Application of convolutional neural networks in medical images: a bibliometric analysis

Huixin Jia[1], Jiali Zhang[1], Kejun Ma[2], Xiaoyan Qiao[3], Lijie Ren[4], Xin Shi[5,6]

[1]School of Computer Science and Technology, Shandong Technology and Business University, Yantai, China; [2]School of Statistics, Shandong Technology and Business University, Yantai, China; [3]School of Mathematics and Information Science, Shandong Technology and Business University, Yantai, China; [4]Department of Neurology, Shenzhen Institute of Translational Medicine, The First Affiliated Hospital of Shenzhen University, Shenzhen Second People's Hospital, Shenzhen, China; [5]School of Health Management/Institute of Health Sciences, China Medical University, Shenyang, China; [6]Immersion Technology and Evaluation Shandong Engineering Research Center, Shandong Technology and Business University, Yantai, China

Contributions: (I) Conception and design: H Jia, X Shi; (II) Administrative support: X Shi, X Qiao; (III) Provision of study materials or patients: L Ren; (IV) Collection and assembly of data: H Jia, J Zhang; (V) Data analysis and interpretation: H Jia, J Zhang, K Ma; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Xin Shi, PhD. School of Health Management/Institute of Health Sciences, China Medical University, No. 77 Puhe Road, Shenyang North New Area, Shenyang 110122, China; Immersion Technology and Evaluation Shandong Engineering Research Center, Shandong Technology and Business University, Yantai, China. Email: 20221010@cmu.edu.cn.

Background: In the field of medical imaging, the rapid rise of convolutional neural networks (CNNs) has presented significant opportunities for conserving healthcare resources. However, with the wide spread application of CNNs, several challenges have emerged, such as enormous data annotation costs, difficulties in ensuring user privacy and security, weak model interpretability, and the consumption of substantial computational resources. The fundamental challenge lies in optimizing and seamlessly integrating CNN technology to enhance the precision and efficiency of medical diagnosis.

Methods: This study sought to provide a comprehensive bibliometric overview of current research on the application of CNNs in medical imaging. Initially, bibliometric methods were used to calculate the frequency statistics, and perform the cluster analysis and the co-citation analysis of countries, institutions, authors, keywords, and references. Subsequently, the latent Dirichlet allocation (LDA) method was employed for the topic modeling of the literature. Next, an in-depth analysis of the topics was conducted, and the topics in the medical field, technical aspects, and trends in topic evolution were summarized. Finally, by integrating the bibliometrics and LDA results, the developmental trajectory, milestones, and future directions in this field were outlined.

Results: A data set containing 6,310 articles in this field published from January 2013 to December 2023 was complied. With a total of 55,538 articles, the United States led in terms of the citation count, while in terms of the publication volume, China led with 2,385 articles. Harvard University emerged as the most influential institution, boasting an average of 69.92 citations per article. Within the realm of CNNs, residual neural network (ResNet) and U-Net stood out, receiving 1,602 and 1,419 citations, respectively, which highlights the significant attention these models have received. The impact of coronavirus disease 2019 (COVID-19) was unmistakable, as reflected by the publication of 597 articles, making it a focal point of research. Additionally, among various disease topics, with 290 articles, brain-related research was the most prevalent. Computed tomography (CT) imaging dominated the research landscape, representing 73% of the 30 different topics.

Conclusions: Over the past 11 years, CNN-related research in medical imaging has grown exponentially. The findings of the present study provide insights into the field's status and research hotspots. In addition, this article meticulously chronicled the development of CNNs and highlighted key milestones, starting with LeNet in 1989, followed by a challenging 20-year exploration period, and culminating in the breakthrough

moment with AlexNet in 2012. Finally, this article explored recent advancements in CNN technology, including semi-supervised learning, efficient learning, trustworthy artificial intelligence (AI), and federated learning methods, and also addressed challenges related to data annotation costs, diagnostic efficiency, model performance, and data privacy.

Introduction

Medical imaging, which can reflect the internal structures or features of the human body, is one of the main bases for modern medical diagnosis (1). Common medical imaging modalities include computed tomography (CT) (2), magnetic resonance imaging (MRI) (3), ultrasound (US) (4), X-rays (5), and positron emission tomography (PET) (6). Convolutional neural networks (CNNs) represent a category of deep-learning algorithms predominantly used to process structured array data, such as images. Characterized by their convolutional layers, CNNs are adept at capturing spatial and hierarchical patterns in data (7). In medical tasks, CNNs are commonly combined with other neural architectures (e.g., recurrent neural networks for sequential data and transformers for pattern recognition) to enhance their applicability and effectiveness in complex medical data analysis. The widespread application of CNNs today is due to their excellent generalization capabilities (8), scalability, adaptability to diverse needs (9,10), as well as their good transferability and interpretability (11,12). The tasks of CNNs in the field of medical imaging mainly include disease classification and grading (13), localization and detection of pathological targets (14), organ region segmentation (15), and image denoising (16), enhancement (17), and fusion (18) to assist clinicians in efficient and accurate image processing.

CNNs usually require a large number of high-quality annotated images for network training. However, when obtaining medical images for CNN training, there may be difficulties in data acquisition (19). First, strict quality control by professionals is required for data acquisition; thus, the large-scale annotation of data incurs significant equipment and labor costs (20,21). Second, while pursuing technological development, patient privacy must be ensured (22). Although techniques such as data augmentation, transfer learning, and semi-supervised learning can partially address these shortcomings (23,24), the processing capabilities of CNNs for computer hardware and data require further improvement. Therefore, it is crucial to use quantitative and qualitative analysis methods to clarify the current status, hotspots, and prospects of CNNs in the field of medical imaging.

This article initially presents the data retrieval strategy and literature grouping methods employed in this study, followed by the results of the bibliometric analysis of the overall trend, countries, institutions, journals, authors, keywords, and references. Latent Dirichlet allocation (LDA) (25,26) was also employed for topic modeling. Subsequently, the literature grouping method was used to delve deeper into the findings of the bibliometric analysis, LDA topic modeling, and milestones. The article concludes by summarizing the research findings and limitations. To date, researchers have predominantly focused on artificial intelligence (AI) or the application of deep-learning models to specific diseases or single image types, and few comprehensive studies have been conducted on CNN technology in medical imaging. By quantitatively analyzing the bibliometric results and qualitatively assessing the field, this study sought to summarize the development and future research directions for CNNs in medical imaging.

Methods

Data collection

In this study, the Science Citation Index Expanded (SCI-E) database of the Web of Science Core Collection (WoSCC) was selected as the source database for the data retrieval. The application of CNNs in medical imaging over the past 11 years (2013–2023) was examined, and a comprehensive search was conducted to retrieve the relevant literature. The language of publication was limited to English, and
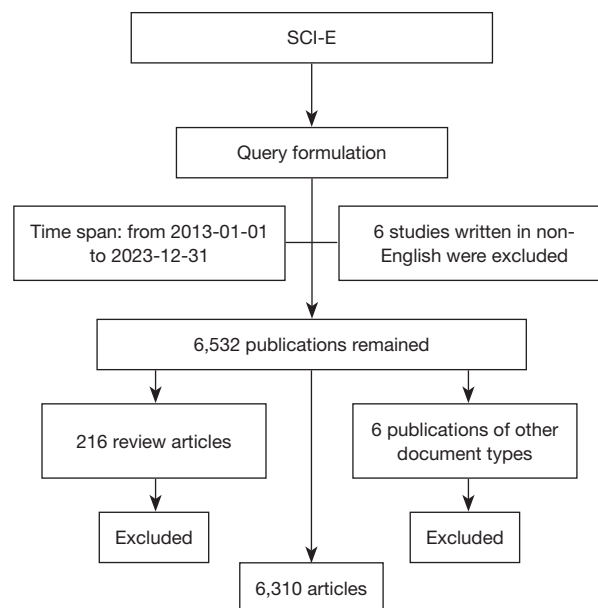
Figure 1 Literature screening flow chart. The authors completed the WoSCC literature retrieval by constructing a retrieval mode and selected relevant articles. SCI-E, Science Citation Index Expanded; WoSCC, Web of Science Core Collection.

the search focused solely on "articles", excluding other document types, such as reviews.

