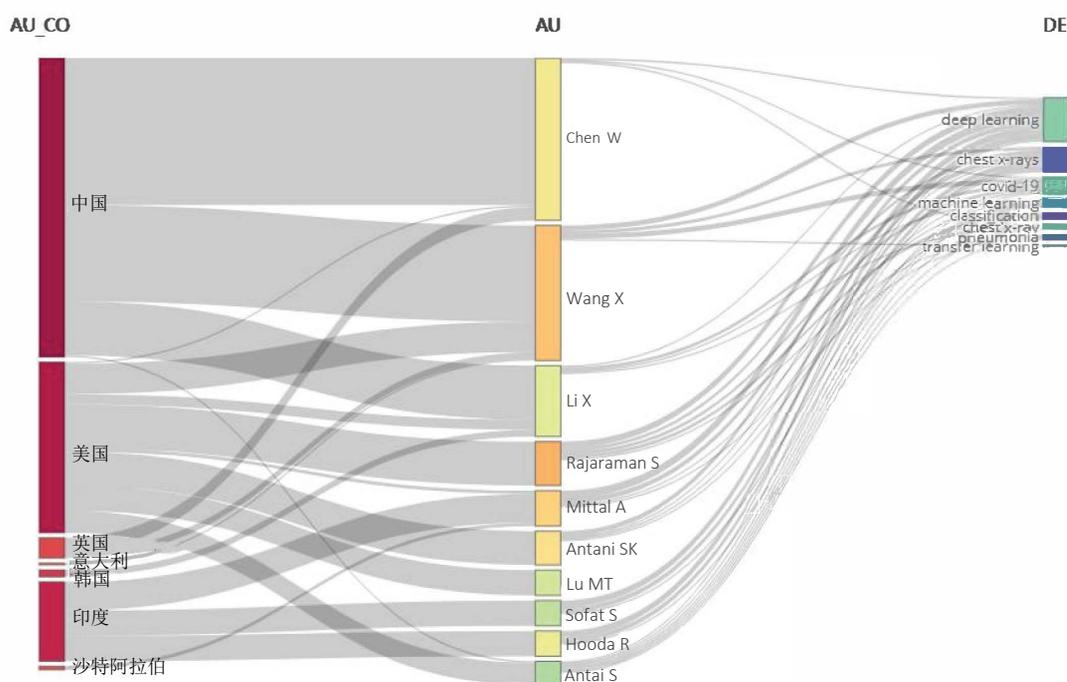


图7 国家、作者及关键词桑基图

Figure 7. Sankey diagram of country, author and key words

### 3 讨论

本文基于 SCIE 和 PubMed 数据库，采用文献计量学分析方法，借助 VOSviewer 和 Bibliometric 分析软件，对近 5 年来胸腔 X 射线影像深度学习相关文献从发文量、作者、出版机构、关键词等方面进行了可视化分析，为进一步的研究提供参考。

既往文献指出学术性文献的增长一般分为四个阶段<sup>[16-17]</sup>，即缓慢增长的初始阶段、指数增长阶段、线性增长阶段和缓慢增长阶段。深度学习作为机器学习的一个热门领域，近年来随着计算机视觉化和临床大数据学科交叉频繁，2017—2021 年间胸腔 X 射线影像的深度学习文献数量呈现逐年增长的趋势，基本达到了指数增长阶段，推测未来该领域发文量将呈现可观的态势。

通过分析纳入文献的国家/地区分布、被引频次、出版期刊与出版商等信息，本研究发现，深度学习在胸腔 X 射线影像的研究仍为近年临床学科交叉学习的新兴领域<sup>[17]</sup>；美国发文量最多，来自英国出版机构的影响力较大，体现了美国学者在基于胸腔 X 射线影像深度学习领域中研究参与度较高。中国论文占统计总文献的 14.09%，反映中国在该研究领域初露头角，具有一定的研究水平，但总体文献影响力仍有待提升。

