ADBA
Computers and Electronics in Medicine
Cutting-Edge Scientific Solutions

e-ISSN: 3023-8609
RESEARCH ARTICLE
Vol.1 / No.1 / 2024 / pp.24-33
https://doi.org/10.69882/adba.cem.2024074

Bibliometric Analysis of Publications on Action Recognition, Convolutional Neural Network, Video Surveillance During 2012-2021

Meryem Yalçınkaya [*,1] **and Ömer Faruk Akmeşe** [β,2]

*Hitit University, Faculty of Engineering, Department of Industrial Engineering, 19030, Corum, Turkiye, βHitit University, Faculty of Engineering, Department of Computer Engineering, 19030, Corum, Turkiye.

ABSTRACT Action recognition based on convolutional neural networks (AR-CNN) has been developing rapidly in recent years. It is of great significance to conduct a deep analysis to understand the recent development of AR-CNN. However, a limited number of studies examining the research status of this field could be found. Therefore, this study aims to quantitatively assess the publications related to the SciVal topic "Action Recognition; Convolutional Neural Network; Video Surveillance (T.561)" in computer vision research. This study focused on six aspects: literature distribution characteristics analysis, the development trend, citation analysis, collaborative analysis, keyword analysis, and thematic evolution, using VOSviewer and Bibliometrix. The relevant publications were retrieved from Scopus in the period 2012–2021. A total of 6633 publications were identified by 9088 different authors; 62% were conference papers, and 35% were research articles. China and the USA contributed 39.7% and 17.9% of the total publications, respectively. The authors' productivity demonstrated variability in alignment with Price's Law, yet exhibited consistency when evaluated under the framework of Lotka's Law. Ling Shao was the most productive author, with 48 papers (0.7%). Chinese Academy of Sciences was the most productive affiliation, with 259 papers (3.9%). The first Bradford site consisted of Computer Science Lecture Notes with 617 publications. A moderately significant correlation was revealed between the country's publications and GDP per capita. The overall results show that the number of AR-CNN-related documents has increased significantly in recent years, with rapid growth from 2016. Although publications on AR-CNN were published mainly in European journals, China led the scientific production.

KEYWORDS
Action recognition
Deep learning
Convolutional neural network
Video surveillance
Artificial intelligence
Bibliometric analysis

INTRODUCTION

Action recognition (AR) is used to decipher an action/activity component from a video or image scene. It is a fundamental and challenging topic in computer vision (Ahad 2011). It has drawn the attention of many researchers due to its varying applications, such as security systems, medical systems, virtual reality, autonomous vehicles, ambient intelligence, human behavior analysis, robotics, human-computer interaction, etc. (Khan *et al.* 2020; Sudha *et al.* 2017; Xia *et al.* 2015; Oh *et al.* 2011; Popoola and Wang 2012). At its very early stage, action recognition was used to categorize human actions, security reasons, and surveillance systems (Oh *et al.* 2011; Bobick and Davis 1996; Rosales and Sclaroff 2001; Karpathy *et al.* 2014). Later, with the breakthrough and fast development of deep learning technology, action recognition has advanced considerably. Much research was conducted based on deep learning to recognize human actions in videos (Simonyan and Zisserman 2014; Donahue *et al.* 2015; Zhou *et al.* 2018). Since videos are 3D Spatiotemporal signals, the main idea behind the majority of these studies is to extend Convolutional Neural Networks (CNNs) to include the temporal information contained in videos. CNN is a deep model that ob-

Manuscript received: 24 April 2024,
Revised: 15 May 2024,
Accepted: 15 May 2024.

[1]meryemyalcinkaya@hitit.edu.tr
[2]ofarukakmese@hitit.edu.tr (**Corresponding author**)

tains complicated hierarchical features via convolutional operation alternating with sub-sampling operation on the raw input images (LeCun *et al.* 1998). Since CNN was adopted in video-based human action recognition (HAR), many extensions of CNN (such as deep neural network (DNN), Recurrent Neural Networks (RNNs), 3D CNNs, etc.) have been proposed to improve both the accuracy and efficiency of human action recognition from videos (Karpathy *et al.* 2014). This increasing interest in action recognition based on CNN (AR-CNN) resulted in numerous research papers. While bibliographic data for AR-CNN-related studies are increasingly available, how to review all research and discover the research trends in AR-CNN based on these bibliographic data represents a challenging research question. Thus, it is necessary to implement bibliometric analytical techniques to evaluate the growing literature on AR-CNN.

Bibliometrics is a useful tool to evaluate and quantify the growth of literature for a particular subject (A. 1969). There are three bibliometric studies in action recognition. Aryanfan etc., examined the characteristics of HAR literature from 1987 to 2015 based on the Web of Science Core Collection (WOS) (Bi *et al.* 2017). Chen and Deng analyzed the evolution of CNNs in many computer vision applications, such as fault and image recognition diagnosis, seismic detection, image classification, etc., by using the bibliometric method in literature from 2011 to 2020 (Chen and Deng 2020). Ci, etc., analyzes the publishing trends, major countries or regions, research topics, and research fronts based on computer vision and graphics papers from 2010 to 2020 (Ci *et al.* 2021). However, these studies remain at the general level for the subject AR-CNN and are limited to the WOS data only. Also, there is no analysis to reveal the research correlation between scientific output and Gross Domestic Product (GDP) per capita, the relationship between journals and the papers they publish, and the productivity of authors. That is why this study is focused on AR-CNN with different quantitative and qualitative parameters such as characteristics of subject literature, productivity, and relationship, international collaboration percentage, etc.

This study uses the Scopus database to conduct a bibliometric analysis of the 2012-2021 AR-CNN research. The objectives of this article are a multi-angle assessment of research productivity and an analysis of the significant publication patterns, research directions, and trends in the field of AR-CNN.