To obtain more comprehensive data, the search strategy involved taking the union of CNNs and medical imaging. Considering the varying language habits of different authors, the keywords for CNNs were set to "convolutional neural network" or "convolutional neural networks" or "convolution neural network" in three forms. The medical image search not only used the phrase "medical image" as a whole, but also combined specific categories of medical imaging, such as radiographic imaging, CT imaging, US imaging, MRI, and nuclear imaging with medical, clinical, diagnostic, and patient-related keywords. Some irrelevant search terms were removed from this search. In addition, this search adopted a cautious approach, generally using the abstract for "and" searches, and subject terms for the "not" searches (*Figure 1*). Finally, this search used the following search formula: ab=("convolutional neural-network" or "convolutional neural networks" or "convolution neural network") and (ab=("medical image" or "medical images" or "medical imaging" or "medical scans") or (ab=(ct or mri or x-ray or ultraso* or pet or pet/ct or mr) and ts=(disease$ or diagnos* or medical or patient or cli nical))) not ts=(crop or plant or "fracture detection" or "road safety" or "fault diagnosis" or "non-destructive testing" or "defect detection" or "ultrasonic signal classification" or "damage diagnostics

" or rock or "dissociation core" or "seismic imaging" or "galaxy cluster mass") and la=(english) and py=(2013–2023) and dt=(article).

A total of 6,310 articles were retrieved. Notably, the number of publications and citations increased annually, until a stable trend was observed last year (*Figure 2*). All the literature was retrieved until December 31, 2023, and the data were exported locally in an Excel, Tab delimited file and Plain text file used for the LDA, VOSviewer and CiteSpace analysis, respectively. All the data used for the searches (e.g., titles, keywords, authors, countries and regions, publishers, date of publications, and citations) were downloaded from the WoSCC database.

Data analysis

Bibliometric analysis is a statistical tool used to evaluate a large amount of scientific output and has been widely used in scientific mapping in recent years (27). This is because a bibliometric analysis can draw conclusions about the connections between articles, journals, authors, keywords, citations, and co-citation networks, which helps researchers to identify research topics and trends, as well as future research directions (28,29). This study used Excel 2019, VOSviewer 1.6.16 (30), CiteSpace 5.8.R3 (31,32), Scimago Graphica 1.0.18, and LDA topic modeling to perform
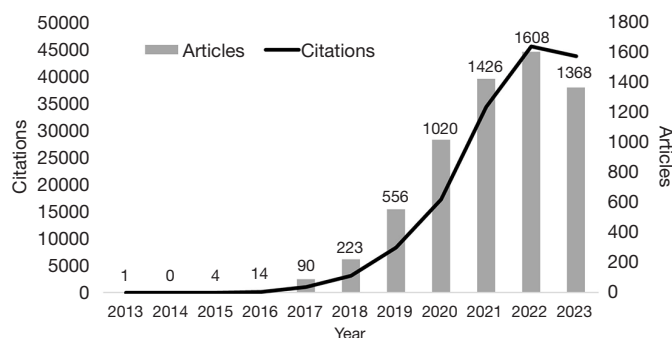
Figure 2 Statistical chart of the number of published articles and citations from 2013 to 2023.

the bibliometric analysis on all 6,310 articles. VOSviewer and Scimago Graphica were used for the co-occurrence analysis of countries, journals, authors, and other entities, and CiteSpace was used for the cluster analysis of subjects, keywords, and references to help visualize the development trends and research hotspots in this field (33).

LDA is a probabilistic topic model that was proposed by Blei *et al*. in 2003 (34). LDA is a machine-learning technique that can be used to identify thematic information hidden in large document collections or corpora. The LDA topic model is a kind of Bayesian probability model. The three layers contain three parts (i.e., themes, keywords, and articles) with a certain probability every article chose a topic, and at a certain probability from the topic chosen for certain keywords, the article to the subject, subject to the term polynomial distributed (35). LDA topic modeling was carried out on the retrieved literature, and each topic summarized by the LDA was evaluated in terms of the topic quality, topic center, and evolution trend (36). In addition, this study also classified the topics in terms of the following tasks: medical segmentation task, medical classification task, and image processing task.

The knowledge establishment method used a combination of bibliometrics and LDA topic modeling to construct the development trajectory of CNNs in the field of medical image analysis. It began with direct observation (without making assumptions), and induced experience from that observation, and then generalized that experience into conclusions. This study explored some of the factors affecting the scientific research strength of a country in terms of the overall development trend, country, institution, journal, and author perspectives. Based on this keyword analysis and LDA topic modeling, this study also identified the future trends in CNN technology and the hot topics in medical image analysis. By analyzing the citation and co-

citation data of the reference articles, this study examined the development trajectory and technological changes in CNNs since since they were first introduced, as well as the reasons behind the development of CNN technology.

Results

Bibliometrics

Contribution of countries and regions

All the literature was distributed across 97 countries/regions. This study measured national contributions to published articles (*Table 1*). In terms of article production, the top three countries were China (2,385 articles), the United States (1,247 articles), and India (755 articles). In terms of total citations, the top three countries were the United States (55,538 times), China (43,788 times), and India (14,826 times). In terms of time, the average publication time in the United States was four months earlier than that in China (see *Figure 3A* in which the colors closer to purple/yellow indicate an earlier/later average publication time for each country/region). In terms of collaboration breadth, the United States (1,083 articles in 67 other countries), China (1,062 articles in 62 other countries), and the United Kingdom (562 articles in 61 other countries) had the widest collaboration with other countries (see *Figure 3B* in which the colors closer to red/yellow indicate collaborations with more/less countries).

Institutions and journals

This study includes contributions from 6,406 institutions (*Table 2*). The article count represents the number of articles in which an institution participated (if an author from an institution was listed as an author, the article count for the institution was increased by one; thus, the

Table 1 Country/region contribution to article publication

Rank	Country/region	Articles	Top 100 articles	Total citations	Average citations
1	China	2,385	23	43,788	18.36
2	USA	1,274	38	55,538	43.59
3	India	755	7	14,826	19.64
4	South Korea	423	9	11,275	26.65
5	Saudi Arabia	386	1	4,772	12.36
6	UK	194	19	14,136	72.87
7	Germany	307	4	7,503	24.44
8	Canada	271	8	8,993	33.18
9	Japan	221	2	4,486	20.30
10	Australia	213	2	5,604	26.30

sum of the "Articles" does not equal the total number of articles retrieved). In terms of article output, the top three institutions were the Chinese Academy of Sciences (187 articles, 7 articles in the top 100), the University of California (162 articles, 5 articles in the top 100), and Sun Yat Sen University (122 articles). In terms of citation count, the top three institutions were the University of California (9,801 total citations, 60.5 average citations), Harvard University (7,412 total citations, 69.92 average citations), and Chinese Academy of Sciences (6,197 total citations, 33.14 average citations). In addition, the citing and cited relationship between the journals can show the interconnections between disciplines. The left side of the figure is the citing journal, and the right side is the cited journal (*Figure 4*).

Keywords

The keyword analysis showed that among 13,741 keywords, 100 keywords were cited more than 54 times (*Figure 5A*). Notably, the keyword "convolutional neural network" (2,665 times) was cited the most frequently, followed by "deep learning" (2,299 times), and "classification" (1,038 times). In terms of the strongest citation bursts, 13 keywords have been prominently cited (see *Figure 5B* in which the different colored lines in the graph represent the intensity of the citation bursts). The keyword burst in this field mainly occurred from 2015 to 2017, during which time, there was a notable increase in the focus on issues such as the "brain", "segmentation", and "quantification".

References

The 6,310 articles analyzed cited a total of 130,336 references, of which 106 references were cited more than

100 times. In terms of the citation count, the most popular were residual neural network (ResNet) (37), which was introduced by He (1,602 times), U-Net (38), which was introduced by Ronneberger (1,419 times), and AlexNet (39), which was introduced by Krizhevsky (1,228 times) (*Table 3*). In terms of the strongest citation bursts, the research advances with the greatest breakthrough were U-Net, which was introduced by Ronneberger (strength 138.16), VGGNet (42), which was introduced by Zisserman (strength 108.95), and the Adam (43) algorithm, which was introduced by Kingma (strength 97.81) (*Figure 6A*). This study also conducted a clustering analysis on the titles of references (*Figure 6B*).

