



# Artificial intelligence and multimodal data fusion for smart healthcare: topic modeling and bibliometrics

Xieling Chen<sup>1</sup> · Haoran Xie<sup>2</sup> · Xiaohui Tao<sup>3</sup> · Fu Lee Wang<sup>4</sup> · Mingming Leng<sup>2</sup> · Baiying Lei<sup>5</sup>

Accepted: 18 January 2024  
© The Author(s) 2024

## Abstract

Advancements in artificial intelligence (AI) have driven extensive research into developing diverse multimodal data analysis approaches for smart healthcare. There is a scarcity of large-scale analysis of literature in this field based on quantitative approaches. This study performed a bibliometric and topic modeling examination on 683 articles from 2002 to 2022, focusing on research topics and trends, journals, countries/regions, institutions, authors, and scientific collaborations. Results showed that, firstly, the number of articles has grown from 1 in 2002 to 220 in 2022, with a majority being published in interdisciplinary journals that link healthcare and medical research and information technology and AI. Secondly, the significant rise in the quantity of research articles can be attributed to the increasing contribution of scholars from non-English speaking countries/regions and the noteworthy contributions made by authors in the USA and India. Thirdly, researchers show a high interest in diverse research issues, especially, cross-modality magnetic resonance imaging (MRI) for brain tumor analysis, cancer prognosis through multi-dimensional data analysis, and AI-assisted diagnostics and personalization in healthcare, with each topic experiencing a significant increase in research interest. There is an emerging trend towards issues such as applying generative adversarial networks and contrastive learning for multimodal medical image fusion and synthesis and utilizing the combined spatiotemporal resolution of functional MRI and electroencephalogram in a data-centric manner. This study is valuable in enhancing researchers' and practitioners' understanding of the present focal points and upcoming trajectories in AI-powered smart healthcare based on multimodal data analysis.

**Keywords** Multimodal data fusion · Artificial intelligence · Smart healthcare · Topic modeling · Bibliometric analysis

---

✉ Haoran Xie  
hrxie@ieec.org

<sup>1</sup> School of Education, Guangzhou University, Guangzhou, China

<sup>2</sup> Department of Computing and Decision Sciences, Lingnan University, Hong Kong, China

<sup>3</sup> School of Sciences, University of Southern Queensland, Toowoomba, Australia

<sup>4</sup> School of Science and Technology, Hong Kong Metropolitan University, Hong Kong, China

<sup>5</sup> School of Biomedical Engineering, Shenzhen University, Shenzhen, China

## Abbreviations

AI	Artificial intelligence
EHRs	Electronic health records
NLP	Natural language processing
DNNs	Deep neural networks
IoMT	Internet of Medical Things
RQs	Research questions
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SNA	Social network analysis
H-index	Hirsch index
ACP	Average citations per paper
MK	Mann–Kendall
C/Y	Yearly citation rate
DBM	Deep Boltzmann machine
MRI	Magnetic resonance imaging
fMRI	Functional MRI
EEG	Electroencephalography
ECG	Electrocardiogram
PET	Positron emission tomography
AD	Alzheimer’s disease
MCI	Mild cognitive impairment
NSST	Nonsubsampled shearlet transform
MM-DPSN	Multimodal deep polynomial networks stacked
m-PCNN	Multi-channel Pulse coupled neural network
DCAE	Deep coupling autoencoder
GNNs	Graph neural networks
ELM	Extreme learning machine
OCT	Optical coherence tomography
3D	Three dimensional
WoS	Web of Science
FREX	Frequent and exclusive terms
CNN	Convolutional neural network
BCO-MDP	Bacterial colony optimization with multi-dimensional population
IoTs	Internet of Things
GANs	Generative adversarial networks
STM	Structural topic model
SHM	Smart health monitoring
XAI	Explainable AI

## 1 Introduction

The comprehensive examination of integrating various types of data for smart healthcare using artificial intelligence (AI) is emerging as a pivotal and dynamic area of study. Interest in advancing smart healthcare through combining diverse data types using AI is steadily growing across academia, medical or healthcare institutions, and pertinent government sectors. Attention has been paid to specific topics like merging medical signals for smart healthcare, data-driven systems for intelligent healthcare using various

data, and integrating medical data from different sources. These studies largely employed systematic approaches with limited samples involved. No study encompasses the integration of multimodal data fusion, smart healthcare, and AI concurrently. Given the importance of this interdisciplinary field, understanding its challenges and trends is crucial for fostering its future progress. Employing topic models and bibliometric approaches, the present research comprehensively examines existing scientific work regarding AI-driven integration of diverse data in smart healthcare on a global scale.

## 1.1 AI-powered multimodal data fusion in smart healthcare

The healthcare industry has experienced a remarkable transformation driven by significant developments of AI and the proliferation of various data sources (Soni et al. 2020). The integration of multimodal data fusion, an innovative approach that combines information from various sources, carries the capacity to transform healthcare practices in the era of smart healthcare (Yang et al. 2022). This study presents a comprehensive study on AI-powered smart healthcare based on multimodal data fusion using topic modeling and Bibliometric Analysis, exploring the intersection of AI, healthcare, and data analytics.

Smart healthcare implements technologies like wearable gadgets, the Internet of Medical Things (IoMT), advanced machine learning, and wireless devices to effortlessly retrieve medical records, link people, resources, and establishments, and efficiently oversee and address healthcare demands (Muhammad et al. 2021). It encompasses the utilization of diverse digital instruments such as wearables, telehealth, electronic health records (EHR), AI, Internet of Things (IoTs), and big data analytics to transform healthcare provision. The utilization of healthcare devices based on IoT, denoting tools embedded with sensors, connectivity, and software for gathering and transmitting healthcare data, represents a facet of smart healthcare by enabling continuous health monitoring, real-time data gathering, and remote patient care. Nonetheless, IoT-based healthcare devices are not the exclusive focus of smart healthcare.

Smart healthcare utilizes AI's capabilities to handle, scrutinize, and decipher extensive quantities of diverse data, encompassing EHRs, medical images, wearable gadgets, genomics, patient-provided information, and social media content (Nguyen et al. 2022). By synthesizing information from these diverse sources, medical professionals can acquire a comprehensive outlook on the well-being of patients, enabling more precise diagnostics, personalized treatment plans, and proactive interventions (Tian et al. 2019; Huynh et al. 2020).

The key to achieving effective multimodal data fusion lies in applying advanced AI, natural language processing (NLP), and deep neural networks (DNNs) to uncover concealed patterns, identify associations, and draw significant understandings from the fusion of (un)structured information (Ahmed et al. 2020; Noorbakhsh-Sabet et al. 2019). As a result, healthcare professionals can make data-driven decisions with higher accuracy and efficiency, resulting in enhanced patient results and the overall quality of healthcare.

As the healthcare industry progresses towards more data-driven and patient-centric approaches, the implementation of multimodal data fusion supported by AI will become increasingly crucial (Hartl et al. 2021; Holzinger et al. 2022). The potential benefits are far-reaching, including improved disease detection and early diagnosis, more effective treatment planning, better management of chronic conditions, and enhanced healthcare delivery overall (Flores et al. 2021; Albahri et al. 2023).

However, alongside these advancements come various challenges, encompassing worries about data privacy and security, ethical deliberations, and the requirement for AI models that are sturdy and easy to understand. Addressing these issues will be vital to ensuring the successful adoption and integration of AI-powered multimodal data fusion in smart healthcare.

## 1.2 Literature review

Driven by the escalating significance of AI and fusion methodologies across healthcare research, scholars have conducted assessments covering pertinent subjects like AI in healthcare applications and data-driven smart healthcare systems. For instance, Muhammad et al. (Muhammad et al. 2021) meticulously surveyed fusion schemes for multimodal medical signals in smart healthcare based on 105 research papers between 2014 and 2020. Specifically, Muhammad et al. delved into (1) IoMT applications, (2) multi-sensor data fusion levels, and 3) recent advancements in multimodal medical data fusion. (Cai et al. 2019) offered a comprehensive overview of techniques regarding multimodal data-powered smart healthcare applications from January 2013 to September 2018, concentrating on multimodal semantic perception, data fusion, cross-border knowledge fusion, and decision-making systems. (Shaik et al. 2023) delivered an extensive overview of multimodal medical data fusion, exploring diverse approaches like feature selection, rule-powered applications, machine learning, DNNs, and NLP for data fusion and analysis. (Albahri et al. 2023) systematically reviewed 64 contributions regarding AI trustworthiness in healthcare, evaluating quality, bias risk, and data fusion. (Sujith et al. 2022) systematically reviewed smart health monitoring (SHM) using DNNs and AI, addressing recent advancements and challenges in SHM, focusing on features, role of deep learning and AI, structure, data security, and limitations based on studies from 2020 to 2021. (Mohsen et al. 2022) comprehensively analyzed fusion approaches, diseases, outcomes, machine learning algorithms, and available multimodal medical datasets based on 34 studies. (Guo et al. 2020) examined AI research in healthcare based on 1473 publications, emphasizing publication growth, research characteristics, patterns, and hotspots. (Chen et al. 2023a) adopted the topic model and bibliometrics to analyze 351 papers about AI-powered information fusion for smart health, identifying contributors, visualizing collaboration, major research topics, future directions, and distributions of contributors. A summary of relevant reviews is listed in Table 1.

The previously mentioned analyses primarily relied on synthesis or systematic methods. The systematic approaches often entail an arduous coding process, and these studies tended to cover a relatively limited number of articles. In terms of research focus, existing reviews typically concentrated on isolated aspects like multimodal data analysis, information fusion, AI, or healthcare. For instance, (Sujith et al. 2022) explored SHM using DNNs and AI, (Muhammad et al. 2021) tackled multimodal medical signals fusion, and (Guo et al. 2020) and (Albahri et al. 2023) delved into AI for healthcare. Others, such as multimodal data-driven smart healthcare (Cai et al. 2019), multimodal medical data fusion, and multimodal information fusion for smart healthcare (Shaik et al. 2023), addressed specific facets. While (Mohsen et al. 2022) and (Chen et al. 2023a) provided more focused assessments on the combined utilization of AI and fusion technologies in healthcare, a comprehensive analysis considering multimodal data fusion, smart healthcare, and AI simultaneously, especially through quantitative analysis employing machine learning approaches on a large scale, has been lacking. Consequently, the understanding of AI-driven multimodal

**Table 1** A summary of relevant reviews

Studies	Research focus	Methods	Analysis aspects
Muhammad et al. (2021)	multimodal medical signal fusion schemes for smart healthcare systems	Systematic survey	IoMT applications, multi-sensor data fusion levels, and advancements in multimodal medical data fusion
Cai et al. (2019)	multimodal data-powered smart healthcare applications	Survey	Multimodal semantic perception and entity association mining, multimodal data fusion and cross-border knowledge fusion, and panoramic interactive and intelligent decision-making system
Shaik et al. (2023)	multimodal medical data fusion	Survey	Data representation, multiple information integration methods, challenges, and trends associated with multimodal fusion
Albahri et al. (2023)	AI trustworthiness and explainability in healthcare	Systematic review	Legitimacy, morality, and robustness standards for AI policymakers, assessment quality, bias risk, and data fusion for medical AI trustworthiness, research gaps to guide future study and assist researchers
Sujith et al. (2022)	SHM using DNNs and AI	Systematic review	Features of SHM and IoMT, the role of DNNs and AI in SHM, structure of IoTM and SHM, data security in SHM, limitations of SHM and IoMT
Mohsen et al. (2022)	AI for electronic health record and imaging data fusion	Systematic review	Fusion approaches, diseases and clinical outcomes, machine learning for multimodal fusion, and multimodal medical datasets
Guo et al. (2020)	AI for healthcare	Bibliometrics	Research activity characteristics, publication patterns, and research hotspots
Chen et al. (2023a)	AI-powered information fusion for healthcare	Topic model and bibliometrics	Contributors and scientific collaboration, research topics, prospects, and topic distributions

data fusion in smart healthcare remains limited. Vital questions such as major research topics and their evolution, prominent contributing countries/regions, institutions, and authors within this field are yet to be thoroughly explored.

### 1.3 Research objectives and questions

To fill the knowledge gap and facilitate research on AI-powered smart healthcare based on multimodal data analysis, this research conducts a thorough investigation of the present body of literature on multimodal data fusion and AI applications in smart healthcare, using advanced topic modeling and bibliometric analysis techniques. Via this all-encompassing analysis, our objective is to enhance the present comprehension of the cutting-edge within this field and recognize possible pathways for additional exploration and ingenuity.

The focus of this paper centers on two critical aspects: topic modeling and bibliometric analysis. Topic modeling is a powerful NLP technique that aids in the discovery and extraction of latent themes and subjects from vast text corpora. By applying topic modeling to healthcare-related literature, we can identify prevalent research areas, emerging trends, and the interconnections between various topics, shedding light on the current state and prospective pathways for integrating AI and multimodal data analytics in smart healthcare.

Furthermore, bibliometric analysis complements topic modeling by quantitatively assessing the publication patterns, research productivity, and impact of academic literature within this domain. By systematically reviewing scientific publications, citation networks, and collaborations among researchers, we can gain valuable insights into the growth trajectory of this field, key contributors, and the most influential research contributions.

Specially, three research questions (RQs) will be addressed in this study.

RQ1: What are the publication patterns, leading studies, journals, countries/regions, institutions, and authors?

RQ2: How is the scholarly collaboration among countries/regions, institutions, and authors in terms of co-authorship?

RQ3: Which noteworthy topics are addressed, and how do these topics evolve in terms of research prominence over time?