MATERIAL AND METHODS

The Scopus database was used for the bibliometric analysis of "Action Recognition; Convolutional Neural Network; Video Surveillance" (T.561). In order to reach reliable and accurate details on this subject, 6633 publications were obtained on September 26, 2022, for the period 2012-2021. Downloaded information included authors' names, paper titles, publishing years, document type, subject, the Scopus categories of the publication, names of journals, country, institution, and citations for each publication. Since the T.561 topic is only defined in Scival, not Scopus, a download strategy was developed for transferring data from Scival to Scopus using the Selenium Python library. The SciVal is based on output and usage data from Scopus, the world's largest abstract and citation database for peer-reviewed publications. The SciVal uses the Scopus data from 1996 to the current date, which covers 48 million publication records, 22,000+ journals, and 5,000+ publishers (SciVal 2023).

The general research performance of the retrieval literature was processed by Microsoft Excel 365. The VOSviewer software (1.6.18) was utilized to evaluate the co-authorships among countries and institutions and produce a keyword co-occurrence analysis. VOSviewer is a robust tool that uses clustering algorithms and functionalities based on the strength of the connections among items to facilitate network analyses (**?**). Also, Bibliometrix and its user interface, Biblioshiny, were used to generate a bibliometric map. Bibliometrix and R Shiny platform incorporate various analyses, such as overview, conceptual structures, intellectual structure, etc. (Aria and Cuccurullo 2017).

- Literature distribution characteristics, including categories, productive publication sources, authors, affiliations, countries, citations, and publication trends, were analyzed using statistical methods. Simple linear regression analysis was used to estimate publication numbers for the following years.
- The network analysis method was used to visualize scientific collaborations among authors, affiliations, and countries.
- Price's Law analysis, Bradford's Law, and Lotka's Law were used to explain scientific productivity and the relationship between authors and the quantities of their papers. The correlations between the Gross Domestic Product (GDP) per capita and publication productivity of countries on the topic T.561 were investigated with Spearman's rank correlation coefficient.
- Keyword analysis by the author Keyword analysis was conducted to provide important information about research trends that concern researchers. For deeper analysis, the thematic evolution was compared between two time periods, i.e., 2012-2018 and 2019-2021; it provides us with a global view of the changes.

Other databases, such as the WOS and Google Scholar, could also be considered for collecting bibliometric information, but Scopus has significant advantages. Scopus includes most journals indexed in the Web of Science (WOS) across various disciplines. As 99% of journals indexed in the WOS overlap with the Scopus, about 34% of journals indexed in the Scopus overlap with the WOS. This means that Scopus covers about 66% of its journals exclusively compared to WOS. In addition, in Natural Sciences and Engineering (NSE), the Scopus covers 38% of all Ulrich's journals in this field, while the WOS covers 33% (Mongeon and Paul-Hus 2016). Yang & Meho (2006) observed that the WOS should not be used alone for locating citations to an author or title and that Scopus and Google Scholar can help identify many valuable citations not found in the WOS. The WOS and the Scopus journal coverage results in variations in research output volume, rank, and global share in different countries. China and India are two significant exceptions, where research output volume in the Scopus is significantly higher than in the WOS. Considering its largest curated databases covering scientific journals, books, conference proceedings, etc., Scopus is a more convenient choice for a more detailed and comprehensive overview of the bibliographic data source (Kiduk and Meho 2006).

RESULTS

The distribution of subject categories

Throughout 2012–2021, 6633 publications of different types appeared: articles (2341, 35.3%), conference papers (4124, 62.2%), chapters (75, 1.1%), conference reviews (43, 0.65%), reviews (32, 0.48%), and others (18, 0.27%).

The Scival database covered 27 top-level and 334 lower-level subject areas based on the Scopus All Subject Journal Classification (ASJC). As shown in Figure 1, most of the publications related to AR-CNN focused on three top-level categories, including "Computer Science" (5981, 90.2%), "Engineering" (2291, 34.6%), and

"Mathematics" (1324, 20%). The first lower level is "Computer Vision and Pattern Recognition" (2284, 34.4%) for the "Computer Science" category, "Electrical and Electronic Engineering" (1118, 16.9%) for "Engineering" and "Theoretical Computer Science" (680, 10.3%) for "Mathematics". The results followed by Instrumentation (192, 2.9%) for Physics and Astronomy, General Materials Science (192, 2.7%) for Material Science, and Information Systems and Management (204, 3.1%) for Decision Sciences. It is noted that since a paper can be mapped to different categories, the total percentage in Figure 1 is more than 100%.

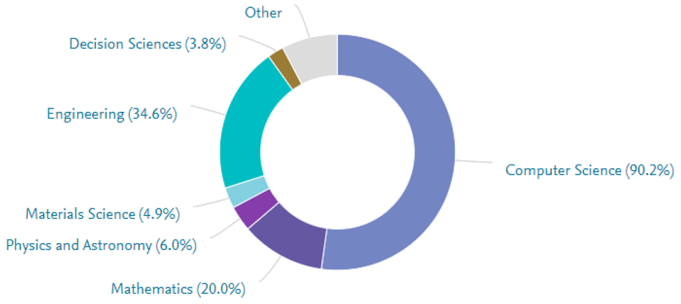

Figure 1 The distribution of subject categories

Development of publications

Figure 2 shows the trend of cumulative and annual publications related to AR-CNN research from 2012 to 2021 and the predictions for the next four years. An increasing trend was observed over the years, approximated by the linear growth curve following the equation $y = 85.85x + 295$ (where x is the number of points per year and y is the number of publications per year). This means that between 2012 and 2021, the data tended to be linear, with a linear annual growth rate of 86 publications per year. Based on the linear model, it could be estimated that 1240 (95% confidence interval; 1085-1395) publications could be published in 2022, while in 2025, this number is expected to increase to 1497 (1209-1785).

Since 2012, the number of publications has been observed to increase despite some fluctuations, revealing that research on AR-CNN has been a growing passion of researchers. More than half of the publications were published after 2018. One of the key reasons for this increase might be related to the improvements in deep learning algorithms, computer vision cloud services, and mobile devices.