LDA topic modeling

Evaluation indicator

In this study, the LDA topic modeling method was used to cluster all the retrieved articles into 30 topics. To objectively evaluate the quality of these topics, the study employed indicators such as topic size and document prominence (*Table 4*). In the context of LDA topic modeling, understanding how "Topic Size" and "Document Prominence" are calculated can provide insights into the referential role of these evaluation indicators. The retrieved document collection is denoted as D ($D = \{d_1, \cdots, d_{|D|}\}$), where $|D|$ is the total number of retrieved documents. The corresponding corpus for this document collection D is denoted as C ($C = \{c_1, \cdots, c_{|C|}\}$), where $|C|$ represents the total number of words in the corpus. The Document-Corpus Matrix, denoted as $DC_{|D| \times |C|}$, is constructed using frequency counts of words, where $p(d_m, c_n)$ represents the frequency
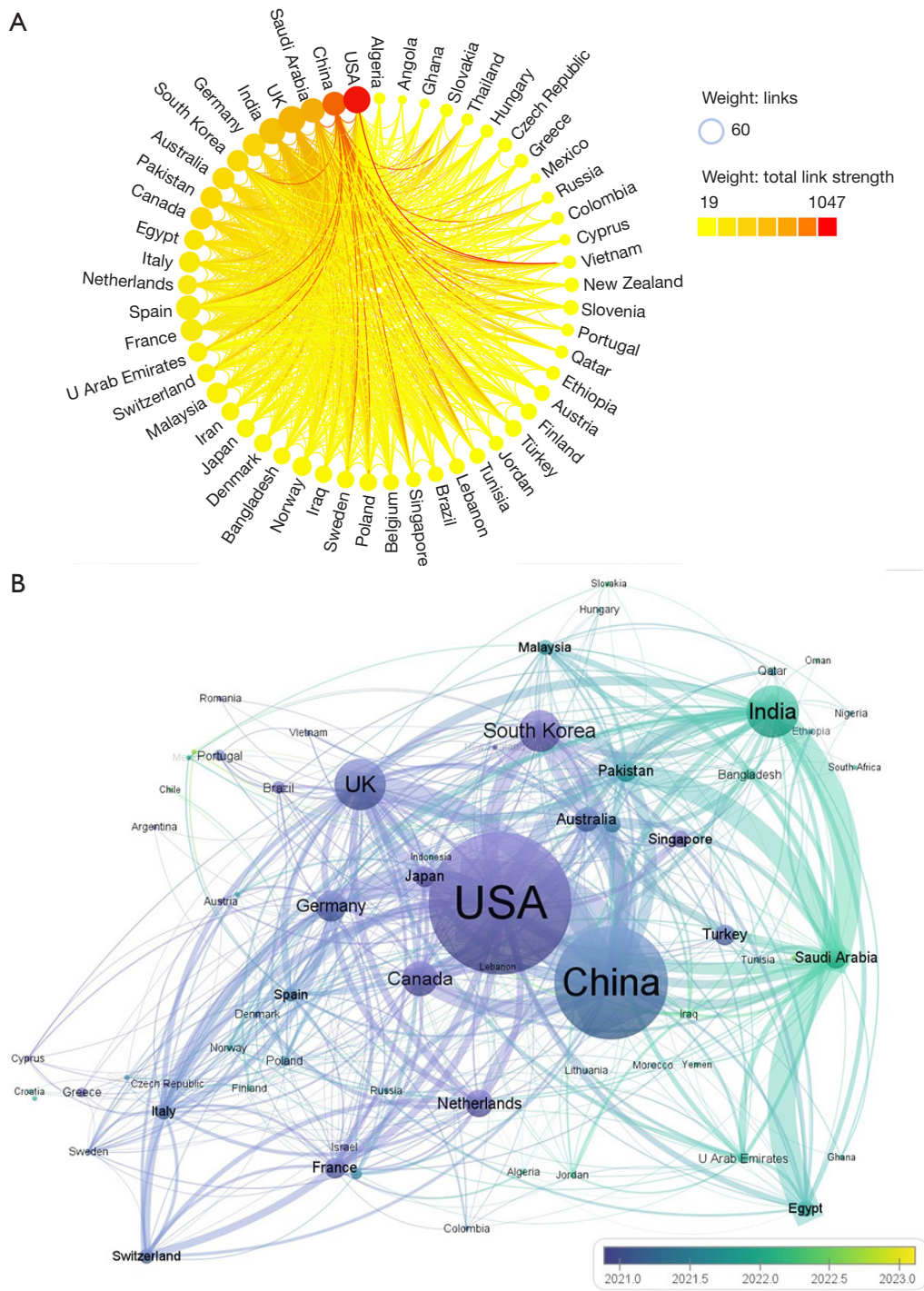
Figure 3 Visualization of the contribution of countries/regions to the number of publications. (A) The extent of cooperation among countries/regions. (B) The average publication time of countries/regions in this field.

Table 2 Statistics on the number of publications, total citations, and average citations of institutions

| Rank	Institution	Articles	Top 100 articles	Total citations	Average citations
1	Chinese Academy of Sciences	187	7	6,197	33.14
2	University of California System	162	5	9,801	60.50
3	Sun Yat-Sen University	122	0	2,351	19.27
4	Shanghai Jiao Tong University	114	2	2,774	24.33
5	Harvard University	106	8	7,412	69.92
6	Fudan University	100	0	1,796	17.96
7	Stanford University	91	5	4,560	50.11
8	Zhejiang University	86	0	1,915	22.27
9	University of Texas System	85	2	4,962	58.38
10	Yonsei University	81	4	1,794	22.15
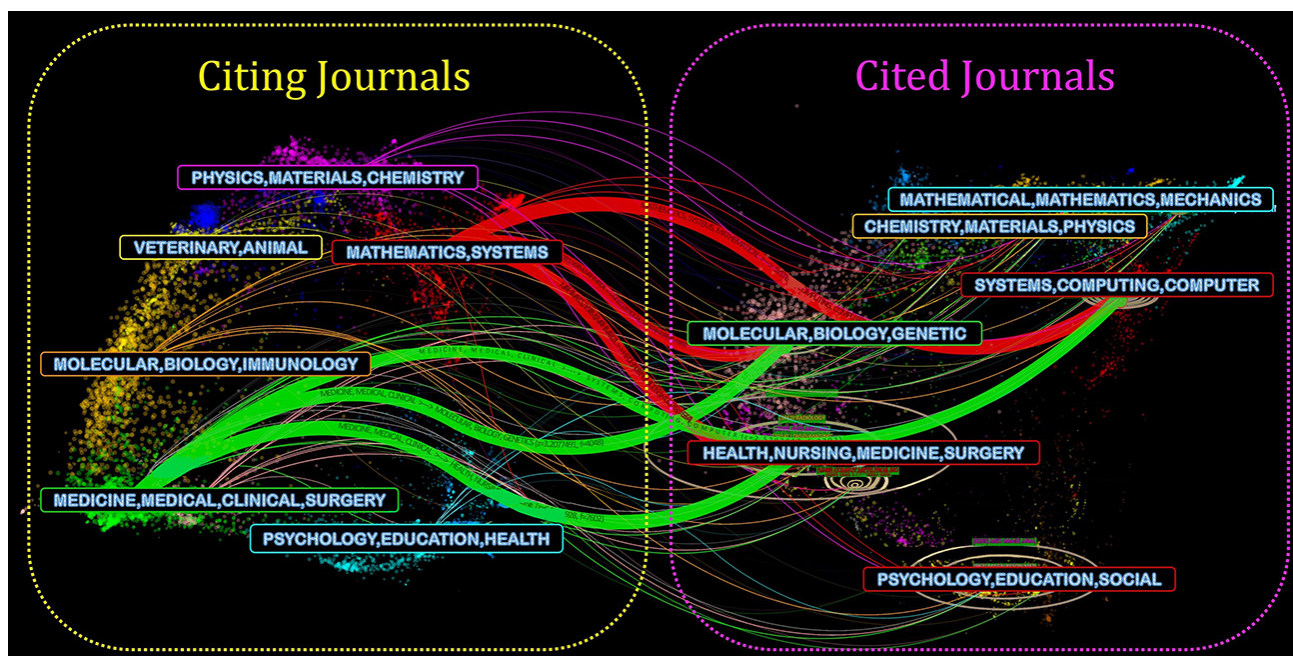
Figure 4 Citing and cited relationships between journals from different categories.

of corpus word c_n in document d_m, for $m = \{1, 2, \cdots, |D|\}$ and $n = \{1, 2, \cdots, |C|\}$.

According to the principles of LDA, the matrix $DC_{|D| \times |C|}$ is comprised of the Document-Topic Matrix $DT_{|D| \times |T|}$ and the Topic-Corpus Matrix $TC_{|T| \times |C|}$. In this decomposition, the Document-Topic Matrix DT ultimately represents the score of documents d_m in topic t_k as p (d_m, t_k), for $m = \{1, 2, \cdots, |D|\}$ and $k = \{1, 2, \cdots, |T|\}$, where $|T|$ is the total number of topics. The Topic-Corpus Matrix TC ultimately represents the probability of a topic in corpus words as p (t_k, c_n), for $k = \{1, 2, \cdots, |T|\}$, $n = \{1, 2, \cdots, |C|\}$. The decomposed topic matrix is denoted as T ($T = \{t_1, \cdots, t_{|T|}\}$).