The RQs are formulated by consulting prior bibliometric investigations that, akin to the present study, strive to grasp the research panorama of a domain. Instances encompass research like AI (Vega Hernández et al. 2023), AI in the healthcare sector (Guo et al. 2020), employment of blockchain in management (Tandon et al. 2021), readiness for an ethical AI society (Wamba et al. 2021), football performance analysis (Principe et al. 2022), sentiment analysis (Cui et al. 2023a), and integration of information and AI in smart healthcare (Chen et al. 2023a). As per past literature, addressing these queries can furnish a cutting-edge grasp of research related to the amalgamation of healthcare information with AI, and bestow significant ramifications to researchers and project initiators for its subsequent advancement.

The rationales behind investigating each of these queries are demonstrated as follows. Initially, addressing RQ1 allows scholars to (1) comprehend the worldwide progression of scientific knowledge and the trajectories in the field's advancement (Cui et al. 2023b), (2) apply the outcomes of influential scholarly works, (3) pinpoint suitable platforms for sharing and publishing research concerning the amalgamation of healthcare information with AI (Swacha 2021), and (4) recognize pivotal authors to learn from Oliveira et al. (2019). Subsequently, tackling RQ2 assists in comprehending patterns of collaboration and associations, as well as identifying potential academic partners (Wu et al. 2021a).

Lastly, outcomes derived from RQ3 will aid in understanding the historical and present scholarly landscape concerning the amalgamation of healthcare data with AI, ensuring that researchers are updated about pressing matters that require their focus (Shao et al. 2021). Furthermore, the results also provide insights into the evolving patterns of research themes and provide insight into the potential directions of the area in the times ahead (Mustak et al. 2021). These revelations empower scholars, policymakers, and practitioners to remain cognizant of cutting-edge research while venturing into scientific and technological endeavors (Jeyaraj and Zadeh 2020).

Therefore, the primary contributions of this research to the academic community can be outlined as follows: (1) introduce the first structural topic model (STM)-driven bibliometric analysis of AI-powered smart healthcare based on multimodal data examination, (2) uncover key contributors (countries/regions, institutions, and authors) to share their research insights, (3) visualize collaborations among prominent contributors (countries/regions, institutions, and authors), (4) identify prevalent research topics and potential future paths, (5) enhance comprehension of the historical, current, and forthcoming academic panorama concerning multimodal data fusion employing AI in smart healthcare, and 6) employ topic model-driven bibliometric methodologies for literature assessment, circumventing the constraints of manual coding or qualitative analysis techniques.

## 2 Data and methods

This study followed the three phases outlined in the Flow Diagram of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Moher et al. 2009) to identify, choose, and critically evaluate relevant studies (Pian et al. 2021). The adoption of the PRISMA aimed to certify robustness and reduce bias in the review procedure (Demir 2021). The collected data underwent analysis using topic modeling and bibliometrics, incorporating a statistical trend examination and social network analysis (SNA). An elaborate description of the procedure is presented below.

### 2.1 Literature search

This study systematically searched the literature in the Web of Science (WoS) database. The search terms (Table 2) used in this study were derived from previous reviews related to AI [e.g., (Goodell et al. 2021; Wamba et al. 2021)], multimodal data analysis [e.g., (Lu et al. 2023; Zhang et al. 2020a)], and healthcare/medical [e.g., (Jimma 2023; Guo et al. 2020)].

We conducted searches for the terms in titles, abstracts, and keywords, resulting in a total of 829 initial hits. Among these hits, there were 779 journal articles written in English from Science Citation Index Expanded (SCI-E), and 50 papers from Social Science Citation Index (SSCI) databases.

### 2.2 Selection criteria and study selection

The data search produced a 829 articles, which were then organized in Mendeley. From this collection, 48 duplicate articles were automatically eliminated, leaving 781 unique articles that underwent screening based on specified inclusion/exclusion criteria (Table 3) adapted from Chen et al. (Chen et al. 2023a).

**Table 2** Search terms used in this study

AI-related	“Intelligent agent*” or “expert system*” or “fuzzy logic” or “neuro-fuzzy” or “text mining” or “text-mining” or “deep network*” or “fuzzy system*” or “graph mining” or “markov chain” or “neural fuzzy” or “neural-fuzzy” or “bayes network*” or “decision tree” or “deep learning” or “deep-learning” or “fuzzy control” or “kernel method” or “kinetic model” or “random forest” or “visual search” or “bayes learning” or “feature coding” or “fundus imaging” or “naive bayesian” or “neural control” or “neural network*” or “robust control” or “computer vision” or “features mining” or “fuzzy reasoning” or “graph embedding*” or “hybrid coupling” or “knowledge graph*” or “regression tree*” or “retinal imaging” or “bayesian network*” or “fuzzy clustering” or “fuzzy set theory” or “machine learning” or “machine-learning” or “naive bayes tree*” or “rough set theory” or “semantic mapping” or “genetic algorithm*” or “k-means*” or “nearest neighbor*” or “neural nets model*” or “ontology matching” or “bayesian inference” or “hesitant fuzzy set” or “human intelligence” or “k-means clustering” or “k-nearest neighbour*” or “question answering” or “sentiment analysis” or “speech recognition” or “swarm intelligence” or “affective computing” or “classif* tree” or “deep belief network*” or “fuzzy cognitive map*” or “machine translation” or “markov random field*” or “pattern recognition” or “ambient intelligence” or “human-machine system*” or “machine intelligence” or “semantic orientation” or “tensor-train network*” or “information retrieval*” or “intelligent computing” or “man-machine interface*” or “feature representation” or “gaussian mixture model*” or “information extraction” or “intuitive intelligence” or “reinforcement learning” or “support vector machine” or “support vector network*” or “artificial intelligence” or “association rule mining” or “bayesian belief network*” or “brain-machine interface*” or “decision support system*” or “deep polynomial network*” or “ensemble classif*” or “semantic topic analysis” or “support vectors machine” or “conditional random field*” or “evolutionary computation” or “fuzzy c-means clustering” or “intelligent robot system*” or “named entity recognition” or “networked control system*” or “sentiment classif*” or “bayesian learning” or “feature learning” or “switched network cluster*” or “artificial neural-network*” or “artificial neural network*” or “gradient boosting machine” or “word sense disambiguation” or “computational intelligence” or “deep contour-aware network*” or “naive bayes classif*” or “neural dynamic programming” or “fuzzy classif*” or “natural language generation” or “natural language processing” or “feature selection” or “support vector classif*” or “fuzzy knowledge representation” or “natural language understanding” or “artificial bee colony” or “artificial fish swarm” or “artificial general intelligence” or “fuzzy inference”
AND	
Fusion-related	“Fusion”
AND	
Multimodal-related	“Multimodal*” OR “multi-modal*”
AND	
Medical/health-related	“Medical*” OR “medicine*” OR “medication” OR “health*” OR “healthcare” OR “hospital*” OR “clinical” OR “clinic*” OR “disease*”

To be eligible for inclusion, an article must meet the following criteria: (i) emphasize the utilization of AI-centered methods (for instance, conventional machine learning techniques, deep learning methodologies, logic and measurements), (ii) employ multimodal data fusion methods and mechanisms, and (iii) address medical/health-related issues or topics (e.g., predictive, preventive, personalized, and participatory solutions for smart health, processing and utilization of medical/clinical images, signals, and texts, along with the detection and evaluation of human activities for health/medical intentions). Studies that solely concentrated on gesture, motion, movement recognition, emotion recognition, or wireless sensor applications without relevance to health purposes were excluded. Similarly,



**Table 3** Inclusion and exclusion criteria

Inclusive criteria	AI methods	
	IA1	Machine learning
	IA2	DNNs
	IA3	Reasoning and metrics
	Multimodal data fusion	
	IM1	Multimodal data fusion domains or mechanisms
	Medical/health-related topics	
	IH1	Solutions for smart healthcare (e.g., predictive, personalized)
	IH2	Medical or clinical image processing applications
	IH3	Medical or clinical signal processing applications
	IH4	Medical or clinical NLP applications
	IH5	Medical or clinical integrated processing applications
	IH6	Generic processing applications for health or medical purposes
	IH7	Human activity detection and assessment for health or medical purposes
	IH8	Multimodal analysis for doctor-patient relationship prediction
	IH9	Safety (e.g., food security, fall prevention)
Exclusive criteria	E1	Wireless sensor applications not for health purposes
	E2	Emotion recognition not for health purposes
	E3	Gesture, motion, or movement recognition not for health purposes
	E4	Computer-assisted detection or diagnosis without AI
	E5	Food, animal, plant, chemical, biometric, microbial, or agriculture-related topics
	E6	Environment and smart home-related topics (e.g., underwater acoustic target recognition)
	E7	Studies that lacked original data (e.g., commentaries, opinion papers, theoretical papers)
	E8	Not full articles (e.g., conference abstracts, brief reports)
	E9	Written in other languages than English

studies that centered on computer-aided detection/diagnosis but did not utilize AI methods or technologies were also excluded. Moreover, studies focusing on food, animals, plants, chemical, biometric, microbial, environment, agriculture, or smart home-related issues were excluded. Finally, research lacking primary data (such as editorials, commentaries, viewpoint articles, and theoretical pieces) or not composed in the English language were also omitted.

The 781 articles underwent two rounds of screening. In the first round of screening, the title and abstract and each of the 781 articles were filtered by two authors (X.C. and X.T.) based on the inclusion/exclusion criteria. The screening process entails three stages. Initially, when determining the inclusion of an article, we first focus on AI methodologies. Articles not aligning with AI methods were promptly excluded without further evaluation of other criteria linked to multimodal data fusion or medical/health-related issues. Any article involving technologies associated with machine learning (IA1), DNNs (IA2), or reasoning and metrics (IA3) was classified as AI method-related. In the subsequent stage, each article categorized as AI method-related underwent assessment for relevance to multimodal data fusion. Those lacking relevance were excluded outright without consideration of other criteria pertaining to medical/health-related topics. Articles involving technologies within multimodal data fusion domains/mechanisms (IM1) were classified as multimodal data fusion-related. In the final stage, each article identified as multimodal data fusion-related was scrutinized for relevance to medical/health-related topics. Articles not meeting this criterion were excluded. Articles touching upon areas like solutions for smart healthcare (IH1), medical or clinical image processing applications (IH2), medical or clinical signal processing applications (IH3), medical or clinical NLP applications (IH4), medical or clinical integrated processing applications (IH5), generic processing applications for health or medical purposes (IH6), human activity detection and assessment for health or medical purposes (IH7), multimodal analysis for doctor-patient relationship prediction (IH8), and safety (e.g., food security, fall prevention) (IH9) were classified as medical or health-related. Whenever discrepancies occurred, the third author (H.X.) was consulted to decide if the article should be selected. A total of 53 articles were reduced, resulting in 728 articles that underwent the second round of screening.

In the second round of screening, a thorough evaluation of the 728 articles' eligibility was carried out through full-text reading based on the same screening process by the three authors (X.C., X.T., and H.X.) to finalize the inclusion of 683 articles. The inter-rater agreement reached 97%, and any discrepancies were resolved through consensus discussions. As per (Huang et al. 2022a), the screening procedure detailed above serves as a method of quality control, ensuring a thorough evaluation of articles using clearly outlined inclusion and exclusion criteria. The PRISMA flow diagram in Fig. 1 presents each stage of the search and the paper selection process.

Quality assessment, often a crucial step in systematic literature reviews to ensure comprehensive coverage of predefined dimensions for analysis, wasn't utilized in this research. This choice arose from employing a topic modeling-based bibliometric approach, distinct from systematic analysis that scrutinizes specific, predetermined dimensions within a restricted article set. The study's objective is to analyze bibliometric attributes and research topics within the 683 eligible articles using automated machine modeling, diverging from the reliance on predefined codes or categories characteristic of systematic analysis or meta-analysis studies.

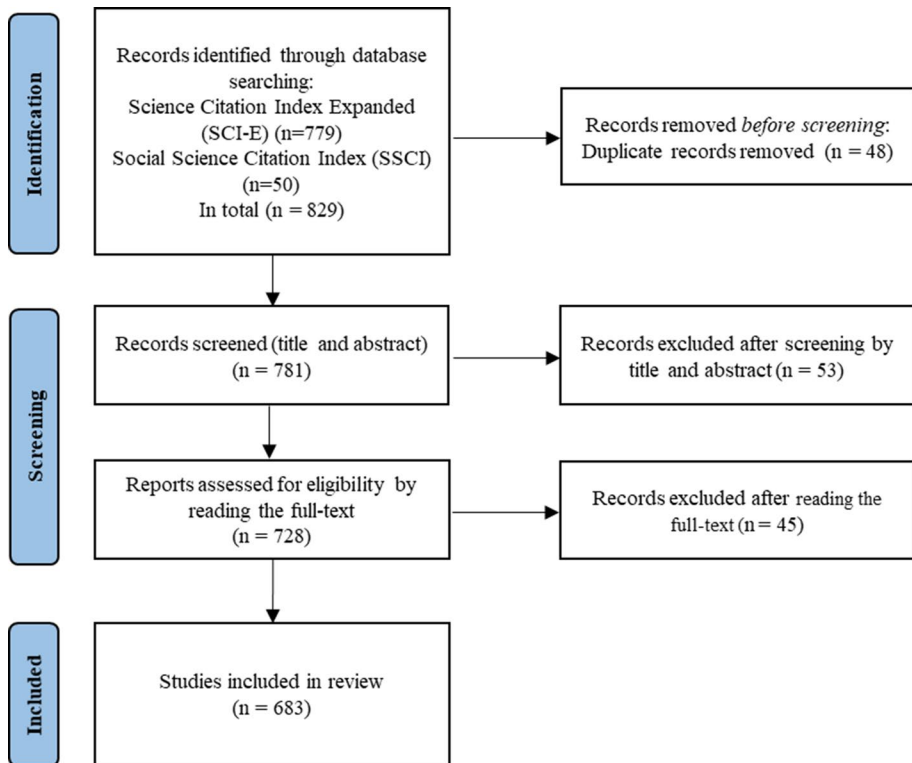


Fig. 1 An overview of data search

### 2.3 Data analysis for answering research questions

From the pool of 683 papers selected for their relevance to AI-powered multimodal data fusion in smart healthcare, we pursued answers to the three RQs. Employing a topic modeling-driven bibliometric analysis, our data analysis predominantly delved into the metadata or bibliographic details linked to each paper, encompassing authors, titles, abstracts, keywords, publication year, and journal titles. Nonetheless, the textual content within the 683 papers was not utilized or reviewed, as our study's aim did not involve pinpointing specific elements akin to systematic reviews. This divergence stems from our analytical approach—STM-based bibliometrics—which focuses on uncovering latent themes within extensive datasets rather than relying on manual coding through predetermined schemes.