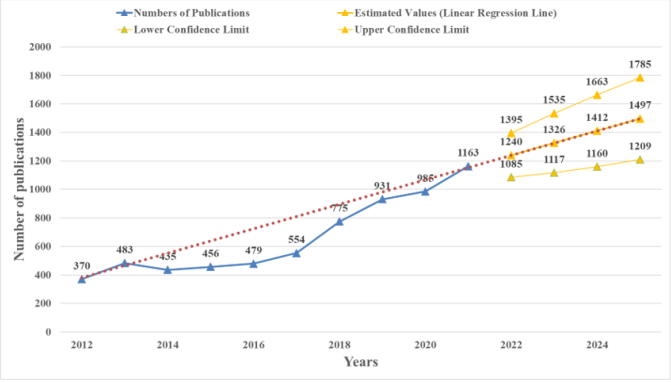

Figure 2 The number of publications by year on AR-CNN

Citations

Overall, 6633 publications received 130,616 citations with 1.76 field-weighted citation impact (FWCI). Table 1 shows the percentage of publications for each threshold based on the number of citations received. Only 6.9% (456) papers have received more than 50 citations, 8% (532) papers received more than or equal to 20 citations, and the remaining papers received less than 20 citations. However, the FWCI of AR-CNN is more than 1. The output is more cited than expected according to the global average.

Active Authors

A total of 6633 publications were produced by 9088 authors. Of these, 4183 authors published articles, some co-authors, and 6917 authors published conference papers. The top productive five authors in terms of citations per publication (CPP) were Andrew P. Zisserman (20 papers, 502 CPP), Du Tran (20 papers, 439.9 CPP), Rahul Sukthankar (15 papers, 380.2 CPP), Lorenzo Torresani (20 papers, 334.4 CPP) and Cordelia Schmid (33 papers, 234.1 CPP). The top 5 authors producing the highest number of publications were Ling Shao (48 papers, 0.72%), Limin Wang (39 papers, 0.59%), Cordelia Schmid (33 papers, 0.5%), Yu Qiao (32 papers, 0.48%), and Anastasios Tefas (32 papers, 0.48%).

Active Institutions

According to the results, 5208 organizations contributed to the analyzed publications. The top 10 organizations that contributed most to the literature were: the Chinese Academy of Sciences (259, 3.9%), CNRS (118, 1.78%), Shanghai Jiao Tong University (118, 1.78%), University of Chinese Academy of Sciences (109, 1.64%), Sun Yat-Sen University (96, 1.45%), Peking University (94, 1.42%), Beijing Institute of Technology (91, 1.37%), CAS - Institute of Automation (91, 1.37%), University of Electronic Science and Technology of China (80, 1.21%), and Institut National de Recherche en Informatique et en Automatique (77, 1.16%).

The top productive ten organizations in terms of CPP are Institut National de Recherche en Informatique et en Automatique (77 papers, 125.6 CPP), Chinese Academy of Sciences (259, 36.6 CPP), CAS - Institute of Automation (90, 32.7 CPP), CNRS (118, 26.9 CPP), University of Electronic Science and Technology of China (80, 15.6 CPP), University of Chinese Academy of Sciences (109, 15.5 CPP), Shanghai Jiao Tong University (118, 14.9 CPP), Peking University (94, 14 CPP), Beijing Institute of Technology (91, 10.7 CPP) and Sun Yat-Sen University (96, 10 CPP). Figure 3 shows the number of publications of the institution over time from 2012 to 2021, and there is an increasing trend, with slight fluctuations in individual years. It was noted that the outputs of government and academic institutions named the Chinese Academy of Sciences fluctuated with high figures in 2018.

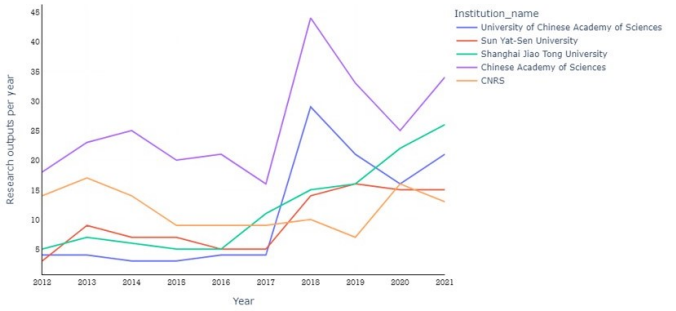

Figure 3 Affiliation production over time

Table 1 General citation structure

Number of citations	Number of publications	Publications %
$\geq$ 50 Citations	456	6.9
$\geq$ 20 Citations	532	8.0
< 20 Citations	5645	85.1

Active Sources

Six thousand six hundred thirty-three publications were published in a wide range of 488 journals, 1055 conference proceedings, 30 books, 23 book series, and one trade publication. The top 22 sources with the highest number of publications are shown in Table 2, and more than 40% of 6633 publications are from these 22 sources.

Lecture Notes in Computer Science (book series) was the most productive source, followed by three conference proceedings, as shown in Table 2. IEEE Access (journal) and Multimedia Tools and Applications (journal) ranked 5th and 6th most productive sources, respectively.

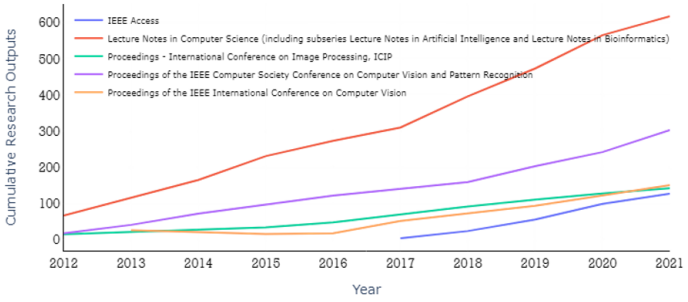

Figure 4 Top 5 journals with the highest number of publications

In Fig. 4, the trends of publication quantities of the top 5 productive journals during ten years were revealed. While the number of publications of the sources was close to each other in 2012, a divergence occurred over time. Especially since 2017, the rate of the annual increase of “Lecture Notes in Computer Science” has accelerated more than other sources, and the number of publications in “Lecture Notes in Computer Science” is almost equal to the total number of publications in the other four sources.