Based on this, the size of a topic t_k, is represented by the sum of its probability in the corpus as shown in Eq. [1]. Document Prominence refers to the number of documents included in a topic, where these documents must not only have the highest probability in that topic but also exceed an
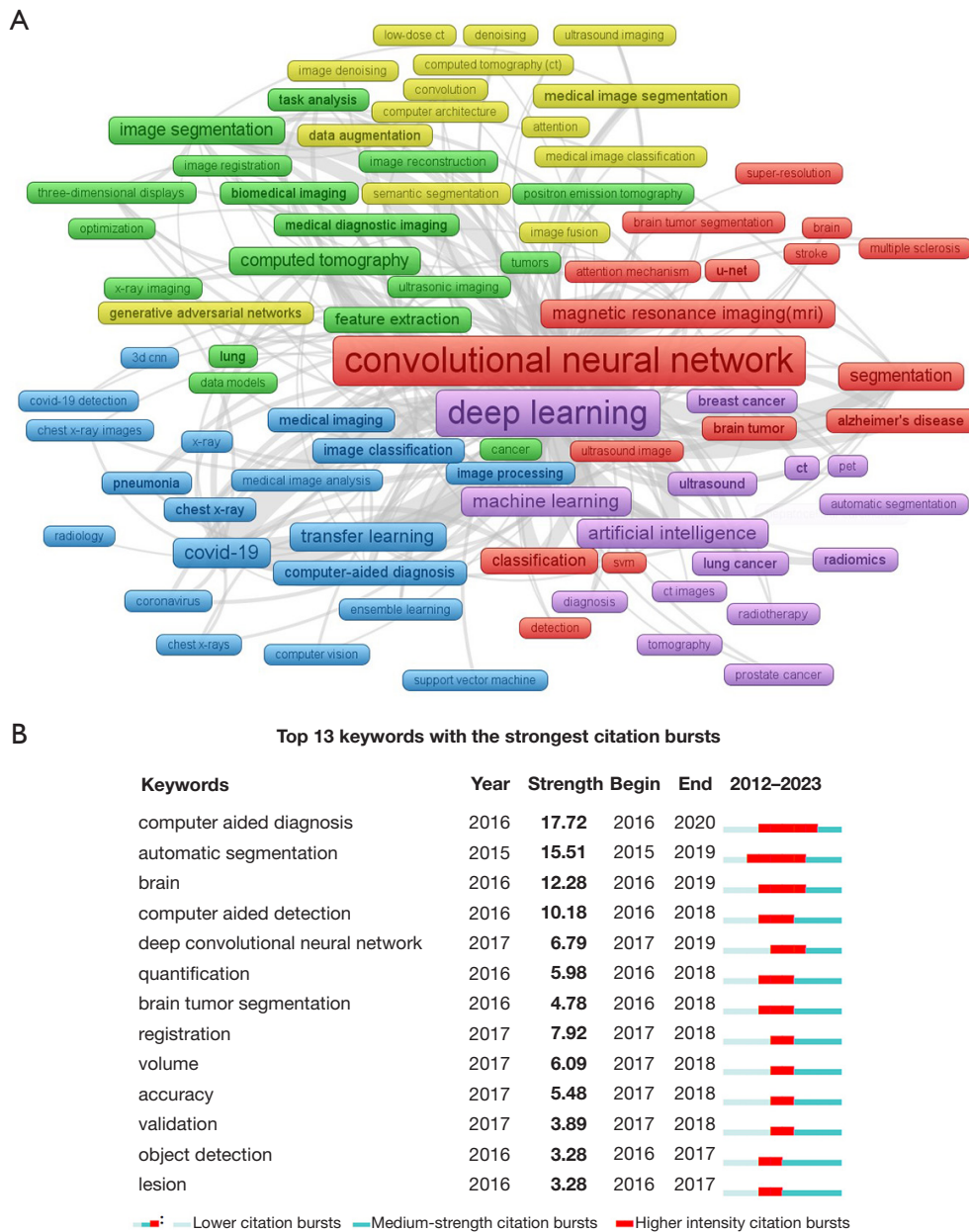
Figure 5 Visualization of the keyword data analysis. (A) Classification and display of high-frequency author keywords. (B) Keywords with the strongest citation bursts; light blue, blue, and red indicate low, medium, and high citation burst intensity, respectively.

average threshold δ (generally $\delta = 0.2$) as shown in Eq. [2].

$$\text{Topic size}(t_k) = \sum_{n=1}^{|C|} p(t_k, c_n), \quad \forall k = \{1, 2, \cdots, |T|\} \tag{1}$$

$$\text{Doc Prominence}(t_k) = \sum_{m=1}^{|D|} \begin{cases} 1 & \text{if} \begin{cases} p(t_k, d_m) = \max_{k \in \{1,2,\cdots,|T|\}} p(t_k, d_m) \\ \& p(t_k, d_m) > \delta \end{cases} \\ 0 & \text{otherwise} \end{cases}, \quad \forall k = \{1, 2, \cdots, |T|\} \tag{2}$$

Table 3 Statistics for total citations of references (14,37-52)

Rank	References	Citations
1	He K, 2016, Conference on Computer Vision and Pattern Recognition (CVPR), Deep residual learning for image recognition	1,602
2	Ronneberger O, 2015, Lecture Notes in Computer Science, U-Net: Convolutional Networks for Biomedical Image Segmentation	1,419
3	Krizhevsky A, 2012, Communications of the ACM, ImageNet classification with deep convolutional neural networks	1,228
4	Karen S, 2015, International Conference on Learning Representations (ICLR), Very Deep Convolutional Networks for Large-Scale Image Recognition	1,103
5	Szegedy C, 2015, Conference on Computer Vision and Pattern Recognition (CVPR), Going deeper with convolutions	984
6	Kingma DP, 2014, arXiv, Adam: A Method for Stochastic Optimization	746
7	Litjens G, 2017, Med Image Anal, A Survey on Deep Learning in Medical Image Analysis	639
8	Shelhamer E, 2015, Conference on Computer Vision and Pattern Recognition (CVPR), Fully Convolutional Networks for Semantic Segmentation	484
9	Milletari F, 2016 Fourth International Conference on 3D Vision (3DV), V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation	476
10	Huang G, 2017, Conference on Computer Vision and Pattern Recognition (CVPR), Densely Connected Convolutional Networks	466
11	LeCun Y, 2015, Nature, Deep Learning	398
12	Srivastava N, 2014, Journal of Machine Learning Research, Dropout: A Simple Way to Prevent Neural Networks from Overfitting	379
13	Deng J, 2009, Conference on Computer Vision and Pattern Recognition (CVPR), Imagenet: A large scale hierarchical image database	370
14	LeCun Y, 1998, Proceedings of the IEEE, Gradient-based learning applied to document recognition	363
15	Russakovsky O, 2015, Int J Comput Vis, Imagenet large scale visual recognition challenge	335
16	Shin HC, 2016, IEEE Trans Med Imaging, Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning	312
17	Kamnitsas K, 2017, Med Image Anal, Efficient Multi-Scale 3D CNN with fully connected CRF for Accurate Brain Lesion Segmentation	302

Therefore, topics that simultaneously possess a smaller Topic Size and a higher Document Prominence are of higher quality, as they are more concentrated in the corpus and exhibit stronger relevance between the documents and the topics, and thus have a greater referential value.

Topic summary

To clearly present these topics, this study created summaries around the first keyword of each topic. The study systematically reviewed the disease domains and corresponding technological issues studied in each topic (*Table 5*). Approximately two-thirds of the first topic words were related to diseases (e.g., nodules, brain, and Alzheimer's disease), while one-third of the first topic words were related to technology (e.g., denoising, multimodality, and image registration). In the research of disease topics, while medical tasks like detection and diagnosis appeared in almost every topic, different diseases still tended to have specific inclinations toward particular medical imaging tasks. For example, topics like nodules (60), breast (61), and Alzheimer's disease (62) leaned more toward image classification tasks, while topics like bone (63,64), prostate (65), and stroke (44) leaned more toward image segmentation tasks.