RQ1 was investigated through a quantitative analysis of article and citation counts over the years. To capture non-linear trends in the annual paper output, a polynomial modeling approach was employed. The academic performance of journals, nations/regions, academic institutions, and researchers was evaluated using bibliometric measures, including the Hirsch index (H-index) and average citations per article (ACP). The productivity and impact of actors were assessed according to the number of articles contributed and citations earned. The H-index was employed to assess contributors from both a standpoint

of excellence and volume, while ACP was calculated by dividing the citation count by the article count.

To respond to RQ2, Gephi (Bastian et al. 2009) and SNA were used to visualize the connections between researchers, institutions, or nations/regions as separate entities. In a cooperative network of institutions, for instance, the size of each node represented the institution’s productivity, and the link’s width between nodes indicated the degree of their cooperation.

Addressing RQ3, the study utilized topic modeling and keyword analysis techniques. In addition to predefined keywords, phrases extracted from paper titles and abstracts were included for keyword analysis. By ranking these phrases and keywords based on their frequency in the corpus of evaluated publications, frequently researched study subjects were identified. This research employed the intricate STM (Chen et al. 2023b; Roberts et al. 2014) as an advanced probabilistic technique for topic modeling, as demonstrated by its plate diagram presented in Fig. 2. In the context of an article collection indexed by  $D$ , wherein vector  $x_d$  denotes article-level covariates,  $K$  signifies the number of topics, the total word count in a sampled article  $d$  is denoted as  $\{w_{d,n}\}$ , and  $V$  represents the size of the vocabulary. The fundamental pseudocode of the STM involves a tripartite process. The initial phase is dedicated to the estimation of the topic prevalence parameter  $\theta_d$  for every word within article  $d$ , accomplished by logistic-normal generalized linear modeling. Subsequently, the second phase centers on approximating the topical content model  $\beta$ , which portrays words as a probabilistic blend of each topic ( $k$ ). In the third and final step, for each word within the article ( $n \in \{1, \dots, N_d\}$ ), each topic is selected through sampling from a multinomial distribution over  $\theta_d$ .

In this study, the modeling process involved three primary steps. First, terms from titles, abstracts, and keywords, were pre-processed by removing numbers, punctuation, and stop-words. Next, using the term frequency-inverse document frequency technology, unimportant words were filtered out based on a threshold of 0.05. Within the pool of potential models encompassing topic quantities spanning from 5 to 30, models with higher performance in terms of semantic coherence and exclusivity were manually compared and evaluated. Then, the Mann–Kendall (MK) test was utilized to analyze the annual proportions of topics and identify significant increasing/decreasing tendencies

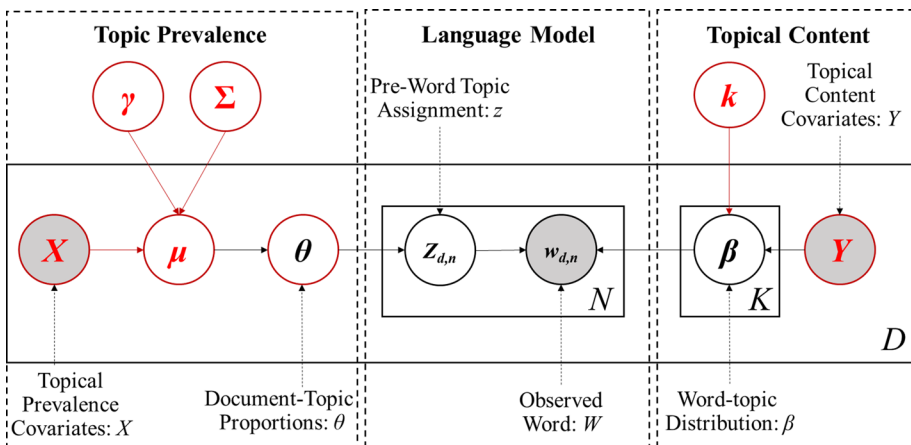


Fig. 2 Plate diagram of STM (Adapted from (Chen et al. 2023b) and (Roberts et al. 2014))

( $p < 0.05$ ). The MK test was also used to examine the evolution of discovered keywords and phrases (Mann 1945).

### 3 Results

#### 3.1 Publication trend

The distributions of articles and citations per year are visualized in Fig. 3. A polynomial regression analysis was performed to fit their patterns. The timeframe from 2002 to 2022 was employed. The outcomes illustrate a discernible inclination toward growth in publication count, particularly noticeable since 2013, reflecting an increasing fervor for research within this domain. This is further supported by the positive coefficient of “ $x^2$ ” in the calculated regression model ( $R^2 = 0.8066$ ). The forecasted value for the year 2023 stands at  $y1 = 0.9920057 \times 2023^2 - 3985.488 \times 2023 + 4003012 = 181.9068$ . There has been a significant surge in citations, particularly since 2013. This trend is likewise highlighted by the positive coefficient of “ $x^2$ ” in the computed model ( $R^2 = 0.8045$ ). In a parallel manner, a regression model was also employed to predict the trajectory of citation variation, yielding  $y2 = 19.16552 \times 2023^2 - 77002.1 \times 2023 + 77343270 = 3453.002$ .

#### 3.2 Top studies

Based on the number of citations, the highest-ranked 10 articles out of the 683 papers related to the integration of healthcare information with AI are showcased in Table 4. Notably, 5 articles feature both ranking compilations [i.e., (Suk et al. 2014; Zhang et al. 2020b; Katsigiannis and Ramzan 2017; Yin et al. 2018; Shi et al. 2017)]. The main

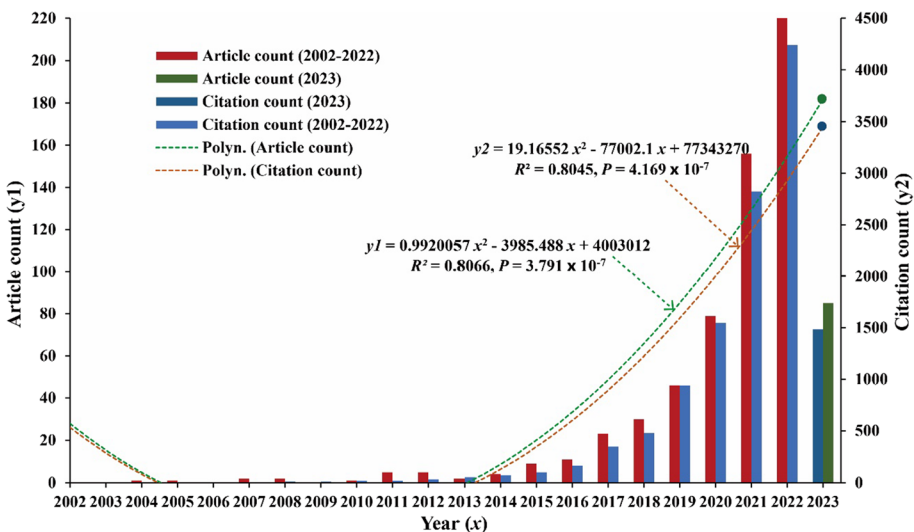


Fig. 3 Trend analysis of article count

content of these five studies is elaborated as follows. The paper by Suk et al. (Suk et al. 2014) received a total of 492 citations with a C/Y value of 49.2, proposing a novel deep learning method that employed a deep Boltzmann machine (DBM) to uncover latent hierarchical features within three-dimensional (3D) patches of neuroimaging data and then created a combined feature depiction for paired MRI and positron emission tomography (PET) patches using a multimodal DBM. The method outperformed previous approaches, achieving high accuracies in categorizing Alzheimer's disease (AD), mild cognitive impairment (MCI), and healthy control subjects. The paper by Zhang et al. (Zhang et al. 2020b) received a total of 328 citations with a C/Y value of 82, presenting IFCNN, an innovative image-blending framework grounded in convolutional neural networks (CNNs). It captured significant attributes from various input images using dual convolutional layers and amalgamated them utilizing elementwise-max, elementwise-min, or elementwise-mean principles depending on the image characteristics. The architecture was entirely convolutional, enabling comprehensive end-to-end training without necessitating post-processing steps. The paper by (Katsigiannis and Ramzan 2017) received a total of 300 citations with a C/Y value of 50, proposing a database containing multiple modes of data encompassing electroencephalography (EEG) and electrocardiogram (ECG) signals, which were captured while eliciting emotions through audio-visual stimuli. The signals, along with self-assessments of participants' emotional states, were captured using portable, wearable, and low-cost equipment, enabling potential use in everyday applications. The paper by (Yin et al. 2018) received a total of 275 citations with a C/Y value of 55, introducing a new approach for fusing multimodal medical images using nonsampled shearlet transform (NSST). This involved NSST decomposition for multiscale and multidirectional representations. High- and low-frequency bands were fused with parameter-adaptive pulse-coupled neural networks and a tactic addressing energy preservation and detail extraction, respectively. The fused images were rebuilt through reverse NSST. The paper by (Shi et al. 2017) received a total of 234 citations with a C/Y value of 39, introducing an algorithm called multimodal deep polynomial networks stacked (MM-DPSN) designed for diagnosing AD. The MM-SDPN was composed of dual-stage stacked deep polynomial networks, which independently acquired advanced features from MRI and PET, and subsequently integrated them to enhance feature representation.

According to total citations, another five studies among the top 10 papers are introduced as follows. The paper by (Liu et al. 2014) received a total of 299 citations, presenting an original diagnostic structure utilizing a deep learning architecture aimed at aiding AD detection. The framework employed a zero-masking tactic for merging data and extracting supplementary insights from several data modalities. By enabling the efficient merging of neuroimaging features from multiple modes and potentially needing fewer labeled datasets, the structure exhibited enhanced outcomes in binary/multiclass AD classification. The paper by Hu et al. (Hu et al. 2018) received a total of 188 citations, presenting a method to infer voxel-level transformation using anatomical labels, which were more practical and reliable to obtain than voxel-level correspondence. They employed a CNN for displacement field prediction for aligning labeled structures in image pairs during training, and the method ran in real-time during inference without requiring labels or initialization. The paper by Zhu et al. (Zhu et al. 2015) received a total of 163 citations, presenting a consolidated structure that merged two distinct subspace learning methods: linear discriminant analysis, and locality-preserving projection. This was employed to pick features that were both distinct to the classes and

**Table 4** Top studies

Studies	Title	C
Suk et al. (2014)	“Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis”	492
Zhang et al. (2020b)	“IFCNN: a general image fusion framework based on convolutional neural network”	328
Katsigiannis and Ramzan (2017)	“DREAMER: a database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices”	300
Liu et al. (2014)	“Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer’s disease”	299
Yin et al. (2018)	“Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampling shearlet transform domain”	275
Shi et al. (2017)	“Multimodal neuroimaging feature Learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer’s disease”	234
Hu et al. (2018)	“Weakly-supervised convolutional neural networks for multimodal image registration”	188
Zhu et al. (2015)	“Subspace regularized sparse multitask learning for multiclass neurodegenerative disease identification”	163
Wang and Ma (2008)	“Medical image fusion using m-PCNN”	162
Ma et al. (2018)	“Deep coupling autoencoder for fault diagnosis with multimodal sensory data”	154

C citation count

robust against noise. The suggested approach demonstrated effectiveness in the context of multiclass classification and surpassed alternative cutting-edge techniques on the AD Neuroimaging Initiative dataset. The paper by Wang and Ma (Wang and Ma 2008) received a total of 162 citations, presenting a multi-channel Pulse coupled neural network (m-PCNN) to fuse medical images. The article elucidated the mathematical foundation of m-PCNN and introduced the dual-channel model as an exceptional instance. The outcomes showcased that m-PCNN surpassed alternative approaches regarding visual impact and objective assessment. The paper by (Ma et al. 2018) received a total of 154 citations, introducing the deep coupling autoencoder (DCAE) for fault detection using multimodal sensory signals. The DCAE seamlessly integrated feature extraction and data fusion, capturing shared information between different sensory data and learning higher-level joint features.