Top Cited Authors

Table 3 shows the top 15 most cited authors, with a minimum of 1500 papers.

Top 10 Publications by Number of Citations

Table 4 provides a detailed overview of the top 10 publications by the number of citations in the field of action recognition. The leading paper, "Learning Spatiotemporal Features with 3D Convolutional Networks" by Tran et al. (2015), has garnered a total of 4721 citations, averaging 590.13 citations per year and achieving a normalized total citation count of 133.77. This is followed closely by the work of Simonyan et al. (2014) on "Two-Stream Convolutional Networks for Action Recognition in Videos," which has received 4423 citations with an annual citation rate of 491.44 and a normalized citation count of 114.66. Notably, Karpathy et al.’s (2014) research on "Large-Scale Video Classification with Convolutional Neural Networks" has also made significant impact, with 4349 citations and a normalized total of 112.75. Other influential works include Ji et al.’s (2013) "3D Convolutional Neural Networks for Human Action Recognition" and Carreira et al.’s (2017) "Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset," reflecting their substantial contributions to the field with normalized citation counts of 103.77 and 85.38, respectively. The table also highlights the works of Heng Wang et al., Feichtenhofer et al., and Yue-Hei Ng et al., illustrating the breadth and depth of research efforts in this domain. Overall, these publications represent critical advancements in the development and application of convolutional neural networks for action recognition, demonstrating significant academic influence and ongoing relevance in the research community.

COLLABORATION ANALYSIS

Active Countries

According to the retrieved results, the papers covered a total of 86 different countries (or territories). The network visualization map of 15 countries producing at least 35 publications is shown in Figure 5b. China ranked first with a dominant output of 2638 papers or a share of 39.7%. The USA had 1184 papers (17.9%), and India had 523 (7.9%), ranking second and third, respectively. Other top-ranked countries are the UK (6.1%), Japan (4.0%), France (3.7%), South Korea (3.5%), and Australia (3.5%).

The international cooperation analysis studied a network of the leading countries, plotted in Figure 5b. China showed 13.42 average citations and a total link strength of 753; the USA showed 46.56 average citations and a link strength of 600; the UK showed 42.63 average citations and a total link strength of 365; Australia showed 26.19 average citations with a link strength of 194, Singapore showed 20.48 average citations with a link strength of 156, France showed 46.12 average citations with a link strength of 150, Spain showed 16.24 average citations with a link strength of 136 and Pakistan showed 16.92 average citations with a link strength of 120.

The strongest link strength was evidenced by the USA and China, with a 260 link strength, followed by the UK and China with a 105 link strength, Australia and China with a 71 link strength, HongKong and China with a 71 link strength, the China and Singapore with a 68 link strength, the UK and the USA with a 41 link strength, and the UK and Spain with a 39 link strength.

Figure 6 shows the growth trends of publications for the top 5 productive countries from 2012 to 2021. Compared to the four countries, the trend in China has increased at a quicker pace after 2017. China’s linear annual growth rate reached 16 publications per year from 2012 to 2018, up to 146 publications per year in 2018, and then 62 publications per year from 2019 to 2021. This marked increase in 2018 may be associated with China’s efforts to

Table 2 Top 22 journals with the highest number of publications

No	Journals	Publication number	%	CiteScore 2021
1	Lecture Notes in Computer Science	617	9.3	2.1
2	Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition	325	4.9	44.6
3	Proceedings of the IEEE International Conference on Computer Vision	177	2.67	14.1
4	Proceedings - International Conference on Image Processing, ICIP	142	2.14	-
5	IEEE Access	127	1.91	6.7
6	Multimedia Tools and Applications	124	1.87	5.3
7	Proceedings - International Conference on Pattern Recognition	101	1.52	2.1
8	ACM International Conference Proceeding Series	96	1.45	1.0
9	IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops	96	1.45	-
10	Neurocomputing	92	1.39	10.3
11	IEEE Transactions on Image Processing	81	1.22	16.4
12	Proceedings of SPIE - The International Society for Optical Engineering	78	1.18	0.9
13	Communications in Computer and Information Science	76	1.15	0.9
14	Advances in Intelligent Systems and Computing	75	1.13	-
15	Pattern Recognition	73	1.1	15.5
16	IEEE Transactions on Circuits and Systems for Video Technology	67	1.01	10.1
17	IEEE Transactions on Pattern Analysis and Machine Intelligence	61	0.92	36.6
18	Pattern Recognition Letters	60	0.9	8.6
19	Proceedings - ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing	58	0.87	5.8
20	Computer Vision and Image Understanding	56	0.84	9.9
21	International Journal of Computer Vision	51	0.77	16.8
22	Proceedings - IEEE International Conference on Multimedia and Expo	51	0.77	-

CiteScore is the ratio of citations of the document received in the current year to the total citations in the previous three years.

implement machine learning interventions in real-world settings, the general increased awareness of action recognition, and the effects of Project 985 (Xuefei 2014). The publications in the US tended to be linear, with an annual growth rate of 17 publications per year since 2016. The publications in India remained constant at around 85 per year during 2018-2020 and then increased at an annual growth rate of 30 publications per year. The growth trends of Japan and the United Kingdom remained constant during the last decade, with a low increase rate.

Figure 7 shows the distribution of publications in SCP (Singular country publications) and MCP (Multiple country publications) status by country. When viewed proportionally, it can be said that England, Australia, Pakistan, and Singapore published more than multiple countries.