通过对核心作者关系的统计和分析，可以反映作者在胸腔 X 射线影像深度学习领域的影响力，包括发表的论文数、被引次数和发文量等。本研究结果显示，发文数量、被引频次排名第一的均为美国学者，总体上共同合作发表多篇文献的作者相对较少，可以初步判断该领域尚未形成主要的核心作者群<sup>[12]</sup>。从作者关系热图可以看出，同一国家的作者合作较为紧密，但是国际间作者合作明显较少，尚不利于提升该研究领域的影响力和参与性。基于此，国内学者应积极与其他国家学者合作交流，尤其是在 COVID-19、人工智能等热门领域方面。

文献的关键词凝聚了一篇文献的主题，本研究发现胸腔 X 射线影像深度学习的热点研究方向集中于深度学习、X 射线、COVID-19、分类、图像分割、工程、人工智能等方面。目前，COVID-19 的诊疗仍是临床研究的主流问题，全球相关中英文文献发表量早已累计破万<sup>[18-20]</sup>，而人工智能结合胸腔 X 射线影像诊断为 COVID-19 的研究方向提供了新思路。在深度学习方面，从关键词方面可以看出目前主要基于影像图像的分割和预测分类，包括通过计算机分割影像实现对疾病的诊断和分级，以及对图像进行目标区域的检测和良恶性分类等。目前，基于胸片的深度学

习在 COVID-19 的应用是研究热点, 未来随着深度学习应用于医学图像的不断完善和进步, 深度学习将在更多的胸部疾病(如气胸、肺水肿、胸腔积液、肺癌等)得到广泛应用, 成为临床诊断和治疗中不可或缺的工具。

本研究仍存在一定的局限性。一方面, 仅对 PubMed、SCIE 两个数据库进行了检索, 文献覆盖面可能不全。另一方面, 从研究热度的角度分析, 深度学习在胸片影像的应用相关研究近几年才出现, 检索主题词和自由词的范围受限, 仍待更加完善的检索方式以扩大检索范围提高查全率<sup>[13]</sup>。