According to the yearly citation rate (C/Y) (Chen et al. 2022), another five studies among the top ten papers are presented in Table 5. The paper by (Holzinger et al. 2021) received a C/Y value of 45.33, advocating the use of graph neural networks (GNNs) for multimodal causability. GNNs were highlighted as essential for multimodal causability because they could directly define causal links between features using graph structures. The paper by (Chen et al. 2020) received a C/Y value of 41.5, introducing pathomic fusion, a comprehensible approach for merging histology images and genomic features to predict survival outcomes. This technique captured mutual feature correlations using a Kronecker product and controlled representation expressiveness through a gate-oriented attention mechanism. The method allowed for feature interpretation and localization across each modality and improved prognostic determinations compared to unimodal deep networks. The paper by (Khan et al. 2020) received a C/Y value of 38.5, presenting an automated approach for classifying brain tumor types utilizing deep learning and multimodal data. The method involved five fundamental stages: applying linear contrast enhancement, performing feature extraction through deep learning utilizing VGG16 and VGG19 architectures, implementing correntropy-based collaborative learning in tandem with extreme learning machine (ELM) for feature curation, fusion of robust covariant features using partial least square, and final classification with ELM. The paper by (Jin et al. 2022) received a C/Y value of 36.5, evaluating choroidal neovascularization in age-related macular degeneration using multimodal DNNs with optical coherence tomography (OCT) and angiography images, enhancing computer-aided diagnosis systems. The paper by (Venugopalan et al. 2021) received a C/Y value of 35.33, improving AD and MCI analysis by integrating imaging, genetic, and clinical data. Deep learning techniques, including stacked denoising autoencoders and 3D CNNs, classified patients into AD, MCI, and controls, surpassing shallow models. The study identified the hippocampus, amygdala, and Rey Auditory Verbal Learning Test as key features, aligning with AD literature.

### 3.3 Journal analysis

A total of 233 journals were found. In Table 6, the leading 15 journals in terms of article quantity are presented, ranked from the complete relevant article pool (683). The primary trio of journals comprises Biomedical Signal Processing and Control, IEEE Journal of Biomedical and Health Informatics, and IEEE Access. These journals encompass 36.60 percent of the aggregate articles, with the foremost 3 accounting for 13.03 percent. When considering both the H-index and citation counts, IEEE Transactions on Medical Imaging and IEEE Journal of Biomedical and Health Informatics hold positions within the top 3.



**Table 5** Top studies

Studies	Title	C/Y
Zhang et al. (2020b)	“IFCNN: a general image fusion framework based on convolutional neural network”	82.0
Shi et al. (2017)	“Multimodal neuroimaging feature Learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer’s disease”	55.0
Katsigiannis and Ramzan (2017)	“DREAMER: a database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices”	50.0
Suk et al. (2014)	“Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis”	49.2
Holzinger et al. (2021)	“Towards multi-modal causability with graph neural networks enabling information fusion for explainable AI”	45.3
Chen et al. (2020)	“Pathomic fusion: an integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis”	41.5
Shi et al. (2017)	“Multimodal neuroimaging feature Learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer’s disease”	39.0
Khan et al. (2020)	“Multimodal brain tumor classification using deep learning and robust feature selection: a machine learning application for radiologists”	38.5
Jin et al. (2022)	“Multimodal deep learning with feature level fusion for identification of choroidal neovascularization activity in age-related macular degeneration”	36.5
Venugopalan et al. (2021)	“Multimodal deep learning models for early detection of Alzheimer’s disease stage”	35.3

C/Y annual citations

**Table 6** Top journals

Journals	AC	TC	ACP	H	MK test	2002–2012			2013–2017			2018–2023		
						Trend	AC	TC	AC	TC	AC	TC	AC	TC
Biomedical signal processing and control	31	237	7.65	8	0.004	72	↑↑↑	0	0	1	5	30	232	
IEEE journal of biomedical and health informatics	30	941	31.37	10	0.002	83	↑↑↑	0	0	1	4	29	937	
IEEE access	28	373	13.32	10	0.007	62	↑↑↑	0	0	0	0	28	373	
Frontiers in neuroscience	20	96	4.80	6	0.006	68	↑↑↑	0	0	1	9	19	87	
Computers in biology and medicine	18	52	2.89	3	0.005	57	↑↑↑	0	0	0	0	18	52	
Medical image analysis	17	423	24.88	9	0.004	77	↑↑↑	0	0	1	5	16	418	
Sensors	15	129	8.60	6	0.004	71	↑↑↑	0	0	1	0	14	129	
Information fusion	14	800	57.14	8	0.080	44	↑	1	16	1	82	12	702	
IEEE transactions on medical imaging	13	538	41.38	10	0.008	71	↑↑↑	0	0	3	4	10	534	
Medical physics	12	187	15.58	6	0.040	55	↑↑	1	1	3	23	8	163	
Multimedia tools and applications	11	73	6.64	4	0.007	55	↑↑↑	0	0	0	0	11	73	
Neurocomputing	11	302	27.45	8	0.003	73	↑↑↑	0	0	1	2	10	300	
International journal of imaging systems and technology	10	58	5.80	3	0.024	52	↑↑	0	0	1	0	9	58	
Journal of healthcare engineering	10	43	4.30	4	0.002	70	↑↑↑	0	0	0	0	10	43	
Neural computing and applications	10	218	21.80	5	0.013	62	↑	1	0	0	2	9	216	

R ranking, H H-index, AC article count, TC citation count, ACP average citations per article, increasing (decreasing) trend but not significant ( $p > 0.05$ ); ↑↑↑(↓↓↓), ↑↑↑(↓↓↓), ↑↑↑(↓↓↓): significantly increasing (decreasing) trend ( $p < 0.05$ ,  $p < 0.01$ , and  $p < 0.001$ , respectively); S MK test statistics

By utilizing the yearly article count spanning 2002 to 2022 and employing the MK trend assessment, we calculated the overall significance level and trend for each journal, denoted by upward (↑) or downward (↓) indicators. The higher the count of these symbols, the more notable the trend. Except for Information Fusion and Neural Computing and Applications, all the other 13 listed journals demonstrate a dramatically increasing trend. Additionally, we divided the two decades into three intervals: 2002–2012, 2013–2017, and 2018–2023, as depicted in Table 6.

### 3.4 WOS categories

A total of 82 WoS categories were found. Table 7 displays the foremost 15 WoS categories with the highest article counts, arranged according to the complete pool of relevant articles (683). The leading quintet encompasses “engineering, electrical & electronic”, “computer science, artificial intelligence”, “engineering, biomedical”, “computer science, interdisciplinary application”, and “computer science, information system”. From both the H-index and citation count standpoints, “engineering, electrical & electronic”, “engineering, biomedical”, and “radiology, nuclear medicine & medical imaging” stand within the highest three positions.

Using the yearly count of articles spanning from 2002 to 2022 and applying the MK trend test, we calculated the overall significance level and the trend for each WoS category. Except for five categories: “engineering, electrical & electronic”, “computer science, artificial intelligence”, “engineering, biomedical”, “radiology, nuclear medicine & medical imaging”, and “computer science, theory & method”, all the other 10 listed categories exhibit a notably ascending trend. Additionally, we divided the 20 years into three phases: 2002–2012, 2013–2017, and 2018–2023, as illustrated in Table 7.

### 3.5 Top countries/regions, institutions, and authors

In the span from 2002 to 2023, a cumulative of 683 articles were published in the research, involving a comprehensive representation of 57 countries/regions. These articles demonstrate a wide array of geographic origins, as illustrated by Table 8 which lists the top 15 countries/regions based on their article count. Among these, the leading trio comprises China, USA, and India. Remarkably, China stands out with 371 articles, making up around 54.32% of the entire article count, followed by the USA (125 articles, 18.30%), and India (73 articles, 10.69%). Ranked by both citation count and H-index, China (4203 citations, H index of 34) and the USA (3715 citations, H index of 37) emerge as the leading two countries/regions in this field of study.

Through employing the annual count of articles within the timeframe of 2002–2022 and applying the MK trend test, computations were carried out for the general significance level and the trend for each country/region collectively. Except for four countries/regions: the USA, the UK, Italy, and France, all the remaining 11 listed countries/regions portray a conspicuously increasing trend. The table further outlines the article and citation counts of countries/regions within the three periods: 2002–2012, 2013–2017, and 2018–2023. The findings demonstrate that China exhibited substantial progress in ranking during the most recent period.

During the period spanning from 2002 to 2023, a collective of 683 research articles engaged a cumulative of 962 institutions in their publication. These articles reflect an array of geographic origins, with Table 9 providing insight into the top 17 institutions based on

**Table 7** Top WoS categories

WOS categories	AC	TC	ACP	H	MK test		2002–2012		2013–2017		2018–2023		
					<i>p</i>	<i>S</i>	Trend	AC	TC	AC	TC	AC	TC
Engineering, electrical & electronic	165	2983	18.08	29	0.143	48	↑	3	10	10	190	152	2783
Computer science, artificial intelligence	147	2864	19.48	26	0.927	4	↑	5	0	14	56	128	2808
Engineering, biomedical	146	3080	21.10	32	0.093	55	↑	4	22	16	202	126	2856
Computer science, interdisciplinary application	112	2059	18.38	20	0.020	72	↑↑	2	17	7	183	103	1859
Computer science, information system	109	1755	16.10	23	0.014	72	↑↑	1	49	4	179	104	1527
Radiology, nuclear medicine & medical imaging	99	2967	29.97	33	0.101	53	↑	5	15	16	166	78	2786
Mathematical & computational biology	75	1362	18.16	17	0.005	84	↑↑↑	3	16	2	100	70	1246
Neuroscience	60	942	15.70	18	0.002	92	↑↑↑	0	12	7	94	53	836
Medical informatics	57	1381	24.23	13	0.015	73	↑↑	2	18	3	195	52	1168
Computer science, theory & method	46	1212	26.35	13	0.404	25	↑	3	0	1	93	42	1119
Telecommunication	43	641	14.91	13	0.002	76	↑↑↑	0	7	0	78	43	556
Instruments & instrumentation	35	585	16.71	12	0.002	88	↑↑↑	1	3	2	33	32	549
Imaging science & photographic technology	34	699	20.56	14	0.006	80	↑↑↑	0	0	5	19	29	680
Computer science, software engineering	25	198	7.92	6	0.004	66	↑↑↑	0	0	0	2	25	196
Biology	21	222	10.57	4	0.024	52	↑↑	1	0	0	0	20	222

Similar to Table 6

**Table 8** Top productive countries/regions

Countries/regions	AC	TC	ACP	H	MK test			2002–2012		2013–2017		2018–2023	
					<i>p</i>	<i>S</i>	Trend	AC	TC	AC	TC	AC	TC
China	371	4203	11.33	34	0.002	96	↑↑↑	2	22	17	262	352	3919
USA	125	3715	29.72	37	0.367	30	↑	8	10	18	161	99	3544
India	73	1240	16.99	17	0.009	81	↑↑↑	3	39	4	37	66	1164
UK	55	1591	28.93	16	0.069	55	↑	2	24	3	231	50	1336
South Korea	39	774	19.85	14	0.020	68	↑↑	1	0	2	20	36	754
Saudi Arabia	29	308	10.62	10	0.000	94	↑↑↑↑	0	26	1	17	28	265
Australia	26	488	18.77	12	0.008	71	↑↑↑	0	3	2	27	24	458
Egypt	26	353	13.58	9	0.010	64	↑↑	0	0	1	40	25	313
Germany	25	602	24.08	12	0.006	77	↑↑↑	0	0	3	56	22	546
Canada	21	642	30.57	11	0.011	74	↑↑	0	0	5	39	16	603
Spain	20	644	32.20	11	0.040	60	↑↑	1	0	2	0	17	644
Italy	17	662	38.94	7	0.112	40	↑	1	0	1	84	15	578
Pakistan	15	178	11.87	7	0.005	64	↑↑↑	0	0	0	1	15	177
France	13	485	37.31	7	0.080	44	↑	1	0	3	4	9	481
Turkey	10	69	6.90	4	0.012	51	↑↑	0	0	0	5	10	64

Similar to Table 6

their article count. Positioned at the helm, the Chinese Academy of Sciences and Shanghai Jiao Tong University secure the top two spots among these institutions, contributing together to 49 articles, followed by Fudan University (16 articles), Sichuan University (15 articles), and Chongqing University of Posts and Telecommunications (14 articles). Ranked by both citation count and H-index, Chinese Academy of Sciences (240 citations, H index of 9), University of North Carolina at Chapel Hill (332 citations, H index of 9), and Korea University (330 citations, H index of 9) are the top 3 institutions in this research area.

Through utilizing the yearly article count within the 2002–2022 timeframe and implementing the MK trend test, calculations were conducted for both the overall significance level and trend pertaining to each institution collectively. All the 17 institutions featured on the list display a highly prominent ascending trend. The table further provides an overview of the article and citation counts for institutions during three distinct periods: 2002–2012, 2013–2017, and 2018–2023. The findings highlight that numerous institutions have notably elevated their rankings in the most recent period, with Fudan University, Sichuan University, and Chongqing University of Posts and Telecommunications notably excelling.