General Patterns of Collaboration

Table 5 shows collaboration metrics by assigning publications as international, national, institutional, and single-authorship based on T.561 data. International collaboration measures the proportion of publications with at least one international co-author. National collaboration measures the proportion of publications carried out by several authors/institutions from a single country. Institutional

Table 3 Top 15 most cited authors (2012-2021) with characteristics of publication outputs

No	Name	Citations	Number of publications	Citations per Publication	Field-Weighted Citation Impact	h-index
1	Zisserman, Andrew P.	9863	20	493.2	24.93	122
2	Schmid, Cordelia	7650	33	231.8	12.77	102
3	Torresani, Lorenzo	6526	20	326.3	15.13	36
4	Tran, Du	6437	15	429.1	19.4	17
5	Sukthankar, Rahul	5648	15	376.5	19.58	48
6	Wang, Limin	5368	39	137.6	11.68	29
7	Qiao, Yu	4861	32	151.9	11.87	52
8	Niebles, Juan Carlos	2554	25	102.2	6.08	33
9	van Gool, Luc J.	2242	18	124.6	10.83	123
10	Ghanem, Bernard	2076	26	79.8	6.04	40
11	Fei-Fei, Li	1955	18	108.6	7.28	87
12	Laptev, Ivan	1877	17	110.4	7.04	47
13	Shao, Ling	1852	48	38.6	3.05	75
14	Shah, Mubarak A.	1604	23	69.7	4.82	87
15	Fernando, Basura	1575	21	75	5.25	24

Field-Weighted Citation Impact: The ratio of citations received relative to the expected world average for the subject field, publication type, and publication year. The average world FWCI is 1.00

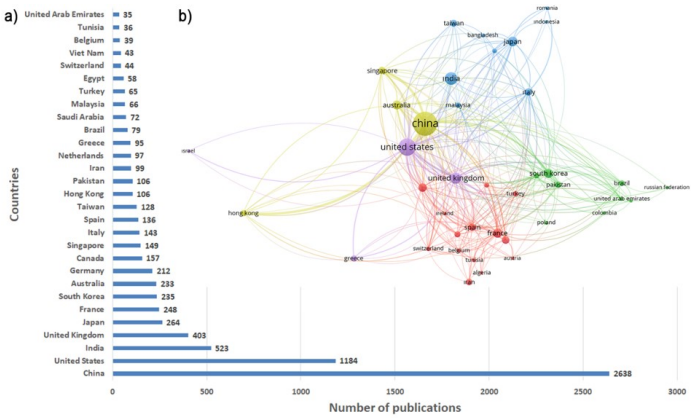

Figure 5 a) Bar chart showing the productivity of the top 15 countries in the world. b) Network visualization map for international cooperation of countries on Action Recognition, Convolutional Neural Networks, and Video Surveillance. Footnote: As the size of the circle increases, the number of publications increases. Clusters are separated by colors. The thickness of the lines expresses the strength of cooperation between countries.

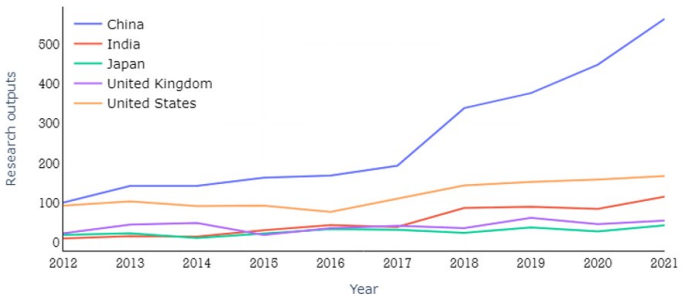

Figure 6 Country Production over Time

collaboration measures the proportion of publications carried out by several authors from the same institution in a single country. Most of the retrieved publications had only institutional collaboration (n=3132; 47.2%), followed by only national collaboration (n = 1909; 28.8%) and international collaboration (n =1385; 20.9%). The rest of the publications belong to the category "single authorship" or "no collaboration" (n=206; 3.1%). Nonetheless, in terms of impact, an international collaboration (31.7 citations/document, 2.64 FWCI) exceeds both national (20.4; 1.79) and institutional collaboration (14.2;1.42). Chinese Academy of Sciences was among the top for coauthorship collaboration between the institutions with the highest number of publications.

Table 4 Top 10 publications by the number of citations

Paper	Author	Year	Total Citations (TC)	TC per Year	Normalized TC
Learning Spatiotemporal Features with 3D Convolutional Networks	Tran et al.	2015	4721	590.13	133.77
Two-Stream Convolutional Networks for Action Recognition in Videos	Simonyan et al.	2014	4423	491.44	114.66
Large-Scale Video Classification with Convolutional Neural Networks	Karpathy et al.	2014	4349	483.22	112.75
3D Convolutional Neural Networks for Human Action Recognition	Ji et al.	2013	3818	381.80	103.77
Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset	Carreira et al.	2017	2910	485.00	85.38
Action Recognition with Improved Trajectories	Heng Wang et al.	2013	2474	247.40	67.24
Convolutional Two-Stream Network Fusion for Video Action Recognition	Feichtenhofer et al.	2016	1815	259.29	57.26
Beyond short snippets: Deep networks for video classification	Yue-Hei Ng et al.	2015	1513	189.13	42.87
Temporal Segment Networks: Towards Good Practices for Deep Action Recognition	Limin Wang et al.	2016	1337	191.00	42.18
Dense Trajectories and Motion Boundary Descriptors for Action Recognition	Heng Wang et al.	2013	1326	132.60	36.04

Field-Weighted Citation Impact: The ratio of citations received relative to the expected world average for the subject field, publication type, and publication year. The average world FWCI is 1.00

Table 5 Publications by the amount of international, national, and institutional collaboration

Metric	%Percentage	Number of publications	Citations	Citations per Publication	Field-Weighted Citation Impact
International collaboration	20.90%	1,385	43,880	31.7	2.64
Only national collaboration	28.80%	1,909	39,015	20.4	1.79
Only institutional collaboration	47.20%	3,132	44,494	14.2	1.42
Single authorship (no collaboration)	3.10%	206	704	3.4	0.52

Figure 8 shows network graphs of coauthorship relations among authors. Coauthorship relations represent whether an author has written a paper with another author. Figure 8 presents the three collaboration clusters among the authors with 50 or more publications. All the institutions in the red cluster belong to China. It can be seen that the cooperation between authors is mainly focused on the same country or neighboring countries.