Image types in topics

Different subjects usually have a certain preference in the study of image types. This study conducted a statistical analysis of the image types adopted by these topics (*Table 6*).
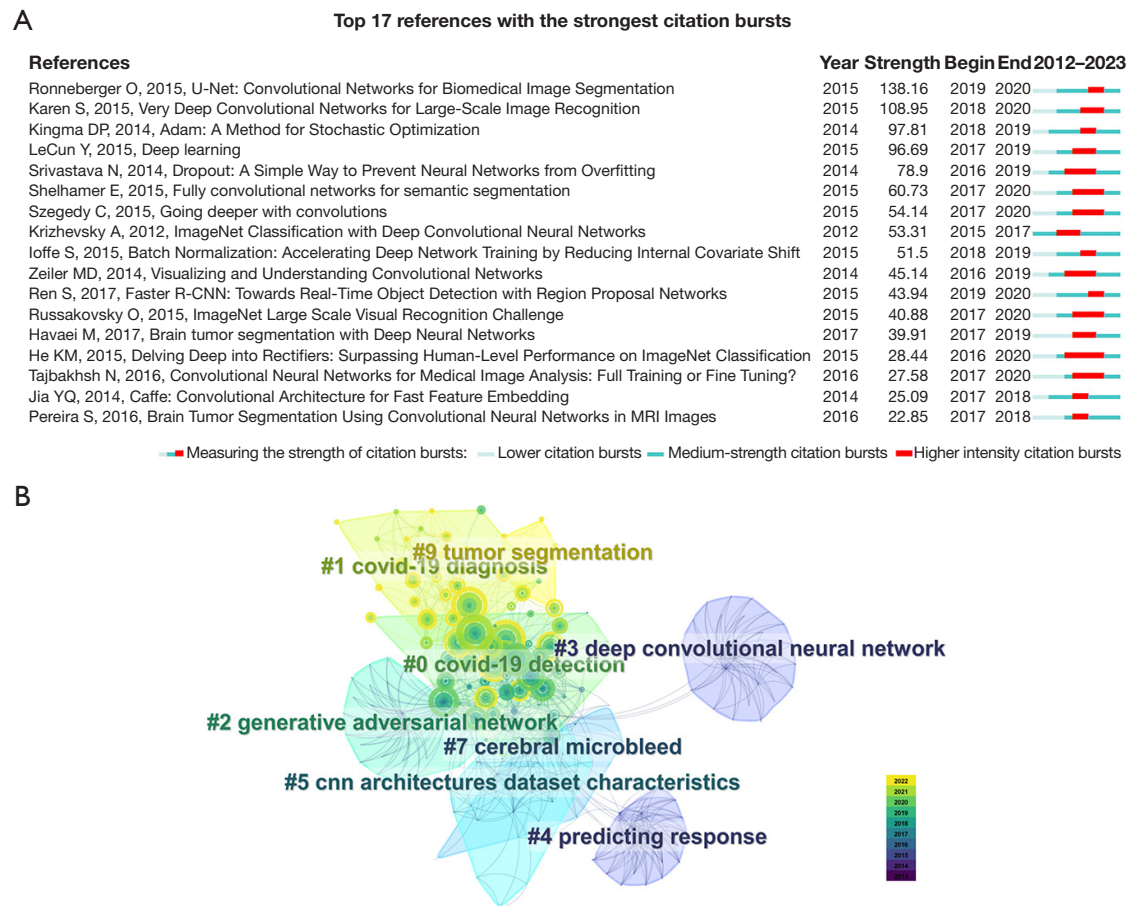
A

Top 17 references with the strongest citation bursts

References	Year	Strength	Begin	End	2012–2023
Ronneberger O, 2015, U-Net: Convolutional Networks for Biomedical Image Segmentation	2015	138.16	2019	2020	
Karen S, 2015, Very Deep Convolutional Networks for Large-Scale Image Recognition	2015	108.95	2018	2020	
Kingma DP, 2014, Adam: A Method for Stochastic Optimization	2014	97.81	2018	2019	
LeCun Y, 2015, Deep learning	2015	96.69	2017	2019	
Srivastava N, 2014, Dropout: A Simple Way to Prevent Neural Networks from Overfitting	2014	78.9	2016	2019	
Shelhamer E, 2015, Fully convolutional networks for semantic segmentation	2015	60.73	2017	2020	
Szegedy C, 2015, Going deeper with convolutions	2015	54.14	2017	2020	
Krizhevsky A, 2012, ImageNet Classification with Deep Convolutional Neural Networks	2012	53.31	2015	2017	
Ioffe S, 2015, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift	2015	51.5	2018	2019	
Zeiler MD, 2014, Visualizing and Understanding Convolutional Networks	2014	45.14	2016	2019	
Ren S, 2017, Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks	2015	43.94	2019	2020	
Russakovsky O, 2015, ImageNet Large Scale Visual Recognition Challenge	2015	40.88	2017	2020	
Havaei M, 2017, Brain tumor segmentation with Deep Neural Networks	2017	39.91	2017	2019	
He KM, 2015, Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification	2015	28.44	2016	2020	
Tajbakhsh N, 2016, Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?	2016	27.58	2017	2020	
Jia YQ, 2014, Caffe: Convolutional Architecture for Fast Feature Embedding	2014	25.09	2017	2018	
Pereira S, 2016, Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images	2016	22.85	2017	2018	

Measuring the strength of citation bursts: — Lower citation bursts — Medium-strength citation bursts — Higher intensity citation bursts

B

Figure 6 Visualization of the reference data analysis. (A) Top 17 references with the strongest citation bursts; light blue, blue, and red indicate low, medium, and high citation burst intensity, respectively (15,37-41,43,46,48,49,53-59). (B) Reference cluster display, different colors distinguish among cluster topics; circles indicate the research time of the included subtopics.

CT and MRI were widely studied in Topic 22 and Topic 20, respectively, accounting for 73% and 67% of the total topics, respectively. X-ray and US images were widely studied in Topic 12 and Topic 13, respectively, accounting for 40% and 43% of the total topics, respectively. PET images were only involved in six topics, accounting for 20% of the total topics. It is evident from this statistical analysis that CT and MRI were the most widely used imaging modalities, followed by X-ray and US images, while PET was the least used.

No imaging technology is perfect. For instance, X-rays and CT scans involve exposure to radiation, which poses potential health risks with long-term or high-dosage use. MRI, while providing detailed images, is expensive and sensitive to metal implants, which limits its applicability to some patients. US and PET scans struggle with issues such as low image or poor spatial resolution. These limitations necessitate technical solutions to aid in diagnosis, such as enhancing the image clarity in images with low contrast and resolution. Imaging techniques with low radiation dosage but high noise levels require sophisticated denoising methods to maintain their diagnostic accuracy. Further, variability in image quality due to patient movement or differences in imaging protocols across institutions calls for robust standardization and normalization techniques. Additionally, integrating heterogeneous data from multiple imaging sources, such as combining MRI, CT, and PET scans, presents significant challenges in data fusion and alignment. These technical issues underscore the need for continual advancements in image processing algorithms and machine-learning techniques to address the shortcomings of current imaging technologies, thereby improving the overall quality and reliability of medical diagnostics.

Table 4 LDA topic modeling effect evaluation

Rank	Topic label	Topic size	Doc prominence
1	COVID-19	109.23	99
2	Liver	146.24	99
3	Nodule	149.78	90
4	Vessel	156.18	84
5	Bone	159.13	89
6	Breast	164.84	69
7	Brain	166.46	69
8	Alzheimer's disease	178.54	70
9	Denoising	180.06	47
10	Prostate	180.44	63
11	PET	180.65	84
12	Radiomics	182.65	50
13	Diabetes	184.71	56
14	Stroke	185.85	58
15	Lymphatic	187.17	71
16	Cardiac	189.77	84
17	Knee	190.83	72
18	Registration	192.74	55
19	Dental	198.51	59
20	Uncertainty	201.06	52
21	Compression	206.14	44
22	Pancreas	206.76	55
23	Muscle	209.14	67
24	Multi-modal	210.29	56
25	Retrieval	211.67	67
26	Cortical	212.5	45
27	Tuberculosis	215.21	57
28	Cervical	225.69	48
29	Hemorrhage	228.35	51
30	Hypertension	230.42	28

LDA, latent Dirichlet allocation; COVID-19, coronavirus disease 2019; PET, positron emission tomography.

Evolutionary trends

The topics had different temporal evolution trends, and this study conducted a statistical analysis to examine the annual publication volume (see *Figure* 7 in which PY denotes the Year of Publication). The results showed that the research area of coronavirus disease 2019 (COVID-19) detection (Topic 18) has been very popular since 2020. The popularity of topics such as multiple sclerosis (Topic 2), Alzheimer's disease (Topic 7), and brain tumors (Topic 26) has been increasing over the years. The publication volume of the topic of nodule detection and classification (Topic 27) began to rise in 2017, peaked in 2020, and has gradually declined since then. In terms of the overall number of publications on each topic, COVID-19 pneumonia (Topic 18, 597 publications), Alzheimer's disease (Topic 7, 290 publications), and breast cancer (Topic 15, 274 publications) were the top three topics with the most publications.