### 参考文献

- 侯宇青阳, 全吉成, 王宏伟. 深度学习发展综述[J]. 舰船电子工程, 2017, 37(4): 59-111. [Hou YQY, Quan JC, Wang HW. Review of deep learning development[J]. Ship Electronic Engineering, 2017, 37(4): 59-111.] DOI: [10.3969/j.issn.1672-9730.2017.04.002](https://doi.org/10.3969/j.issn.1672-9730.2017.04.002).
- 高明旭, 李靖, 朱绪平, 等. 深度学习研究方法研究综述[J]. 中国科技信息, 2019, (10): 56-57. [Gao MX, Li J, Zhu XP, et al. A survey of research on deep learning methods[J]. China Science and Technology Information, 2019, (10): 56-57.] DOI: [10.3969/j.issn.1001-8972.2019.10.018](https://doi.org/10.3969/j.issn.1001-8972.2019.10.018).
- 李莉, 黄韬, 王新宇, 等. 胸腔 X 射线影像数据库 -MIMIC-CXR 数据探索[J]. 中国循证心血管医学杂志, 2021, 13(6): 653-656, 660. [Li L, Huang T, Wang XY, et al. Thoracic X-ray image Database-MIMIC-CXR data exploration[J]. Chinese Journal of Evidence-Based Cardiovascular Medicine, 2021, 13(6): 653-656, 660.] DOI: [10.3969/j.issn.1674-4055.2021.06.04](https://doi.org/10.3969/j.issn.1674-4055.2021.06.04).
- 刘栋, 李素, 曹志冬. 深度学习及其在图像物体分类与检测中的应用综述[J]. 计算机科学, 2016, (12): 20-30. [Liu D, Li S, Cao ZD. A review of deep learning and its applications in image object classification and detection[J]. Computer Science, 2016, (12): 20-30.] DOI: [CNKI:SUN:JSJA.0.2016-12-004](https://doi.org/CNKI:SUN:JSJA.0.2016-12-004).
- 陈莹. 迁移学习: 教 AI 提取抽象知识[N]. 科技日报, 2018-1-8(8). [Chen Y. Transfer learning: teaching AI to extract abstract knowledge[N]. Science and Technology Daily, 2018-1-8(8).]
- 张诗乐. 基于 ESI 和 InCites 数据库对我国科研论文产出力和学术影响力的统计分析[D]. 河南: 新乡医学院, 2015. [Zhang SY. The statistical analysis of outputs and academic impact about scientific papers in China based on ESI and InCites databases[D]. Henan: Xinxiang Medical University, 2015.]
- 康国光, 沈振锋, 徐跃进, 等. 学生满意度研究: 现状、演进路径及前沿——基于 Web of Science 数据库[J]. 现代情报, 2014, 34(8): 29-36, 41. [Kang GG, Shen ZF, Xu YJ, et al. The analysis of evolution pathway, research hotspots and research frontiers on student satisfaction——based on 1511 articles from Web of Science[J]. Modern Information, 2014, 34(8): 29-36, 41.] DOI: [10.3969/j.issn.1008-0821.2014.08.005](https://doi.org/10.3969/j.issn.1008-0821.2014.08.005).
- 董文军. 基于 Web of Science 及 ESI 的学科数据统计分析[J]. 情报杂志, 2009, 28(z1): 27-31. [Dong WJ. Statistical analysis of subject data based on Web of Science and ESI[J]. Journal of Intelligence, 2009, 28(z1): 27-31.] DOI: [CNKI:SUN:QBZZ.0.2009-S1-008](https://doi.org/CNKI:SUN:QBZZ.0.2009-S1-008).
- 林芸峰, 田玲, 张宏梁, 等. PubMed 数据库 1998-2008 年抗抑郁药文献剂量学分析[J]. 中国药房, 2010, 21(1): 90-92. [Lin YF, Tian L, Zhang HL, et al. Bibliometric analysis on antidepressive agents in PubMed database during 1998~2008[J]. China Pharmacy, 2010, 21(1): 90-92.] DOI: [CNKI:SUN:ZGYA.0.2010-01-044](https://doi.org/CNKI:SUN:ZGYA.0.2010-01-044).
- 万晓霞. 近 10 年 SCI 人格心理学研究文献计量分析[J]. 心理科学进展, 2009, 17(6): 1281-1286. [Wan XX. Research of personality psychology bibliometric analysis based on the database of SCI in recent ten years[J]. Advances in Psychological Science, 2009, 17(6): 1281-1286.] DOI: [CNKI:SUN:XLXD.0.2009-06-024](https://doi.org/CNKI:SUN:XLXD.0.2009-06-024).
- 胡臻, 张阳. 基于普赖斯定律与综合指数法的核心作者和扩展核心作者分析——以《西南民族大学学报》(自然科学版)为例[J]. 西南民族大学学报(自然科学版), 2016, 42(3): 351-354. [Hu Z, Zhang Y. Analysis of core authors and extended core authors based on price law and comprehensive index method——Take Journal of Southwest University for Nationalities (Natural Science Edition) for example[J]. Journal of Southwest University for Nationalities (Natural Science Edition), 2016, 42(3): 351-354.] DOI: [10.11920/xnmdzk.2016.03.019](https://doi.org/10.11920/xnmdzk.2016.03.019).
- 郑彩琴, 陈小清, 黄笑云, 等. 基于 PubMed、SCIE 数据库国内医用红外热成像相关研究的文献计量学分析[J]. 中国医学影像学杂志, 2021, 29(7): 751-755. [Zheng CQ, Chen XQ, Huang XY, et al. Bibliometric