SCIENTIFIC PRODUCTIVITY AND RELATIONSHIP ANALYSIS

Price's Law

Price's Law is the most commonly used indicator to analyze productivity in a specific discipline. It states that half of the literature on a subject will be contributed by the square root of the total number of authors publishing in that area. For AR-CNN over the entire period, at least 95 primary authors (the square root of 9088 authors) were required to publish 1939 papers, whereas according to Price's calculations, 95 authors were needed to produce 3316 papers (the

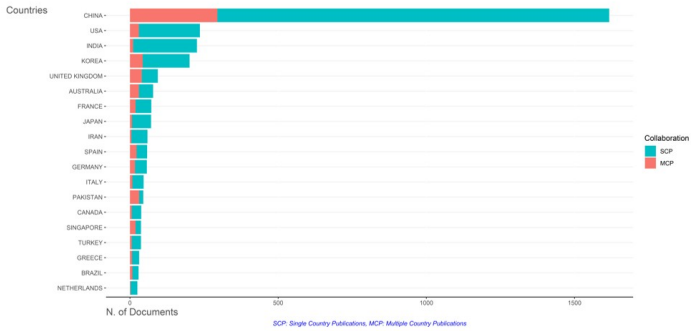

Figure 7 Corresponding Author's Country

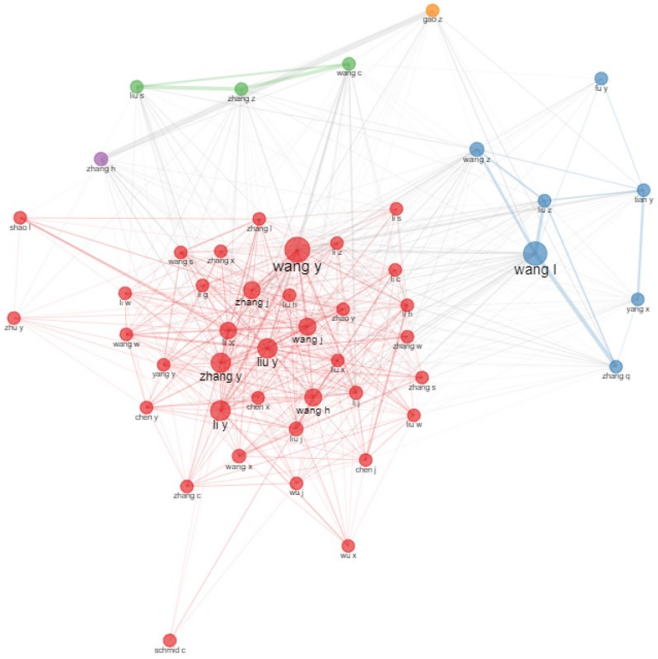

Figure 8 Network visualization of coauthorship collaboration

half of 6633 outputs). Therefore, our data did not support Price's Law.

Bradford's Law

Bradford's Law is a bibliometric indicator of the quantitative relationship between journals and their published papers (Bradford 1948). In this Law, the journals are arranged in descending order of productivity and divided into equal zones (core, allied, and alien). Each zone would contain a similar number of publications, but the number of journals would increase upon moving from one zone to another.

The total 6633 publications are divided into three groups. The first zone (core journals) contained 15 journals with 2260 (34%) publications. The second zone (allied journals) contained 144 journals with 2186 (33%) publications. The third zone (alien journals) contained 1433 journals with 2186 (33%) publications. The summary of the division of zones is as follows. Bradford's algebraic interpretation of the Law is 1: n: n^2. The connection of each zone in this study is 15:144:1433. Here, 15 is the number of journals in the core zone, and Bradford's mean multiplier is 9.78.

Hence,

$$15 : 15 * 9.78 : 15\,(9.78)^2$$

$$15 : 147 : 1433 \approx 1595$$

$$Percentage\ of\ Error = \frac{1595 - 1592}{1592} * 100 = 0.19\%$$

The error percentage is very low (0.19%); therefore, Bradford's Law fits the above data.

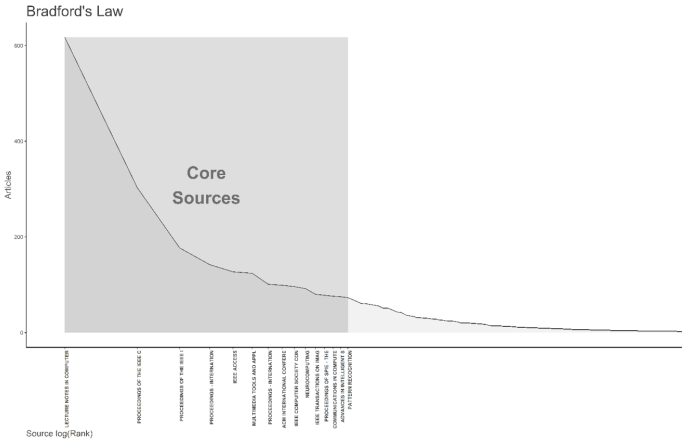

Figure 9 Scattering of Journals over Bradford Zone

Lotka's Law

Lotka's Law Lotka (1926) describes the frequency distribution of authors on scientific productivity in any given field. This Law is expressed as $Y = KX^{-b}$, where x is the number of contributing authors and y is the number of publications, whereas K and b are parameters to be estimated from data. According to this Law, the number of authors that publish n publications in a given subject field is inversely proportional to n^2. This means that for every 100 authors contributing to one publication, 25 will contribute 2, 11 will contribute three, and so on. According to Lotka's Law of Scientific Productivity, only six percent of the authors in a given field will produce more than ten publications. Using the Kolmogorov-Smirnov test (KS-test) of goodness-of-fit, we find that the null hypothesis that the A literature conforms to Lotka's law must be rejected at 0.01 level of significance.