Discussion

Characteristics of publications summarized through bibliometric methods

The number of articles published by Chinese authors was twice that of American authors, but the number of articles in the top 100 rankings and the average citation count per publication were only half of those of the United States. The United States, the United Kingdom, Germany, Canada, South Korea, and other countries started research in this field earlier, and thus their publications had an average lead time over other nations. The United States exhibited the highest level of international collaboration in this field, having engaged in 1,083 collaborations with 67 different countries. Among all the publishing institutions, Harvard University, the University of California, and the University of Texas had significant influence, even though the Chinese Academy of Sciences had published the largest number of articles.

Hot topics of focus induced by the LDA method

Following the rapid development of CNN algorithms in 2014/2015, computer-aided diagnosis began to flourish in 2016 and continues to flourish to this day. The shortage of medical resources is a common problem worldwide. The application of large models, such as ResNet and U-Net, in medical image classification, detection, and segmentation has greatly saved medical costs. Among them, the segmentation task of brain tumors has received the most attention. Since 2015, high-impact articles have been published in the field of brain tumor segmentation, and this trend has continued to this day. Since 2016, brain, automatic segmentation, and brain tumor segmentation have

Table 5 Diseases and technologies focused on each topic

Topic	Label	Disease focus of topic	Technology focus of topic
1	Denoising	Brain, chest	Low-dose CT image denoising
2	Compression	Multiple sclerosis	Lesion segmentation
3	Liver	Liver tumor	Lesion classification, organ segmentation
4	Tuberculosis	Chest tuberculosis	Disease detection, fast screening
5	Radiomics	Lung or rectal cancer	Treatment response assessment or prediction
6	Registration	Lung diseases	Image registration
7	Alzheimer's disease	Alzheimer's disease	Diagnosis, classification, prediction
8	Stroke	Stroke, brain	Lesion segmentation
9	Knee	Knee bone and cartilage	Automated segmentation
10	Pancreas	Pancreatic cancer	Diagnosis, segmentation, prediction
11	Prostate	Prostate tumor	Diagnosis, segmentation, prediction
12	Hemorrhage	Intracerebral hemorrhage	Automated diagnosis
13	PET	Brain, amyloid	Low-dose PET image reconstruction
14	Muscle	Muscle	Automated segmentation
15	Breast	Breast cancer	Diagnosis, segmentation, classification
16	Cortical	Cortical surface	Reconstruction, parcellation
17	Cardiac	Cardiac, left ventricle	Quantification, segmentation
18	COVID-19	COVID-19 pneumonia	Diagnosis, prediction
19	Uncertainty	Tuberculosis, stroke	Uncertainty estimation, quantification
20	Retrieval	Brain, dermatoscopic	Image retrieval
21	Vessel	Coronary artery, carotid	Automatic characterization, image correction
22	Multi-modal	Brain tumor	Multi-modal medical image segmentation
23	Diabetes	Diabetic retinopathy	Localization, detection, classification
24	Dental	Dental caries	Diagnosis, age estimate, segmentation
25	Bone	Spine, femur	Automatic segmentation, bone age labeling
26	Brain	Gliomas brain tumor	Grades classification, grading, segmentation
27	Nodule	Lung nodule	Detection, classification,
28	Lymphatic	Lymph node metastasis	Prediction, pretreatment identification
29	Hypertension	Hypertension	Diagnosis, classification, prediction
30	Cervical	Cervical cancer	Segmentation, classification

CT, computed tomography; PET, positron emission tomography; COVID-19, coronavirus disease 2019.

Table 6 Statistics on the types of images contained in each topic

Statistical items in topics	Image type				
	CT	MRI	X-ray	PET	Ultrasound
Number of applications	22	20	12	6	13
Proportion in topic (%)	73	67	40	20	43

CT, computed tomography; MRI, magnetic resonance imaging; PET, positron emission tomography.

Topic / PY	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
2023	46	75	53	53	42	49	61	55	46	62	29	45	53	53	70	48	51	125	54	56	47	52	36	43	55	47	40	52	38	53
2022	49	59	55	55	31	38	83	52	42	47	24	38	45	42	76	36	49	202	52	54	47	50	49	42	43	58	43	44	54	47
2021	52	50	49	44	47	35	55	42	33	34	40	29	47	27	44	34	42	206	36	33	34	31	24	32	43	48	43	26	36	28
2020	31	47	40	37	22	24	43	33	26	26	28	20	24	25	41	23	31	61	26	26	27	24	33	19	31	27	65	24	36	33
2019	25	15	24	18	16	24	28	21	13	16	23	16	18	17	27	13	16	2	14	11	19	20	14	15	15	12	37	18	12	15
2018	5	10	16	9	7	8	15	7	6	2	11	3	11	5	14	2	7	1	5	10	8	5	4	5	6	12	11	8	3	5
2017	8	4	7	2	4	2	5	3	0	2	3	1	3	2	2	1	1	0	0	7	2	2	1	2	6	4	11	3	0	1
2016	0	0	1	1	0	2	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	3	2	0	1	0	0	0	0
2015	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
2014	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2013	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
SUM	216	260	245	195	169	182	290	214	166	189	158	153	202	172	274	157	198	597	187	197	184	184	165	161	200	209	250	175	179	182
	Red (50< Topic articles)									Yellow (25≤ Topic articles ≤50)										Green (Topic articles< 25)										

Figure 7 Statistics for the number of published articles per year and the total number of published articles on the topic of LDA modeling. PY, publication year; LDA, latent Dirichlet allocation.
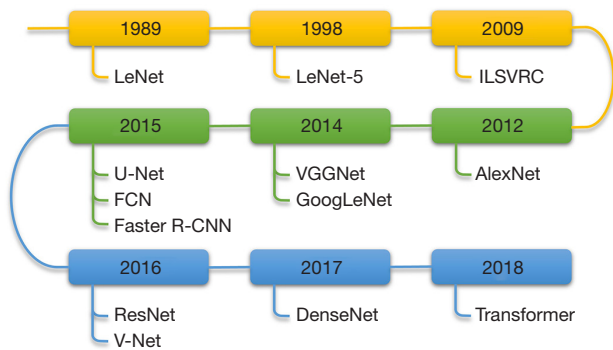
Figure 8 The development of CNNs in the field of medical images. ILSVRC, ImageNet Large Scale Visual Recognition Challenge; VGGNet, the CNNs proposed by the Visual Geometry Group; FCN, fully convolutional network; CNNs, convolutional neural networks.

become popular keywords. This keyword is second only to "computer-aided diagnosis", which shows the cutting-edge progress in the field of brain tumor segmentation.

CNNs have shown great potential and application prospects in segmentation (15), classification (66), detection, and diagnosis (67), but image preprocessing techniques, such as image denoising (68), image registration (69), and image compression (70), can still provide effective preprocessing methods for these tasks to improve the accuracy of experimental results. Multimodality learning (71) and multi-task learning (MTL) (72) represent breakthroughs in the development of CNNs in the past two years. In medical image analysis, multimodality usually refers to the combination of different types of medical images (e.g., CT,

MRI, and PET) or other types of data (e.g., medical records and physiological indicators) to obtain more comprehensive and accurate diagnostic results. MTL is a neural network method that learns multiple related tasks simultaneously in one model. The extensive research on multimodality allows existing data to be more fully used horizontally or vertically in the context of difficult-to-obtain and difficult-to-annotate high-quality medical image training data. The extensive research on multi-tasking has resulted in savings in computational resources, such as computer graphics processing units (GPUs), and reductions in training costs, while allowing CNNs to learn more tasks simultaneously.

Development trends and milestone events

The highlighted references (*Figure 6A*) and highly cited references (*Table 3*) provide a good guide to the development and technological evolution of CNNs. Through a thorough analysis of these articles, this study identified milestone events in the development of CNNs (*Figure 8*), examined the reasons for their occurrence, and predicted the future directions of CNNs.

In 1989, LeCun and his team proposed LeNet, a CNN model. After a decade of improvement, LeNet-5 (1998, cited 363 times) was successfully applied to handwritten digit recognition, and is considered a pioneering model in the field of CNNs (45). The development of CNNs cannot be separated from the support of massive databases. In 2009, Deng J and her team established the ImageNet database (cited 370 times), which contains over 14 million labeled full-size images (46,47). This became the main data source for the ImageNet Large-Scale Visual Recognition Challenge

(ILSVRC) competition, and also provided a fertile ground for the subsequent technological changes in CNNs (*Figure 8*, yellow).