- analysis of medical infrared thermal imaging in China based on PubMed and SCIE databases[J]. Chinese Journal of Medical Imaging, 2021, 29(7): 751-755.] DOI: [10.3969/j.issn.1005-5185.2021.07.023](https://doi.org/10.3969/j.issn.1005-5185.2021.07.023).
- 13 王贝,刘纯青.基于 Citespace 与 VOSviewer 的国内生态网络研究[J].环境科学与管理,2021,46(4):53-58. [Wang B, Liu CQ. Review on domestic ecological network research based on bibliometrics[J]. Environmental Science and Management, 2021, 46(4): 53-58.] DOI: [10.3969/j.issn.1673-1212.2021.04.016](https://doi.org/10.3969/j.issn.1673-1212.2021.04.016).
- 14 Tian DQ. Bibliometric analysis of pathogenic organisms[J]. Biosafety and Health, 2020, 2(2): 95-103. DOI: [10.1016/j.bsheal.2020.05.004](https://doi.org/10.1016/j.bsheal.2020.05.004).
- 15 梁立锋,曾紫云,邹玉如,等.基于中国知网数据库的深度学习文献计量分析[J].岭南师范学院学报,2020,41(2):118-123. [Liang LF, Zeng ZY, Zou YR, et al. Bibliometric analyses of deep learning research based on CNKI[J]. Journal of Lingnan Normal University, 2020, 41(2): 118-123.] DOI: [10.3969/j.issn.1006-4702.2020.02.015](https://doi.org/10.3969/j.issn.1006-4702.2020.02.015).
- 16 石建,石苗茜.基于 SCI 及 ESI 的脑膜炎研究十年发展态势的文献计量分析[J].科学技术与工程,2010,10(30):7396-7401. [Shi J, Shi MQ. 10 year bibliometrics quantitative analysis of development situation for meningitis study based on SCI and ESI[J]. Science Technology and Engineering, 2010, 10(30): 7396-7401.] DOI: [10.3969/j.issn.1671-1815.2010.30.005](https://doi.org/10.3969/j.issn.1671-1815.2010.30.005).
- 17 邱均平.信息计量学[M].武汉:武汉大学出版社,2007. [Qiu JP. Informatics[M]. Wuhan: Wuhan University Press, 2007.]
- 18 王高玉,刘红宁,赵益.基于 SCIE 中医药研究十年发展态势文献计量分析[J].江西中医药大学学报,2016,28(6):111-115. [Wang GY, Liu HN, Zhao Y. Research on the development of Chinese medicine based on SCIE for ten years[J]. Journal of Jiangxi University of Traditional Chinese Medicine, 2016, 28(6): 111-115.] DOI: [CNKI:SUN:XYXB.0.2016-06-037](https://doi.org/CNKI:SUN:XYXB.0.2016-06-037).
- 19 张泽华,郭姗姗,赵志刚,等.基于 CiteSpace 的新冠肺炎研究文献计量分析[J].中国医院药学杂志,2020,40(19):2029-2034. [Zhang ZH, Guo SS, Zhao ZG, et al. Bibliometric analysis of COVID-19 based on CiteSpace[J]. Chinese Journal of Hospital Pharmacy, 2020, 40(19): 2029-2034.] DOI: [10.13286/j.1001-5213.2020.19.05](https://doi.org/10.13286/j.1001-5213.2020.19.05).
- 20 张欣桐,刘景卓,李盼,等.重症新型冠状病毒肺炎的文献计量和可视化分析[J].中国急救医学,2021,41(4):335-340. [Zhang XT, Liu JZ, Li P, et al. The bibliometrics and visualization analysis in the assessment of severe coronavirus disease 2019[J]. Chinese Journal of Critical Care Medicine, 2021, 41(4): 335-340.] DOI: [10.3969/j.issn.1002-1949.2021.04.012](https://doi.org/10.3969/j.issn.1002-1949.2021.04.012).

收稿日期:2022 年 01 月 26 日 修回日期:2022 年 02 月 21 日  
本文编辑:李阳 曹越

引用本文:黄夏璇,陈咏梅,袁师其,等.深度学习运用于胸腔 X 射线影像研究的文献计量学分析[J].医学新知,2023,33(2):91-99. DOI: [10.12173/j.issn.1004-5511.202201031](https://doi.org/10.12173/j.issn.1004-5511.202201031)  
Huang XX, Chen YM, Yuan SQ, et al. Bibliometric analysis of deep learning in chest X-ray imaging research[J]. Yixue Xinzhi Zazhi, 2023, 33(2): 91-99. DOI: [10.12173/j.issn.1004-5511.202201031](https://doi.org/10.12173/j.issn.1004-5511.202201031)