After analyzing data, we determined that in CA, 5666 authors published one paper, which accounted for 62.3% of the total authors (9088), 1378 authors published two papers, which accounted for 15.2% and 640 authors published three papers, which accounted for 7%. This conclusion indicates that only six percent of the authors in AR-CNN will produce more than ten publications, as in Lotka's Law of Scientific Productivity. In other words, the scientific elite of researchers, each of whom publishes many papers, is responsible for advancing the field.

Correlation Analysis

Spearman's rank correlation test was used to assess the relationship between publication numbers and GDP per capita. The test resulted in a p-value of 0.002, rejecting the null hypothesis that there is no significant relationship between publications and GDP per capita at the significance level of 5%. The number of publications

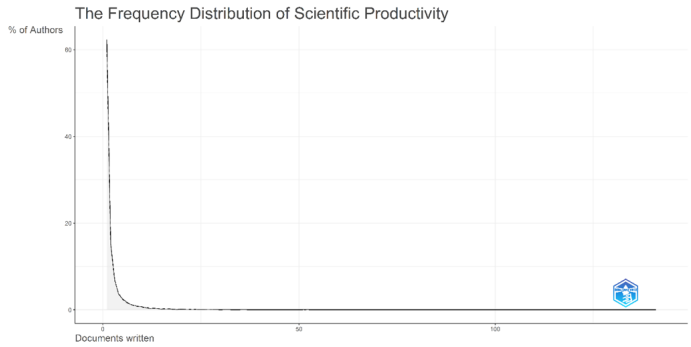

Figure 10 Lotka's Law on the productivity of authors

correlates significantly positively with GDP per capita (r=0.23). The correlation results by country have been found as r=0.72 for the United States, r=0.75 for Australia, r=0.99 for China, r=0.68 for Germany, r=0.96 for India, r=0.66 for Japan and r=0.92 for Pakistan. There is no significant relationship between publications and GDP per capita for other countries.

Keyword Analysis

A total of 8918 different author keywords were used in 6633 publications, and 514 keywords meet the threshold set at the 5-minimum number of occurrences for the keywords. 6991 author keywords appeared only once, 913 keywords appeared only twice, and 320 keywords appeared only three times. The number of keywords used more than three times was 694 (10.5%), which showed that the popular research topics in AR-CNN focused on a small field. A large number of once-only author keywords probably indicates a lack of continuity in research and a wide disparity in the research focuses. The clustering analysis between these keywords is given in the network visualization map (Fig. 11). According to the analysis of the frequency of keywords, the AR-CNN-related research was basically steady, and mainstream research was centered on the following methods: action recognition, human action recognition, and security systems. Except for the seven top keywords, the four frequently used keywords were optical flow, video analysis, feature extraction, and video classification. The keyword "optical flow" indicates the importance of analyzing moving agents, such as creatures or robots, to adapt to their environment.

Thematic Evolution

Figure 12 demonstrates the evolution of keywords in two different stages (2012-2018 and 2019–2021). The left side indicates the themes that were studied during 2012-2018. Six themes are listed in different sizes based on their usage. Themes include action detection, action recognition, deep learning, transfer learning, human action recognition, and feature extraction. Most of the research themes of the initial period were lumped together under these two large topics (human action recognition and action recognition) in the developing period. Feature extraction-related research lines took an important place in both stages. Figure 12 shows that words like transfer learning are highly connected to human action recognition.

DISCUSSION

The objective of this study was to estimate the research productivity of the topic "Action Recognition; Convolutional Neural Network; Video Surveillance". The publications included in the study

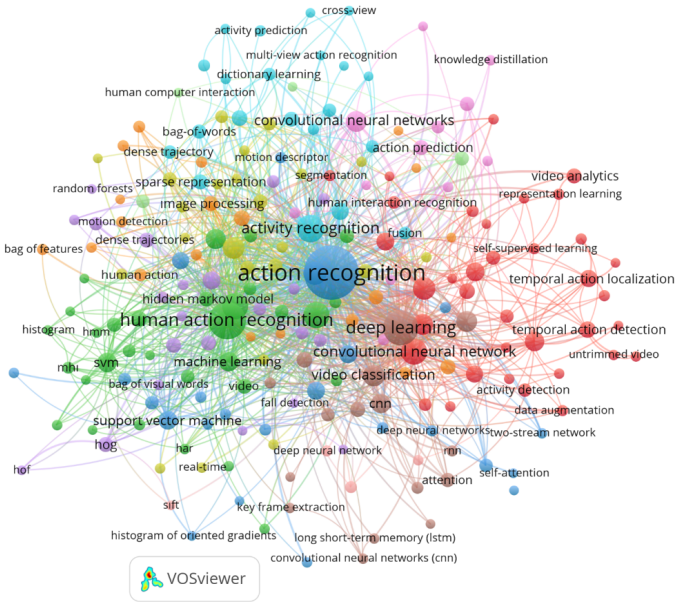

Figure 11 Keyword Analysis

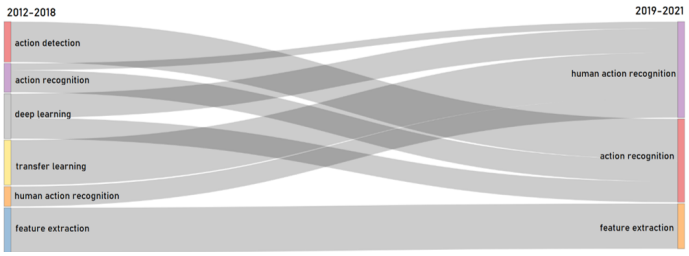

Figure 12 Thematic evolution of the AR-CNN

were collected from the Scopus database during 2012–2021. Of the 6632 publications, 62% were conference papers, 35% were research articles, and 3% were others (chapters, conference reviews, etc.).