In the span from 2002 to 2023, a cumulative of 683 research articles witnessed the contribution of 3170 authors. These articles emanate from a multitude of geographic origins, as evident in the top 17 authors enumerated in Table 10. The top three authors are Dinggang Shen from ShanghaiTech University (12 articles), Xia-An Bi from Hunan Normal University (9 articles), and Yu Liu from Hefei University of Technology (9 articles), followed by Tamer Abuhmed from Sungkyunkwan University (7 articles), and Shaker El-Sappagh from Benha University (7 articles). Dinggang Shen is also the top author ranked by H-index. Ranked by citation count, Anant Madabhushi from Emory University (494

**Table 9** Top institutions

Institutions	Countries/regions										MK test			2018–2023	
	AC	TC	ACP	H	P	S	Trend	2002–2012		2013–2017		AC	TC		
								AC	TC	AC	TC				
Chinese Academy of Sciences	25	240	9.60	9	0.004	77	↑↑↑	0	6	2	35	23	199		
Shanghai Jiao Tong University	24	173	7.21	7	0.004	72	↑↑↑	0	5	1	5	23	163		
Fudan University	16	154	9.63	5	0.004	66	↑↑↑	0	12	0	51	16	91		
Sichuan University	15	122	8.13	4	0.012	51	↑↑	0	0	0	0	15	122		
Chongqing University of Posts and Telecommunications	14	69	4.93	4	0.005	64	↑↑↑	0	0	0	0	14	69		
Hefei University of Technology	12	84	7.00	3	0.002	70	↑↑↑	0	0	0	0	12	84		
Northeastern University	12	106	8.83	6	0.016	49	↑↑	0	0	0	12	12	94		
University of North Carolina at Chapel Hill	12	332	27.67	9	0.010	69	↑↑↑	0	0	2	14	10	318		
University of Science and Technology of China	12	97	8.08	5	0.004	66	↑↑↑	0	0	0	0	12	97		
Korea University	11	330	30.00	9	0.031	54	↑↑	0	0	2	14	9	316		
University of Chinese Academy of Sciences	11	122	11.09	5	0.004	66	↑↑↑	0	6	0	35	11	81		
Jilin University	10	118	11.80	5	0.002	76	↑↑↑	0	0	0	13	10	105		
Peking University	10	66	6.60	4	0.009	53	↑↑↑	0	1	0	21	10	44		
Shanghai University	10	102	10.20	4	0.004	66	↑↑↑	0	0	0	7	10	95		
Sungkyunkwan University	10	59	5.90	5	0.007	55	↑↑↑	0	0	0	0	10	59		
University of Electronic Science and Technology of China	10	130	13.00	4	0.004	66	↑↑↑	0	0	0	5	10	125		
Zhejiang University	10	26	2.60	3	0.031	37	↑↑	0	0	0	0	10	26		

Similar to Table 6

**Table 10** Top authors

Authors	Institutions	AC			MK test			2002–2012		2013–2017		2018–2023			
		AC	TC	ACP	H	P	S	Trend	AC	TC	AC	TC	AC	TC	
Shen, Dinggang	ShanghaiTech University	12	331	27.58	9	0.031	54	↑↑	0	0	2	14	10	317	
Bi, Xia-An	Hunan Normal University	9	65	7.22	4	0.003	68	↑↑↑	0	0	0	0	0	9	65
Liu, Yu	Hefei University of Technology	9	79	8.78	3	0.012	51	↑↑	0	0	0	0	0	9	79
Abuhmed, Tamer	Sungkyunkwan University	7	40	5.71	3	0.005	57	↑↑↑	0	0	0	0	0	7	40
El-Sappagh, Shaker	Benha University	7	40	5.71	3	0.005	57	↑↑↑	0	0	0	0	0	7	40
Chen, Xun	University of Science & Technology of China	6	39	6.50	2	0.012	51	↑↑	0	0	0	0	0	6	39
Gupta, Deep	Visvesvaraya National Institute of Technology Nagpur	6	43	7.17	4	0.007	62	↑↑↑	0	0	0	0	0	6	43
Li, Weisheng	Chongqing University of Posts and Telecommunications	6	35	5.83	3	0.007	62	↑↑↑	0	0	0	0	0	6	35
Calhoun, Vince D	Georgia State University	5	132	26.40	5	0.038	52	↑↑	0	0	2	0	0	3	132
Dai, Yin	Northeastern University	5	84	16.80	4	0.073	31	↑	0	0	0	0	10	5	74
Khan, Muhammad Attique	HITEC University Taxila	5	46	9.20	3	0.009	53	↑↑↑	0	0	0	0	0	5	46
Madabhusi, Anant	Emory University	5	494	98.80	5	0.802	6	↑	4	0	1	0	0	0	494
Singh, Sneha	Indian Institute of Technology	5	90	18.00	4	0.028	45	↑↑	0	0	0	0	0	5	90
Wang, Tianfu	Shenzhen University	5	205	41.00	4	0.016	60	↑↑	0	0	2	30	3	175	
Zhang, Yu	Southern Medical University	5	76	15.20	4	0.042	35	↑↑	0	1	0	0	32	5	43
Zhou, Tao	North Minzu University	5	106	21.20	3	0.021	47	↑↑	0	0	0	0	14	5	92
Zhu, Xiaofeng	University of Electronic Science and Technology of China	5	417	83.40	4	0.351	22	↑	1	0	1	0	1	0	417

Similar to Table 6

citations), Xiaofeng Zhu from University of Electronic Science and Technology of China (417 citations), and Dinggang Shen (331 citations) are the top 3 authors in this research area.

Utilizing the annual article count spanning from 2002 to 2022 and employing the MK test, computations were performed to determine the overall significance level and trend for each author collectively. Except for Vince D. Calhoun from Georgia State University, Anant Madabhushi from Emory University, and Xiaofeng Zhu from the University of Electronic Science and Technology of China, the remaining 14 authors among those listed exhibit a remarkable upward trend. The table also offers a broad perspective on the article and citation counts of authors across three distinct timeframes: 2002–2012, 2013–2017, and 2018–2023. The findings underscore that numerous authors have significantly enhanced their rankings in the latest period, particularly Xia-An Bi, Yu Liu, Tamer Abuhmed, and Shaker El-Sappagh.

### 3.6 Scientific collaboration analysis

In Fig. 4a illustrates the partnerships involving 7 countries/regions, with collaborative frequencies spanning from 9 to 52. Among these, 3 belong to Asia (pink nodes). Notably, the USA and China demonstrated the strongest collaboration, appearing in 52 articles. This was followed by collaborations between China and the UK (18), South Korea and the USA (14), and the UK and the USA (12). In Fig. 4b, collaborations among 10 countries/regions are showcased, featuring collaborative frequencies ranging from 6 to 7. Of the 10 countries/regions, 4 are from Asia (pink nodes) and 3 are from Europe (green nodes). The collaboration among them is close as witnessed by three collaborative clusters: 2) China and Hong Kong, 2) Italy, the UK, and Germany, and 3) Saudi Arabia and Pakistan. Figure 4c displays the partnerships among 7 countries/regions, characterized by a collaborative frequency of 5. Of the 7 countries/regions, 4 are from Asia (pink nodes). They maintain a tight collaborative relationship, as witnessed by the collaborative cluster formed by China, India, and Saudi Arabia. In Fig. 4d, the partnerships involving 9 countries/regions are illustrated, marked by a collaborative frequency of 4. Of the 9 countries/regions, 4 are from Asia (pink nodes) and 2 are from North America (orange nodes). The collaboration among them is close as witnessed by 2 collaborative clusters: 2) the USA, Canada, and Saudi Arabia and 2) Pakistan and South Korea. Figure 4e visualizes the partnerships among 14 countries/regions, characterized by a collaborative frequency of 3. Of the 14 countries/regions, 7 are from Europe (green nodes) and 6 are from Asia (pink nodes). The collaboration among them is closed as witnessed by China, Singapore, Macao, Norway, German, France, the UK, and Iran.

In Figs. 5 and 6 illustrate partnerships among institutions, exhibiting collaborative frequencies spanning from 3 to 10. In Fig. 5a, partnerships among 4 institutions are represented, marked by a collaborative frequency of 10. Of the 4 institutions, 2 are from China (green nodes). These 4 institutions form two clusters of collaboration: (1) Korea University and University of North Carolina at Chapel Hill, and (2) University of Chinese Academy of Sciences and Chinese Academy of Sciences. In Fig. 5b, the partnerships among 7 institutions are showcased, characterized by collaborative frequencies that range from 5 to 6. Of the 7 institutions, 4 are from China (green nodes) and 2 are from South Korea (orange nodes). These 7 institutions form 3 clusters of collaboration: (1) Hefei University of Technology and University of Science and Technology of China, (2) Sichuan University and Chengdu University of Information Technology, and (3) Sejong University, Benha



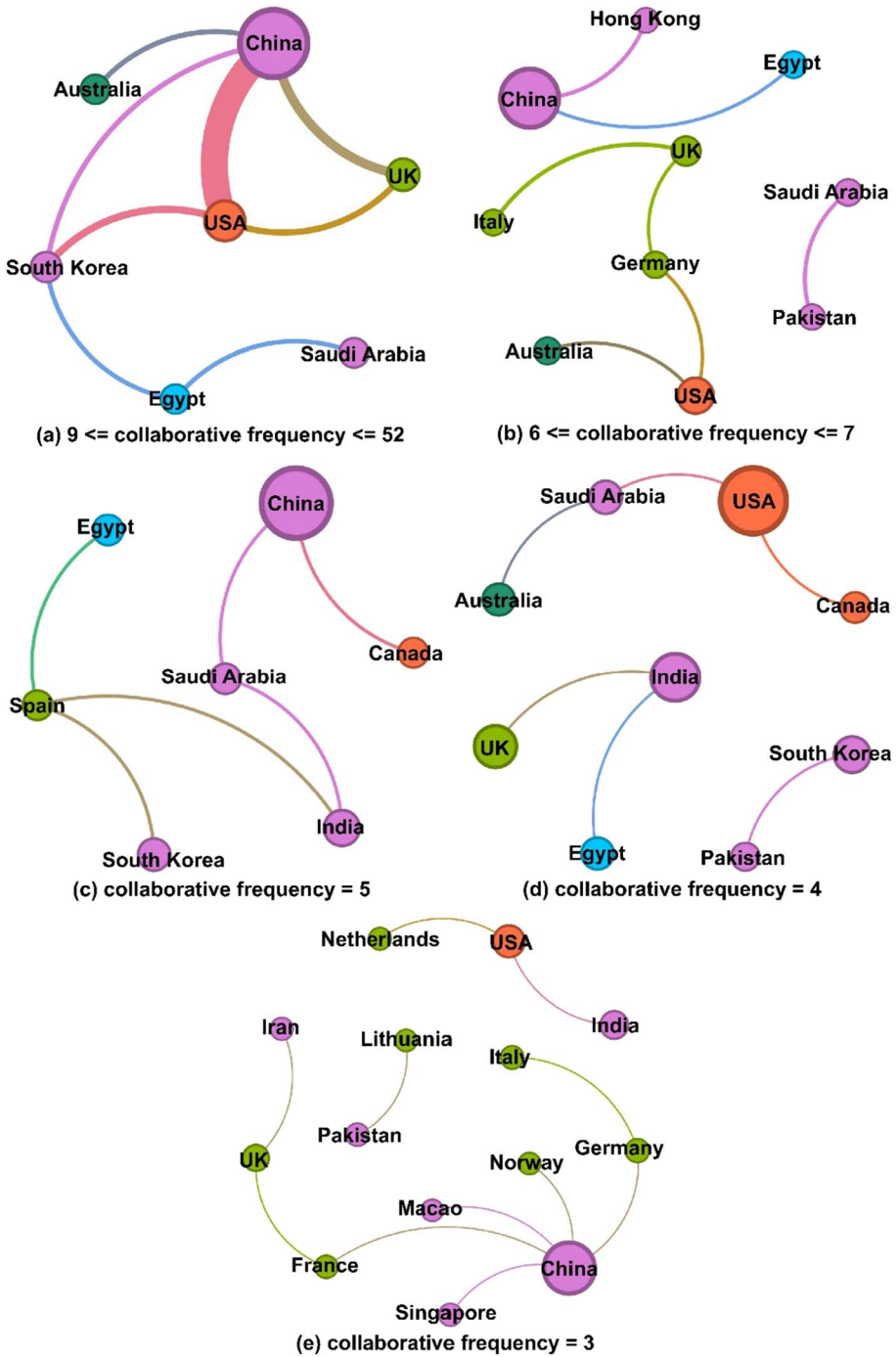


Fig. 4 Country/region collaborations ( $52 < \text{collaborative frequency} < = 3$ )

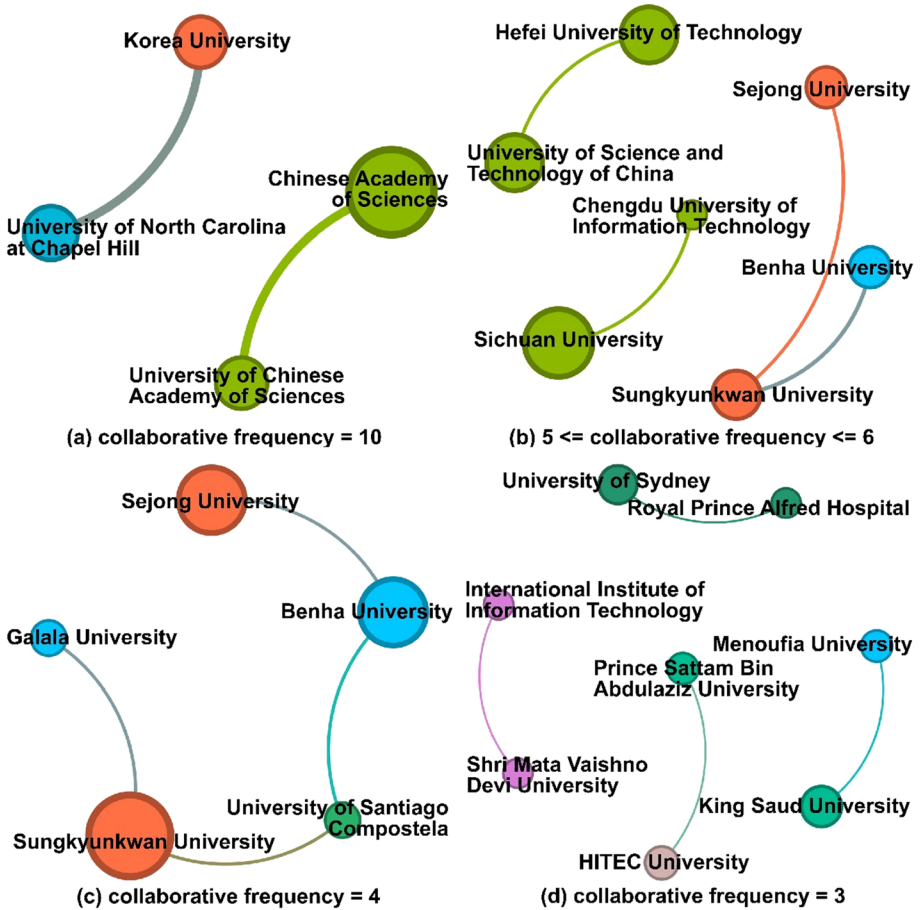


Fig. 5 Institution collaborations ( $10 \leq \text{collaborative frequency} < 3$ )

University, and Sungkyunkwan University. Figure 5c depicts a collaborative network formed by 7 institutions with a collaborative frequency of 4. These 7 institutions include Sejong University, Sungkyunkwan University, Benha University, Galala University, and University of Santiago Compostela. Figure 5d and Fig. 6 illustrate partnerships among 24 institutions, marked by a collaborative frequency of 3. Of the 24 institutions, 9 are from China (green nodes), 4 are from India (pink nodes), and 3 are from Egypt (blue nodes). The collaboration among institutions from the same countries/regions is closed as witnessed by several collaborative clusters, for example, (1) International Institute of Information Technology and Shri Mata Vaishno Devi University, (2) First Hospital of Jilin University and Jilin University, (3) Chinese Academy of Sciences, Fudan University, and University of Science and Technology of China, and (4) Benha University and Galala University.