Following the development of digital technology, the expansion of the scale of the available image data, improvements in the computing speed of GPU hardware, and the continuous improvement of algorithms, AlexNet (citations: 1,228, burst strength: 53.31) came into being. In the 2012 ILSVRC, AlexNet won the championship with an absolute advantage of 10.9 percentage points over the model that placed second. This breakthrough is of milestone significance. The upsurge in the industry has continued to this day, and the emergence of AlexNet can be described as the return of the king of CNNs. In 2014, the Visual Geometry Group of Oxford University proposed VGGNets (citations: 1,103, burst strength: 108.95). Based on AlexNet, VGGNets effectively improved the effect of the model by using smaller convolution kernels and increasing the network depth. Later, GoogLeNet (citation 984, burst strength: 54.14) introduced the inception module into the network architecture to further improve the overall performance of the model (48). Although the depth reached 22 layers, the parameter size was much smaller than the parameter sizes of AlexNet and VGGNet.

The image classification task first caught the attention of CNN researchers, and the use of CNN algorithms in the field of image segmentation and object detection was proposed soon thereafter. In 2015, U-Net (citations: 1,419; burst strength: 138.16) proposed a network and training strategy that effectively leveraged image augmentation techniques to make the most out of insufficiently available annotated samples, thus achieving excellent results in end-to-end training. Faster R-CNNs (burst strength: 43.94) form the basis of multiple first-place award-winning works in multiple areas, and truly realize the end-to-end target detection framework (59). Fully convolutional networks (FCNs) (citations: 484, burst strength: 60.73) demonstrate that CNNs, which are trained end-to-end and pixel-to-pixel, surpass state-of-the-art models for semantic segmentation (*Figure 8*, green) (49). Although CNNs are widely used to solve problems in the fields of computer vision and medical image analysis, most methods can only deal with two-dimensional images, but most medical data used in clinical practice are three-dimensional (3D). In 2016, Milletari *et al.* proposed V-Net (citations: 476), which is a 3D image segmentation method based on a volumetric FCN (50). This work created a precedent whereby CNNs moved from 2D to 3D, and greatly enriched the application scenarios in the

fields of computer vision and medical image analysis.

As the network layers are built deeper and deeper, the difficulty of training also increases. The ResNet model proposed by He in 2015 (in an article published in 2016 that has since had 1,602 citations) uses a residual learning framework to simplify training (51). This model won the championship at the 2015 ILSVRC. The residual network of ResNet is deeper than that of the VGGNet, but its complexity is lower, and the degradation problem caused by a too-deep network is solved by introducing the residual unit, so that the deep network can be trained more stably. DenseNet (466 citations) proposed in 2017 was improved on the basis of ResNet (52). For each layer, all the feature maps of the previous layer are used as input, and its own feature map is used for all subsequent layers. The input of the algorithm adopts a dense multiplexing method, which greatly reduces the number of parameters and achieves high performance with less calculation. Transformer is a neural network architecture in the field of natural language processing. It was first proposed by Google in 2017 (73). It is a neural network model based on the attention mechanism. Since its inception, Transformer has achieved great success in the field of natural language processing and has become one of the most popular deep-learning models. Behind the success of the Transformer model is not only the powerful performance of the model itself, but also the rapid development of deep-learning technology, as well as the support of huge data sets and computing resources. With the development of the Transformer model, more and more researchers and companies have begun to apply it to speech recognition, image processing, and other fields (74,75) (*Figure 8*, blue).

The latest CNN-based techniques

The application of neural networks in medical image processing has advanced the field of medical diagnosis. Similarly, the demand for medical imaging diagnosis has propelled advancements in CNN technologies. Medical imaging demands top-tier neural networks, which has led to the development of semi-supervised learning, supervised learning, trustworthy AI training, and federated learning methods. Semi-supervised learning involves learning from data sets that contain labeled and unlabeled data. The effective use of unlabeled data is highly valuable in improving model performance, reducing manual costs, and handling large-scale medical image data. Learning refers to efficiently using computational resources, data, and time to minimizing computational and storage requirements during the training

process. This leads to rapid, cost-effective machine-learning and deep-learning methods that enhance the efficiency, accuracy, and usability of medical image analysis. Techniques like data augmentation, model compression, and transfer learning contribute to this goal. Trustworthy AI refers to purchasing AI systems with characteristics such as high operability, interpretability, privacy protection, and fairness. These features are crucial in medical diagnostics. Federated learning is a machine-learning method where machine-learning models are trained across multiple data sources without centralizing the data. In federated learning, models are trained locally on devices like smartphones and medical equipment. Updates to these local models are then sent to a central server through encrypted, secure communication protocols. The updates from these local models are aggregated to form a global model that is then sent back to each local device. This process continues iteratively. It preserves data privacy while allowing the continuous optimization of the global model.

Conclusions

Judging from the characteristics of the publications, while there is a certain gap between China's influence in this field and that of developed countries, the gap is continuously narrowing. In terms of technical application, the current research on diseases, such as new techniques for coronavirus detection, brain tumor segmentation, and detection, are CNN hotspots in the field of medical images. In terms of model evolution, LeNet is considered a pioneering model in the field of CNNs. After more than 10 years of silence, the breakthrough of AlexNet gave CNNs new momentum. Since then, U-Net and ResNet have set off a new wave of computer-aided diagnosis. In recent years, research hotspots have included multi-modal and multi-task analysis, and the combination of CNNs and Transformer.

Acknowledgments

Footnote

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Anastasiou E, Lorentz KO, Stein GJ, Mitchell PD. Prehistoric schistosomiasis parasite found in the Middle East. Lancet Infect Dis 2014;14:553-4.
2. Halder A, Dey D, Sadhu AK. Lung Nodule Detection from Feature Engineering to Deep Learning in Thoracic CT Images: a Comprehensive Review. J Digit Imaging 2020;33:655-77.
3. Tang Q, Liu X, Jiang Q, Zhu L, Zhang J, Wu PY, Jiang

3516

Jia et al. Application of CNNs in medical images

Y, Zhou J. Unenhanced magnetic resonance imaging of papillary thyroid carcinoma with emphasis on diffusion kurtosis imaging. Quant Imaging Med Surg 2023;13:2697-707.

4. Qi X, Zhang L, Chen Y, Pi Y, Chen Y, Lv Q, Yi Z. Automated diagnosis of breast ultrasonography images using deep neural networks. Med Image Anal 2019;52:185-98.

5. Apostolopoulos ID, Mpesiana TA. Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med 2020;43:635-40.

6. Ding Y, Sohn JH, Kawczynski MG, Trivedi H, Harnish R, Jenkins NW, Lituiev D, Copeland TP, Aboian MS, Mari Aparici C, Behr SC, Flavell RR, Huang SY, Zalocusky KA, Nardo L, Seo Y, Hawkins RA, Hernandez Pampaloni M, Hadley D, Franc BL. A Deep Learning Model to Predict a Diagnosis of Alzheimer Disease by Using (18)F-FDG PET of the Brain. Radiology 2019;290:456-64.

7. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, Santamaría J, Fadhel MA, Al-Amidie M, Farhan L. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data 2021;8:53.

8. Alzubaidi L, Al-Amidie M, Al-Asadi A, Humaidi AJ, Al-Shamma O, Fadhel MA, Zhang J, Santamaría J, Duan Y. Novel Transfer Learning Approach for Medical Imaging with Limited Labeled Data. Cancers (Basel) 2021;13:1590.

9. Chang JW, Kang KW, Kang SJ. An energy-efficient fpga-based deconvolutional neural networks accelerator for single image super-resolution. IEEE Transactions on Circuits and Systems for Video Technology 2020;30:281-95.

10. Singh S, Anand RS. Multimodal medical image fusion using hybrid layer decomposition with CNN-based feature mapping and structural clustering. IEEE Transactions on Instrumentation and Measurement 2020;69:3855-65.

11. Zhang W, Liljedahl AK, Kanevskiy M, Epstein HE, Jones BM, Jorgenson MT, Kent K. Transferability of the Deep Learning Mask R-CNN Model for Automated Mapping of Ice-Wedge Polygons in High-Resolution Satellite and UAV Images. Remote Sens 2020;12:1085.

12. Zhang Q, Zhu S. Visual interpretability for deep learning: a survey. Frontiers Inf Technol Electronic Eng 2018;19:27-39.

13. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA 2016;316:2402-10.

14. Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. IEEE Trans Med Imaging 2016;35:1285-98.

15. Pereira S, Pinto A, Alves V, Silva CA. Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. IEEE Trans Med Imaging 2016;35:1240-51.

16. Li M, Hsu W, Xie X, Cong J, Gao W. SACNN: Self-Attention Convolutional Neural Network for Low-Dose CT Denoising With Self-Supervised Perceptual Loss Network. IEEE Trans Med Imaging 2020;39:2289-301.

17. Rahman T, Khandakar A, Qiblawey Y, Tahir A, Kiranyaz S, Abul Kashem SB, Islam MT, Al Maadeed S, Zughaier SM, Khan MS, Chowdhury MEH. Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. Comput Biol Med 2021;132:104319.