When the number of publications is evaluated by years, it is seen that there is a linear increase in the number of publications, with a significant increase since 2017. According to the regression analysis, AR-CNN has attracted serious concerns, and more and more publications have been published in the last four years. China ranked first with the highest number of publications, accounting for 39.8% of worldwide publications. It can be predicted that its predominant position will be further enhanced in the next few years. The Chinese Academy of Sciences published the largest number of papers when considering institutions. Through the coauthorship analysis of countries, this study determined that China was at the center of international cooperation. The strongest link strength was evidenced by the USA and China, with a 260 link strength, followed by the UK and China, with a 105 link strength. According to Lotka's Law, authors' production was consistent, meaning that the number of authors who wrote a single paper was approximately 60%. The number of publications is significantly positively correlated with GDP per capita for the United States, Australia, China, Germany, India, Japan, and Pakistan. There is no significant relationship between publications and GDP per capita for other countries. Based on an analysis of article keywords, mainstream research was centered on the following keywords: action recognition, human action recognition, and security systems.

CONCLUSION

This study aims to quantitatively evaluate the publications on the topics of SciVal's "Action Recognition; Convolutional Neural Network; Video Surveillance (T.561)" in computer vision research between 2011 and 2021. It has been observed that there is generally an increase in the annual number of studies produced from 2011 to 2021. AR-CNN has a FWCI above 1 and is cited more than expected based on the global average, with the most prolific author in terms of citations per publication (CPP) being Andrew P. Zisserman (20 articles, 502 CPP). The first organization that contributed the most to the literature was the Chinese Academy of Sciences (259, 3.9%), while the most productive organization in terms of CPP was the Institut National de Recherche en Informatique et en Automatique (77 articles, 125.6 CPP). Lecture Notes in Computer Science was the most productive source. Zisserman, Andrew P. was the most prolific author in the FWCI and h-index, citing per publication. "Learning Spatiotemporal Features with 3D Convolutional Networks" was the most cited publication. China ranked first with an output of 2638 articles or a share of 39.7%. The most productive countries in terms of the number of publications are developed or overpopulated countries. The participation of researchers in multinational studies in developing or underdeveloped countries may enable them to conduct more research on this subject.

Availability of data and material

Not applicable.

Conflicts of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Ethical standard

The authors have no relevant financial or non-financial interests to disclose.

LITERATURE CITED

A., P., 1969 Statistical bibliography or bibliometrics. Journal of Documentation **25**: 348–.

Ahad, M., 2011 *Computer vision and action recognition: A guide for image processing and computer vision community for action understanding*.

Aria, M. and C. Cuccurullo, 2017 bibliometrix: An r-tool for comprehensive science mapping analysis. Journal of Informetrics **11**: 959–975.

Bi, Y., S. Kapoor, and R. Bhatia, 2017 *Proceedings of SAI Intelligent Systems Conference (IntelliSys) 2016: Volume 2*.

Bobick, A. and J. Davis, 1996 An appearance-based representation of action. In *Proceedings of the International Conference on Pattern Recognition (ICPR)*, volume 1, pp. 307–312.

Bradford, S. C., 1948 *Documentation*. Crosby Lockwood.

Chen, H. and Z. Deng, 2020 Bibliometric analysis of the application of convolutional neural network in computer vision. IEEE Access **8**: 155417–155428.

Ci, Y., F. Jiao, W. Tu, and J. Fang, 2021 Analysis on development tendency of computer vision and graphics based on bibliometrics. Journal of Physics: Conference Series **1883**: 012054.

Donahue, J. *et al.*, 2015 Long-term recurrent convolutional networks for visual recognition and description. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2625–2634.

Karpathy, A., G. Toderici, S. Shetty, T. Leung, R. Sukthankar, *et al.*, 2014 Large-scale video classification with convolutional neural networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) pp. 1725–1732.

Khan, M. A. *et al.*, 2020 Human action recognition using fusion of multiview and deep features: an application to video surveillance. Multimedia Tools and Applications pp. 1–27.

Kiduk, Y. and L. I. Meho, 2006 Citation analysis: A comparison of google scholar, scopus, and web of science. Proceedings of the American Society for Information Science and Technology **43**: 1–15.

LeCun, Y., L. Bottou, Y. Bengio, and P. Haffner, 1998 Gradient-based learning applied to document recognition. Proceedings of the IEEE **86**: 2278–2323.

Lotka, A. J., 1926 The frequency distribution of scientific productivity. Journal of the Washington Academy of Sciences **16**: 317–323.

Mongeon, P. and A. Paul-Hus, 2016 The journal coverage of web of science and scopus: a comparative analysis. Scientometrics **106**: 213–228.

Oh, S. *et al.*, 2011 A large-scale benchmark dataset for event recognition in surveillance video. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3153–3160.

Popoola, O. P. and K. Wang, 2012 Video-based abnormal human behavior recognition: a review. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) **42**: 865–878.

Rosales, R. and S. Sclaroff, 2001 Trajectory guided tracking and recognition of actions. In *Proceedings of the International Conference on Pattern Recognition (ICPR)*.

SciVal, 2023 Scival - working with research areas. https://www.research.uky.edu/uploads/working-research-areas, Accessed: 2023-12-06.

Simonyan, K. and A. Zisserman, 2014 Two-stream convolutional networks for action recognition in videos. In *Advances in Neural Information Processing Systems (NIPS)*, volume 27.

Sudha, M. R., K. Sriraghav, S. S. Abisheck, S. G. Jacob, and S. Manisha, 2017 Approaches and applications of virtual reality and gesture recognition: A review. International Journal of Ambient Computing and Intelligence (IJACI) **8**: 1–18.

Xia, L., I. Gori, J. K. Aggarwal, and M. S. Ryoo, 2015 Robot-centric activity recognition from first-person rgb-d videos. In *Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 357–364.

Xuefei, C., 2014 Ideal-oriented policymaking. Chinese Education and Society **44**: 8–18.

Zhou, Y., X. Sun, Z.-J. Zha, and W. Zeng, 2018 Mict: Mixed 3d/2d convolutional tube for human action recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 449–458.

How to cite this article: Yalcinkaya, M., and Akmese, O. F. Bibliometric Analysis of Publications on Action Recognition, Convolutional Neural Network, Video Surveillance During 2012-2021. *Computers and Electronics in Medicine*, 1(1), 24-33, 2024.