In Fig. 7a, partnerships among 4 authors are represented, with collaborative frequencies spanning from 5 to 7. Of the 4 authors, 2 are from China (purple nodes). These 4 authors form two clusters of collaboration: (1) Yu Liu and Xun Chen, and (2) Tamer Abuhmed and Shaker El-Sappagh. Figure 7b illustrates partnerships among 10 authors, characterized by a collaborative frequency of 4. Of the 10 authors, 7 are from China (purple nodes).

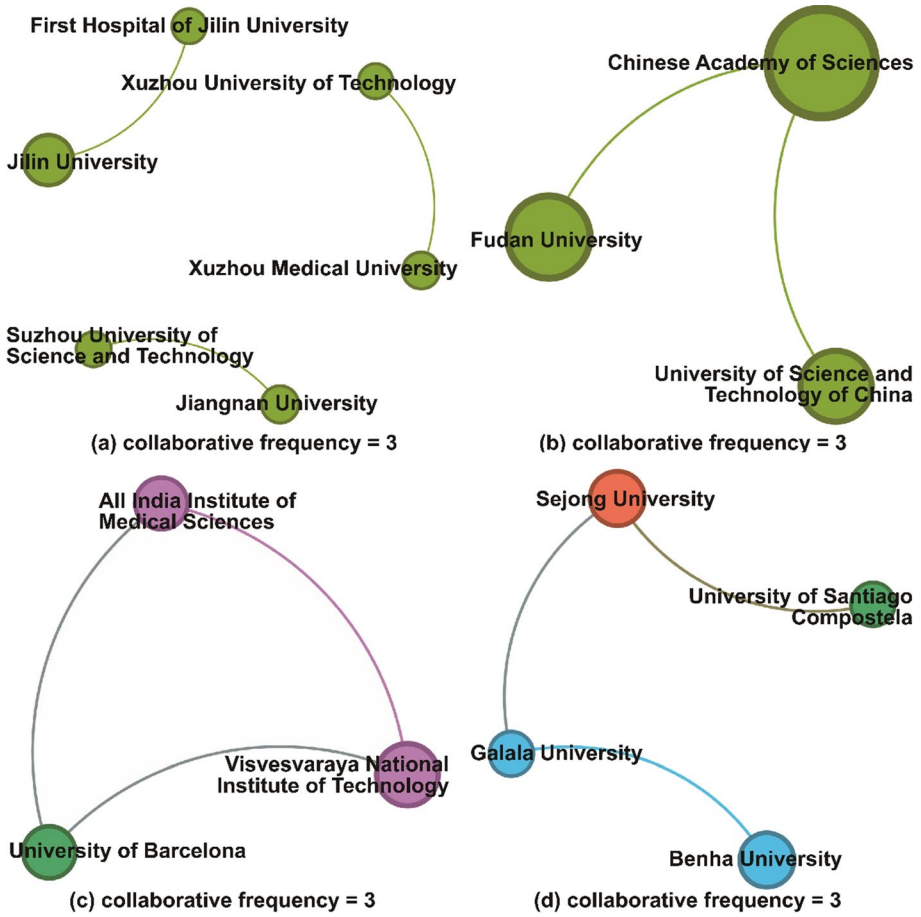


Fig. 6 Institution collaborations (collaborative frequency = 3)

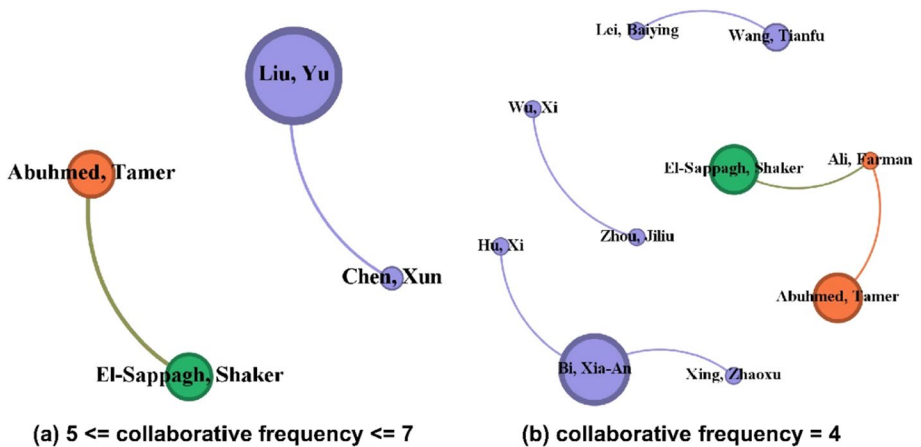


Fig. 7 Author collaborations ( $4 \leq$  collaborative frequency  $\leq 7$ )

The collaboration among them is close as witnessed by three collaborative clusters: (1) Xi Hu, Xia-An Bi, and Zhaoxu Xing, (2) Xi Wu and Jiliu Zhou, and (3) Baiying Lei and Tianfu Wang. Figure 8 showcases partnerships among 57 authors, marked by a collaborative frequency of 3. Of the 57 authors, 33 are from China (pink nodes) and 7 are from India (green nodes). The collaboration among institutions from the same countries/regions is closed as witnessed by several collaborative clusters, for example, (1) Sneha Singh, Radhey Shyam Anand, Deep Gupta, Ashwini M. Bakde, and Manisha Das, (2) Tao Zhou, Kim-Han Thung, Xiaofeng Zhu, and Dinggang Shen.

### 3.7 Top frequently used terms and phrases

The top 50 terms that appear most frequently are visually presented in Table 11. Leading the list is the term “disease”, securing the first position by making an appearance in 210 articles, representing a share of 30.75%. Other prevalent terms encompass “brain” (122 articles, 17.86%), “detection” (107 articles, 15.67%), “system” (107 articles, 15.67%), “imaging” (106 articles, 15.52%), “prediction” (97 articles, 14.2%), “multi-modality”

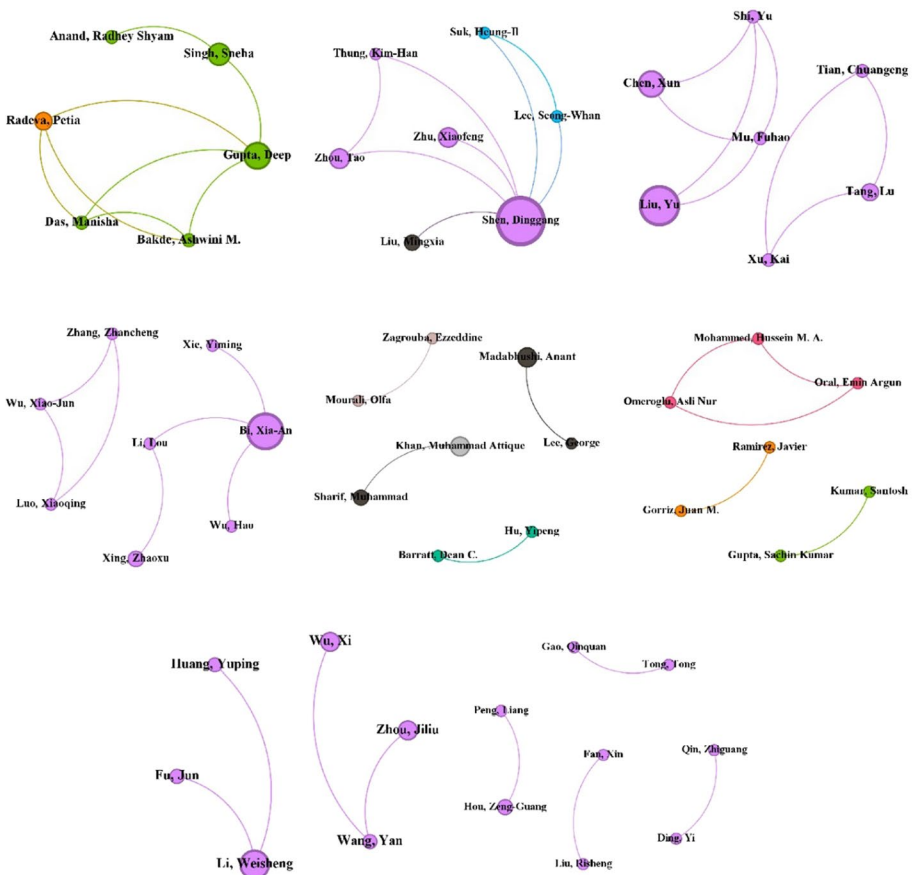


Fig. 8 Author collaborations (collaborative frequency = 3)

**Table 11** Top frequently used terms

Terms	MK test					Terms	MK test										
	A1	A2	A3	A	P		P	S	Trend	A1	A2	A3	A	P	P	S	Trend
Disease	3	11	196	210	30.75	0.0216	67	↑↑	Representation	1	2	54	57	8.35	0.0065	72	↑↑↑
Brain	2	10	110	122	17.86	0.0021	87	↑↑↑	Health	1	4	51	56	8.2	0.0423	54	↑↑
Detection	1	2	104	107	15.67	0.0078	67	↑↑↑	Classifier	2	3	47	52	7.61	0.0103	68	↑↑
System	3	11	93	107	15.67	0.0514	58	↑	Transform	4	7	41	52	7.61	0.3388	29	↑
Imaging	2	9	95	106	15.52	0.0214	67	↑↑	Diagnostic	1	3	47	51	7.47	0.0046	78	↑↑↑
Prediction	0	3	94	97	14.2	0.001	83	↑↑↑↑	Structure	0	4	47	51	7.47	0.0029	82	↑↑↑
Multi-modality	6	4	83	93	13.62	0.7327	11	↑	Activity	0	4	45	49	7.17	0.0029	75	↑↑↑
mri	2	8	82	92	13.47	0.0604	52	↑	Detect	0	1	48	49	7.17	0.0015	75	↑↑↑
Convolutional	0	1	88	89	13.03	0.0008	79	↑↑↑↑	Identify	0	5	43	48	7.03	0.0052	74	↑↑↑
Alzheimer	0	9	77	86	12.59	0.0012	89	↑↑↑	Potential	0	5	43	48	7.03	0.0036	80	↑↑↑
Segmentation	2	4	72	78	11.42	0.0125	63	↑↑	Sensor	2	4	42	48	7.03	0.0584	54	↑
Recognition	1	5	71	77	11.27	0.0103	68	↑↑	Combination	1	4	41	46	6.73	0.0103	68	↑↑
Early	0	3	71	74	10.83	0.0025	80	↑↑↑	Technology	0	3	42	45	6.59	0.0019	82	↑↑↑
Tumor	1	2	64	67	9.81	0.0067	64	↑↑↑	Coefficient	2	3	39	44	6.44	0.6059	15	↑
Visual	0	7	59	66	9.66	0.0037	73	↑↑↑	Function	0	4	40	44	6.44	0.0052	74	↑↑↑
Attention	1	1	63	65	9.52	0.0711	43	↑	Classify	0	2	41	43	6.3	0.0014	80	↑↑↑
Cancer	2	1	61	64	9.37	0.0037	73	↑↑↑	Layer	0	1	42	43	6.3	0.002	73	↑↑↑
Magnetic	0	3	61	64	9.37	0.0005	87	↑↑↑↑	Neuroimaging	0	8	35	43	6.3	0.0029	82	↑↑↑
Automatic	1	7	55	63	9.22	0.0111	70	↑↑	Rate	2	1	40	43	6.3	0.0238	60	↑↑↑
Resonance	0	3	60	63	9.22	0.0008	84	↑↑↑↑	Sample	1	1	41	43	6.3	0.0054	70	↑↑↑
Predict	1	2	58	61	8.93	0.0078	67	↑↑↑	Selection	1	3	38	42	6.15	0.0507	52	↑
Signal	0	3	56	59	8.64	0.0005	87	↑↑↑↑	Subject	1	4	37	42	6.15	0.0362	53	↑↑
Architecture	0	4	54	58	8.49	0.0025	80	↑↑↑	Correlation	1	2	38	41	6	0.0032	78	↑↑↑
Human	2	3	52	57	8.35	0.0238	60	↑↑	Functional	0	4	37	41	6	0.0015	84	↑↑↑
Region	0	3	54	57	8.35	0.0022	77	↑↑↑	Spatial	1	1	39	41	6	0.0033	74	↑↑↑

Similar to Table 6 except for A1, A2, A3, A, and P, representing the article counts for the periods of 2002–2012, 2013–2017, 2018–2023, 2002–2023, and proportion of article count (%), respectively

(93 articles, 13.62%), “mri” (92 articles, 13.47%), “convolutional” (89 articles, 13.03%), “alzheimer” (86 articles, 12.59%), and “segmentation” (78 articles, 11.42%). Utilizing the yearly term frequency data spanning from 2002 to 2022, we also incorporated the outcomes of the non-parametric MK trend test. Across these periods, most of the terms listed, particularly “prediction”, “convolutional”, “magnetic”, “resonance”, “predict”, and “signal” exhibited notably substantial increases in frequency.