18. Liu Y, Chen X, Wang Z, Wang ZJ, Ward RK, Wang X. Deep learning for pixel-level image fusion: Recent advances and future prospects. Inf Fusion 2018;42:158-73.

19. Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H. Gan-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. Neurocomputing 2018;321:321-31.

20. Payer C, Štern D, Bischof H, Urschler M. Integrating spatial configuration into heatmap regression based CNNs for landmark localization. Med Image Anal 2019;54:207-19.

21. Wang X, Chen H, Xiang H, Lin H, Lin X, Heng PA. Deep virtual adversarial self-training with consistency regularization for semi-supervised medical image classification. Med Image Anal 2021;70:102010.

22. Fezai L, Urruty T, Bourdon P, Fernandez-Maloigne C. Deep anonymization of medical imaging. Multimed Tools 2023;82:9533-47.

23. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data 2019;6:60.

24. Xie Y, Zhang J, Xia Y. Semi-supervised adversarial model for benign-malignant lung nodule classification on chest CT. Med Image Anal 2019;57:237-48.

25. Choi S, Seo J. An Exploratory Study of the Research on Caregiver Depression: Using Bibliometrics and LDA Topic Modeling. Issues Ment Health Nurs 2020;41:592-601.

26. Heo GE, Kang KY, Song M, Lee JH. Analyzing the field of bioinformatics with the multi-faceted topic modeling technique. BMC Bioinformatics 2017;18:251.

27. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM. How to conduct a bibliometric analysis: An overview and

guidelines. J Bus Res 2021;133:285-96.

28. Pu QH, Lyu QJ, Su HY. Bibliometric analysis of scientific publications in transplantation journals from Mainland China, Japan, South Korea and Taiwan between 2006 and 2015. BMJ Open 2016;6:e011623.

29. Wei N, Xu Y, Li Y, Shi J, Zhang X, You Y, Sun Q, Zhai H, Hu Y. A bibliometric analysis of T cell and atherosclerosis. Front Immunol 2022;13:948314.

30. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 2010;84:523-38.

31. Chen C. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. J Am Soc Inf Sci Technol 2006;57:359-77.

32. Chen C. Searching for intellectual turning points: progressive knowledge domain visualization. Proc Natl Acad Sci U S A 2004;101 Suppl 1:5303-10.

33. Wu H, Tong L, Wang Y, Yan H, Sun Z. Bibliometric Analysis of Global Research Trends on Ultrasound Microbubble: A Quickly Developing Field. Front Pharmacol 2021;12:646626.

34. Blei DM, Ng A.Y, Jordan MI. Latent dirichlet allocation. J Mach Learn Res 2003;3:993-1022.

35. Shen J, Huang W, Hu Q. PICF-LDA: a topic enhanced LDA with probability incremental correction factor for Web API service clustering. J Cloud Comp 2022. doi: 10.1186/s13677-022-00291-9.

36. Zhang Z, Wang Z, Huang Y. A Bibliometric Analysis of 8,276 Publications During the Past 25 Years on Cholangiocarcinoma by Machine Learning. Front Oncol 2021;11:687904.

37. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells W, Frangi A. editors. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, Springer, Cham, 2015:9351.

38. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Communications of the ACM 2012;60:84-90.

39. Karen S, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:1409.1556.

40. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436-44.

41. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. J Mach Learn Res 2014;15:1929-58.

42. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak JAWM, van Ginneken B, Sánchez CI. A survey on deep learning in medical image analysis. Med Image Anal 2017;42:60-88.

43. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. arXiv:1412.6980.

44. Kamnitsas K, Ledig C, Newcombe VFJ, Simpson JP, Kane AD, Menon DK, Rueckert D, Glocker B. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med Image Anal 2017;36:61-78.

45. LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proceedings of the IEEE 1998;86:2278-324.

46. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Li FF. Imagenet large scale visual recognition challenge. Int J Comput Vis 2015;115:211-52.

47. Deng J, Dong W, Socher R, Li LJ, Li K, Li FF. Imagenet: A large scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009;248-55.

48. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015;1-9.

49. Shelhamer E, Long J, Darrell T. Fully Convolutional Networks for Semantic Segmentation. IEEE Conference on Computer Vision and Pattern Recognition 2015;3431-40.

50. Milletari F, Navab N, Ahmadi SA. V-net: Fully convolutional neural networks for volumetric medical image segmentation. 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 2016:565-71.

51. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016;770-8.

52. Huang G, Liu Z, van der Maaten L, Weinberger KQ. Densely connected convolutional networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017;4700-8.

53. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. Proceedings of the 32nd International Conference on Machine Learning, Lille, France, 2015;37:448-56.

54. Zeiler MD, Fergus R. Visualizing and Understanding Convolutional Networks. Lecture Notes in Computer Science 2014. doi: 10.1007/978-3-319-10590-1_53.

3518

Jia et al. Application of CNNs in medical images

55. Havaei M, Davy A, Warde-Farley D, Biard A, Courville A, Bengio Y, Pal C, Jodoin PM, Larochelle H. Brain tumor segmentation with Deep Neural Networks. Med Image Anal 2017;35:18-31.

56. He K, Zhang X, Ren S, Sun J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015:1026-34.

57. Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, Liang J. Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? IEEE Trans Med Imaging 2016;35:1299-312.

58. Jia J, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T. Caffe: Convolutional Architecture for Fast Feature Embedding. Proceedings of the 22nd ACM international conference on Multimedia 2014:675-8.

59. Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Trans Pattern Anal Mach Intell 2017;39:1137-49.

60. Xie Y, Zhang J, Xia Y, Fulham M, Zhang Y. Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest CT. Inf Fusion 2018;42:102-10.

61. Ribli D, Horváth A, Unger Z, Pollner P, Csabai I. Detecting and classifying lesions in mammograms with Deep Learning. Sci Rep 2018;8:4165.

62. Liu M, Li F, Yan H, Wang K, Ma Y; Alzheimer's Disease Neuroimaging Initiative; Shen L, Xu M. A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease. Neuroimage 2020;208:116459.

63. Deniz CM, Xiang S, Hallyburton RS, Welbeck A, Babb JS, Honig S, Cho K, Chang G. Segmentation of the Proximal Femur from MR Images using Deep Convolutional Neural Networks. Sci Rep 2018;8:16485.

64. Ambellan F, Tack A, Ehlke M, Zachow S. Automated segmentation of knee bone and cartilage combining statistical shape knowledge and convolutional neural networks: Data from the Osteoarthritis Initiative. Med Image Anal 2019;52:109-18.

65. Wang B, Lei Y, Tian S, Wang T, Liu Y, Patel P, Jani AB, Mao H, Curran WJ, Liu T, Yang X. Deeply supervised 3D fully convolutional networks with group dilated convolution for automatic MRI prostate segmentation. Med Phys 2019;46:1707-18.

66. Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE Trans Med Imaging 2016;35:1207-16.

67. Khan AI, Shah JL, Bhat MM. CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. Comput Methods Programs Biomed 2020;196,105581.

68. Sun J, Du Y, Li C, Wu TH, Yang B, Mok GSP. Pix2Pix generative adversarial network for low dose myocardial perfusion SPECT denoising. Quant Imaging Med Surg 2022;12:3539-55.

69. Balakrishnan G, Zhao A, Sabuncu MR, Guttag J, Dalca AV. VoxelMorph: A Learning Framework for Deformable Medical Image Registration. IEEE Trans Med Imaging 2019. [Epub ahead of print]. doi: 10.1109/TMI.2019.2897538.

70. Urbaniak I, Wolter M. Quality assessment of compressed and resized medical images based on pattern recognition using a convolutional neural network. Commun Nonlinear Sci Numer Simul 2021;95:105582.

71. Ding Z, Zhou D, Nie R, Hou R, Liu Y. Brain Medical Image Fusion Based on Dual-Branch CNNs in NSST Domain. Biomed Res Int 2020;2020:6265708.

72. Samala RK, Chan HP, Hadjiiski LM, Helvie MA, Cha KH, Richter CD. Multi-task transfer learning deep convolutional neural network: application to computer-aided diagnosis of breast cancer on mammograms. Phys Med Biol 2017;62:8894-908.

73. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I. Attention is all you need. Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS 2017), 2017;6000-10.

74. Afouras T, Chung JS, Senior A, Vinyals O, Zisserman A. Deep audiovisual speech recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 2018;44:8717-27.

75. Khan S, Naseer M, Hayat M, Zamir SW, Khan FS, Shah M. Transformers in vision: A survey. ACM Comput Surv 2022;54:1-41.