Table 12 showcases the leading 42 commonly employed phrases, with “neural network” emerging as the foremost choice in 163 articles, constituting 23.87%. Other prominently used phrases included “deep learning” (154 articles, 22.55%), “experimental result” (131 articles, 19.18%), “medical image fusion” (108 articles, 15.81%), “medical image” (86 articles, 12.59%), “fused image” (86 articles, 12.59%), “multimodal fusion” (54 articles, 7.91%), “multimodal data” (47 articles, 6.88%), “source image” (44 articles, 6.44%), “magnetic resonance imaging” (40 articles, 5.86%), and “clinical diagnosis” (39 articles, 5.71%). Employing the annual phrase frequency data spanning from 2002 to 2022, we also incorporated the findings of the non-parametric MK trend test. Throughout these temporal spans, most of the phrases listed, particularly “deep learning” and “magnetic resonance imaging”, exhibited substantial changes in frequency.

Table 13 presents the emerging terms and phrases during 2020–2023. Examples of important terms include “interaction”, “encoder”, “covid-19”, “transformer”, “decoder”, “texture”, “heterogeneity”, “u-net”, and “progression”. Examples of important phrases include “multimodal fusion model”, “ct image”, “imaging modality”, “tumor segmentation”, “brain disease diagnosis”, “contextual information”, “coronavirus disease”, “electronic health record”, and “human–computer interaction”.

### 3.8 Topic identification and trend analysis

Within this investigation, the scholars engaged in topic modeling to determine a suitable quantity of topics. To achieve this, they undertook the process using a range of candidates for topic numbers, spanning from 5 to 30. According to the evaluation of semantic coherence and exclusivity performance (Fig. 9), three models were selected for manual comparison. Following a meticulous evaluation by domain experts, who drew upon their profound expertise and substantive understanding of the subject matter, the researchers ultimately opted for a 14-topic model. This selection was made after considering interpretability, relative effectiveness, external validity, and semantic coherence. This choice ensured that the model not only generated distinct topics but also maintained their interpretability. Labels were attributed to each topic by referencing typical terms and papers, and the frequent and exclusive terms (FREX) metric was utilized to pinpoint highly represented terms in each topic (Airoldi and Bischof 2016).

The outcomes, which encompass article proportions and recommended topic labels, are displayed in Fig. 10. Notably, the top three topics with the highest frequencies encompassed *adaptive transformations for enhanced visual data analysis* (12.10%), *neurodegenerative disease prediction with multi-task diagnostics* (10.70%), and *cross-modality MRI for brain tumor analysis* (9.80%). The trend assessment outcomes for the 14 topics are likewise incorporated within the Table. Eleven topics, including *neurodegenerative disease prediction with multi-task diagnostics*, *cross-modality MRI for brain tumor analysis*, *cancer prognosis through multi-dimensional data analysis*, *IoT-enabled sensory monitoring for health management*, *neuroimaging for cognitive impairment detection*, *AI for emotion recognition and post-stroke assessment*, *AI for neuroimaging-based brain disorder*

*prediction, multimedia content analysis for mental health support, AI-assisted diagnostics and personalization in healthcare, advanced signal processing for sleep and gait disorders, and neurological health monitoring via mobile technologies*, manifest a noteworthy and statistically robust escalation in trend, reaching significance at the two-sided  $p=0.05$  threshold. In contrast, the remaining 3 topics do not showcase a statistically significant trend. These trends within the mentioned topics are visually depicted in Fig. 11, which illustrates the dynamic prevalence of each of the 14 topics across time within the entire dataset.

## 4 Discussion

### 4.1 Answers to RQs

Within this research endeavor, a topic-focused bibliometric investigation centered on AI-powered smart healthcare through multimodal data analysis was devised to address the earlier-raised research inquiries. Regarding RQ1, the comprehensive augmentation in the corpus of scholarly articles, graphically depicted in Fig. 3, indicates a mounting fascination with research within this multidisciplinary sphere over the last two decades. A trend analysis further accentuates this trajectory, prognosticating a continued surge. This trajectory underscores the promise of AI-powered smart healthcare anchored in multimodal data analysis, substantiating a flourishing research community and a burgeoning output of academic contributions. Illustrated in Table 6, the scrutiny of journals identified a roster of periodicals inclined to publish works underpinning AI-powered smart healthcare via multimodal data analysis. The distribution of these publications was notably diverse, reflecting a broad spectrum of research perspectives. The findings also reveal a discernible popularity surge, particularly since 2013, for articles elucidating the nexus between AI technology and its augmentation of multimodal health and medical data analysis for the advancement of smart healthcare. These contributions notably found resonance in interdisciplinary journals bridging healthcare, medical research, and information technology, underlining their invaluable contribution to these domains. Surveying the results encompassing countries and regions in Table 8, the notable escalation in published studies can be attributed to two core factors. Firstly, it stems from the escalating researchers' interest hailing from non-English speaking countries/regions, typified by China and South Korea. Secondly, it is complemented by the substantive contributions of authors originating from the USA and India. Notably, China's commanding role is evident, contributing approximately 54% of the scrutinized articles. This dominance extends to the identification of 14 out of the 17 most prolific institutions (Table 9) and 9 out of the 17 most prolific authors (Table 10). Such prominence firmly establishes China as a pivotal player in this domain, with its three top prolific authors solidifying its stature.

Regarding RQ2, Fig. 4 illustrates that countries/regions with robust international collaborations tend to exhibit substantial progress and growth in the field, exemplified by the achievements of the USA, the UK, China, and India. This underscores the significance of global cooperation in propelling the advancement of this emerging domain, facilitating the harnessing of opportunities, and the effective addressing of challenges that may arise. Figures 5, 6, 7 and 8 also imply that institutions and authors hailing from the same geographical regions are inclined to engage in this multidisciplinary research. This pattern could stem from the inherent ease of communication and resource sharing within proximity. However,

**Table 12** Top frequently used phrases

Phrases	AI				A3				A				MK test		Trend		
	A1	A2	A3	A	P	A1	A2	A3	A	P	A1	A2	A3	A		P	S
Neural network	3	8	152	163	23.87	0.0292	69	↑↑	Classification accuracy	0	4	16	20	2.93	0.0061	73	↑↑↑
Deep learning	0	3	151	154	22.55	0.0002	108	↑↑↑↑	Visual quality	0	0	20	20	2.93	0.0013	80	↑↑↑
Experimental result	5	5	121	131	19.18	0.2621	37	↑	Clinical practice	2	1	16	19	2.78	0.0113	71	↑↑
Medical image fusion	5	10	93	108	15.81	0.2285	40	↑	Feature selection	0	2	17	19	2.78	0.0038	81	↑↑↑
Medical image	2	7	77	86	12.59	0.0625	59	↑	Multimodal image	1	1	17	19	2.78	0.0091	73	↑↑↑
Fused image	2	3	51	56	8.2	0.0064	84	↑↑↑	Brain tumor segmentation	0	1	17	18	2.64	0.0088	60	↑↑↑
Multimodal fusion	0	7	47	54	7.91	0.0054	81	↑↑↑	Clinical data	0	4	14	18	2.64	0.0296	50	↑↑
Multimodal data	2	0	45	47	6.88	0.0048	75	↑↑↑	Computer-aided diagnosis	1	2	15	18	2.64	0.0647	52	↑
Source image	1	3	40	44	6.44	0.0261	65	↑↑	Cognitive impairment	0	1	16	17	2.49	0.0043	71	↑↑↑
Magnetic resonance imaging	0	2	38	40	5.86	0.0004	99	↑↑↑↑	Computed tomography	0	4	13	17	2.49	0.0169	67	↑↑
Clinical diagnosis	0	4	35	39	5.71	0.0049	82	↑↑↑	Medical diagnosis	0	1	16	17	2.49	0.0068	62	↑↑↑
Learning approach	1	4	31	36	5.27	0.0221	69	↑↑	Treatment planning	1	5	11	17	2.49	0.0796	47	↑
Complementary information	1	2	31	34	4.98	0.0091	73	↑↑↑	Accurate diagnosis	0	2	14	16	2.34	0.0061	73	↑↑↑
Disease diagnosis	1	4	29	34	4.98	0.0074	75	↑↑↑	Data fusion	0	3	13	16	2.34	0.0076	71	↑↑↑
Early diagnosis	0	0	31	31	4.54	0.0231	39	↑↑	Multimodal image fusion	0	2	14	16	2.34	0.0076	71	↑↑↑
Fusion model	0	2	26	28	4.1	0.0014	85	↑↑↑	Attention mechanism	0	0	15	15	2.2	0.0051	57	↑↑↑
Image fusion	1	2	25	28	4.1	0.034	62	↑↑	Clinical application	1	2	12	15	2.2	0.0199	68	↑↑
Machine learning	0	1	25	26	3.81	0.003	74	↑↑↑	Magnetic resonance image	0	1	13	14	2.05	0.003	74	↑↑↑
Artificial intelligence	0	0	24	24	3.51	0.0022	70	↑↑↑	Medical imaging	0	2	12	14	2.05	0.0088	60	↑↑↑
Feature fusion	0	0	24	24	3.51	0.0013	80	↑↑↑	Multimodal data fusion	0	1	13	14	2.05	0.003	74	↑↑↑
Brain tumor	0	1	19	20	2.93	0.0017	78	↑↑↑	Multimodal feature fusion	0	0	14	14	2.05	0.0051	57	↑↑↑

Similar to Table 11



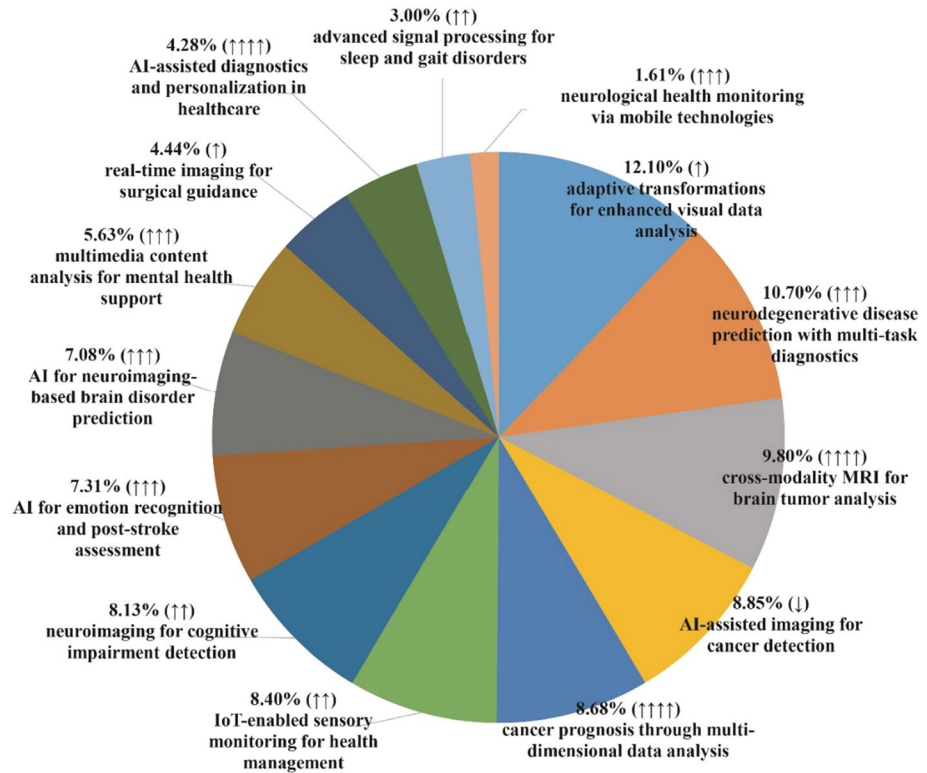
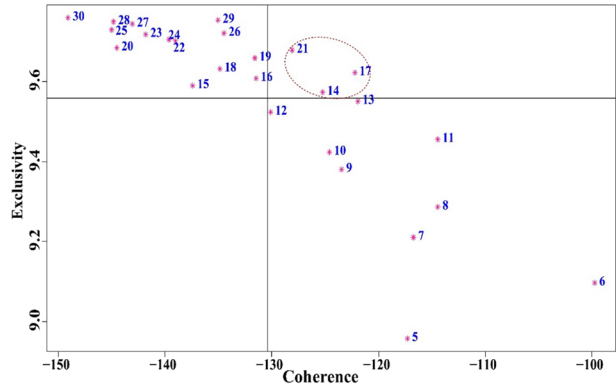
**Table 13** emerging terms and phrases since 2020

Terms	Number of articles					Phrases	Number of articles				
	2020–2021		2022		2023		2020–2021		2022		2023
	2020	2021	2022	2023	2020–2023		2020	2021	2022	2023	2020–2023
Interaction	5	7	11	7	30	Multimodal fusion model	0	2	8	2	12
Encoder	3	2	12	6	23	ct image	3	2	4	2	11
Covid-19	1	6	12	2	21	Imaging modality	2	2	2	3	9
Transformer	0	3	9	8	20	Tumor segmentation	0	2	3	4	9
Decoder	1	2	9	6	18	Multimodal mr image	0	2	10	4	16
Texture	1	8	6	2	17	Brain disease diagnosis	1	1	3	1	6
Heterogeneity	3	4	8	1	16	Contextual information	1	2	2	1	6
u-net	2	3	7	4	16	Coronavirus disease	0	1	5	0	6
Progression	1	4	7	2	14	Covid-19 diagnosis	0	1	4	1	6
Node	2	3	4	3	12	Early stage	1	1	3	1	6
Encode	1	7	1	1	10	Electronic health record	2	1	2	1	6
Resnet	0	1	8	1	10	Heterogeneous data	1	2	3	0	6
Ablation	0	1	4	4	9	Human-computer interaction	0	2	1	3	6
Augmentation	2	3	2	2	9	Multi-scale feature	1	0	3	2	6
Chest	0	0	5	4	9	Attention model	1	2	1	1	5
Contextual	1	3	4	1	9	Cervical cancer	0	1	3	1	5
Coronavirus	0	1	7	1	9	Deep transfer learning	0	0	3	2	5
Dependency	0	2	4	3	9	Depression recognition	0	2	2	1	5
Fine-tuned	2	3	2	2	9	Differential evolution	0	3	2	0	5
Modality-specific	0	2	4	3	9	Fusion feature	0	1	2	2	5
Alteration	1	2	2	3	8	x-ray image	0	0	2	3	5
Attention-based	0	0	6	2	8	Chest x-ray	0	0	3	1	4
Dense	0	2	5	1	8	Covid-19 pandemic	0	0	2	2	4
Grading	2	2	4	0	8	Hybrid model	0	1	2	1	4
Pain	1	1	5	1	8	Image registration	0	0	3	1	4

**Table 13** (continued)

Terms	Number of articles					Phrases	Number of articles				
	2020–2021		2022–2023		2020–2023		2020–2021		2022–2023		2020–2023
	2020	2021	2022	2023			2020	2021	2022	2023	
Relation	0	2	5	1	8	Molecular data	0	0	3	1	4
x-ray	0	0	4	4	8	Adversarial network	0	1	1	1	3
Cross-modal	0	1	5	1	7	Breast cancer diagnosis	0	0	3	0	3
Cross-modality	0	2	3	2	7	Chinese medicine	0	0	3	0	3
Embedding	0	1	6	0	7	Contrastive learning	0	0	0	3	3
Hyper-parameter	0	2	5	0	7	Depression detection	0	0	2	1	3
Spatio-temporal	0	5	2	0	7	Generative adversarial network	0	1	2	0	3
Stress	1	2	3	1	7	Information redundancy	0	0	2	1	3
Deviation	0	3	3	0	6	Noninvasive technique	0	0	2	1	3
Efficientnet	0	1	2	3	6	Non-small cell lung cancer	0	0	1	2	3
Generative	0	2	3	1	6	Pain assessment	0	1	2	0	3
Infection	0	2	4	0	6	Pancreatic cancer	0	1	2	0	3
Lymphoma	0	2	2	2	6	Pathological examination	0	1	1	1	3
Bilateral	0	2	2	1	5	Pixel-level fusion	0	1	1	1	3
Bilinear	0	3	0	2	5	Precise diagnosis	0	1	2	0	3
Blood	0	1	4	0	5	Residual attention	0	0	2	1	3
Cervical	0	1	3	1	5	Residual learning	0	0	3	0	3
Chain	0	2	2	1	5	Self-supervised learning	0	0	2	1	3
Long-range	0	2	2	1	5	Skin cancer	0	0	2	1	3
Manifestation	0	1	3	1	5	Visual question	0	0	2	1	3

**Fig. 9** Semantic coherence and exclusivity of models with topics ranging from 5 to 30



**Fig. 10** Topic proportions, label, and developmental tendencies (↑(↓): increasing (decreasing) trend but not significant ( $p > 0.05$ ); ↑↑(↓↓), ↑↑↑(↓↓↓), ↑↑↑↑(↓↓↓↓): significantly increasing (decreasing) trend ( $p < 0.05$ ,  $p < 0.01$ , and  $p < 0.001$ , respectively))

mitigating negative local effects might involve collaborating with more distant countries/ regions to maximize impact gains. Consequently, we advocate for increased cross-regional partnerships to embrace the evolving challenges of applying AI and multimodal data fusion for smart healthcare.

In response to RQ3 concerning research themes, the results from the analysis of frequently used terms and topics exhibit substantial congruence, reinforcing each other’s findings. This study delineates 11 distinct themes, emerging through the combined interpretation of word/phrase frequency analyses (Tables 11, 12 and 13), topic modeling (Figs. 10 and 11), and an examination of papers focusing on the identified topics and terms. These themes encapsulate the contemporary and evolving research dimensions within the realm of AI-powered smart healthcare predicated on multimodal data analysis. Further elaboration on the formation of each theme is provided below.

### 4.2 Neurodegenerative disease prediction using AI-powered multi-task diagnostics

There has been a growing enthusiasm for neurodegenerative disease prediction using AI-powered multi-task diagnostics and multimodal data fusion, evidenced by the increasing tendencies of *neurodegenerative disease prediction with multi-task diagnostics* (Fig. 11) and frequent and emerging terms/phrases such as “alzheimer”, “disease”, “prediction”, “combination”, and “disease diagnosis” (Table 13).

AI-powered multi-task diagnostics and multimodal data fusion hold great promise in neurodegenerative disease prediction by integrating diverse data sources, enhancing accuracy, enabling personalized risk assessment, and supporting proactive interventions for at-risk individuals. Multimodal data fusion integrates medical imaging (e.g., MRI, PET scans), genetic, clinical, wearable sensor, and cognitive data. AI-powered multi-task diagnostic models leverage this information to assess an individual’s health status and neurodegenerative disease risk, identifying complex patterns associated with early-stage neurodegenerative diseases. Considering multiple biomarkers and risk factors, these models achieve higher accuracy in detecting at-risk individuals and neurodegenerative diseases at earlier stages. AI-powered multi-task diagnostics enable personalized risk

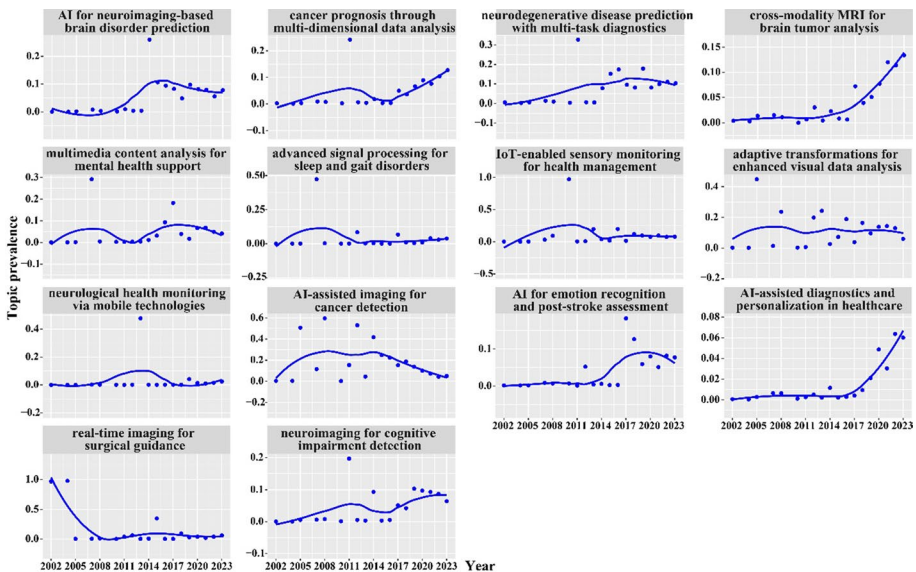


Fig. 11 Trends of the 14 topics

assessment, guiding targeted prevention and personalized care plans. Early prediction empowers proactive interventions, implementing lifestyle modifications, cognitive training, and disease-specific therapies for high-risk individuals. Post-diagnosis, these diagnostics continuously monitor disease progression through longitudinal data integration, facilitating treatment assessment and care plan adjustments. Additionally, AI-powered multi-task diagnostics aid neurodegenerative drug development by identifying patient subgroups for targeted therapies in clinical trials.

Researchers have been interested in jointly applying AI-powered multi-task diagnostics and multimodal data fusion technologies to facilitate neurodegenerative disease prediction. For example, (Liu et al. 2019a) focused on multi-task feature learning by combining fused group lasso and  $\ell_{2,1}$ -norm with mixed norms to capture more adaptable structures. An alternating direction approach of multipliers was utilized to efficiently resolve non-smooth formulation. (Shao et al. 2020) suggested a hypergraph-driven multi-task feature selection technique to classify AD/MCI. They initiated feature selection on individual modalities as distinct tasks and incorporated group-sparsity regularization to simultaneously choose shared features across diverse modalities. Additionally, they integrated a hypergraph-grounded regularization component into the conventional multi-task feature selection process to account for the complex interrelationships among subjects. Ultimately, a multi-kernel support vector machine (SVM) classifier was employed for fusing the selected features from various modalities.

### 4.3 AI for neuroimaging-based brain disorder prediction and detection

Evidenced by the increasing tendencies of *AI for neuroimaging-based brain disorder prediction* and *neuroimaging for cognitive impairment detection* (Fig. 11) and frequent and emerging terms/phrases such as “imaging”, “neuroimaging”, “brain”, “detection”, “prediction”, “cognitive impairment”, and “brain disease diagnosis” (Tables 11, 12 and 13), we formed the second theme as “AI for neuroimaging-based brain disorder prediction and detection”.

Brain disorders and cognitive impairment, a common feature in many brain disorders describing a decrease in cognitive capabilities (such as memory, attention, language skills, executive functions, and problem-solving aptitudes), have been widely discussed by researchers the combination of AI and multimodal data fusion with neuroimaging techniques holds immense promise for tackling brain disorders and cognitive impairment in smart healthcare. For example, AI-powered neuroimaging analysis can identify early biomarkers associated with brain disorders and cognitive impairment. Early detection allows prompt interventions, potentially allowing for disease treatment and personalized care plans. Also, the combination of AI and multimodal data fusion facilitates precision medicine approaches. By tailoring treatments suiting individuals. neuroimaging and cognitive profiles, healthcare providers can optimize treatment efficacy and minimize adverse effects. Furthermore, AI algorithms can analyze large-scale neuroimaging datasets to discover new biomarkers for brain disorders and cognitive impairment. These biomarkers can aid in diagnosis, prognosis, and therapeutic target identification. In addition, the integration of AI with neuroimaging technologies can support telemedicine and remote care, providing access to specialized brain health assessments for individuals in remote or underserved areas.

Researchers have been interested in jointly applying AI, multimodal data fusion, and neuroimaging techniques to tackle brain disorders and cognitive impairment. For instance,

(Zhang et al. 2021) introduced a graph-based DNN that concurrently captured both brain structure/function in MCI. The initial graph topology was derived from structural network data (obtained through diffusion MRI) and was progressively refined by integrating functional information (gathered from functional MRI). This process aimed to enhance the discriminative power between MCI patients and elderly normal controls. (Wang et al. 2023) developed a genetic evolution random neural network clustering approach based on mutual information correlation analysis for merging resting-state functional MRI (fMRI) data with single nucleotide polymorphism data, forming fused features. To prevent the model from getting trapped in local optimal solutions, the traditional genetic evolution algorithm was enhanced with strategies such as elite retention and large variation genetic algorithms. Through multiple independent experiments, the proposed model demonstrated greater effectiveness in identifying AD patients and extracting pertinent pathogenic factors, thus holding promise as a valuable tool in AD research.

#### 4.4 Cross-modality MRI for brain tumor diagnostics

The third theme “cross-modality MRI for brain tumor diagnostics” is formed considering the increased interest in the topic of *cross-modality MRI for brain tumor analysis* (Fig. 11) and the frequent and emerging terms/phrases “tumor”, “brain tumor”, “brain tumor segmentation”, “cross-modality”, and “tumor segmentation” (Tables 11, 12 and 13).

The combination of AI and cross-modality MRI data fusion and analysis is effective for brain tumor segmentation and diagnostics. Specifically, MRI data can be acquired using different imaging sequences, each providing complementary information about the brain’s anatomy and pathology. For instance, T1-weighted images highlight anatomical structures, T2-weighted images emphasize edema and necrotic regions, and contrast-enhanced images highlight areas with increased vascularity, such as tumor borders. By fusing information from these multiple modalities, AI models leverage each sequence’s advantages, leading to an inclusive representation of the tumor and its surrounding tissues. Furthermore, AI-based segmentation techniques, such as deep learning algorithms, excel at learning complex patterns and features within medical images. With access to cross-modality MRI data, these algorithms can extract intricate details and subtle variations indicative of tumor boundaries. In addition, AI-powered cross-modality data fusion enables rapid and automated segmentation, reducing the burden on healthcare professionals and allowing them to focus more on treatment planning and patient care. Once trained on a diverse dataset with cross-modality MRI data, AI models not only can demonstrate excellent generalization across different patient cohorts and imaging protocols even when dealing with challenging cases or novel MRI acquisition techniques, but also be used for predictive analytics for tailoring personalized treatment strategies and monitoring disease progression over time.

Researchers have been interested in leveraging AI’s ability to integrate and process information from multiple MRI sequences to improve segmentation accuracy, save time, enhance image quality, and enable predictive analytics, ultimately leading to better patient care and outcomes in brain tumor diagnostics. (Amin et al. 2020a) proposed a fusion approach for amalgamating the structural and textural information originating from four distinct MRI sequences to enhance brain tumor detection. The fusion process employed a discrete wavelet transform in conjunction with the Daubechies wavelet kernel, leading to an informative representation of the tumor region compared to individual MRI sequences. After fusion, a partial differential diffusion filter was employed to mitigate noise. A

global thresholding technique was implemented to segment the tumor region, which was inputted into a CNN for the ultimate differentiation between tumor and non-tumor regions. Amin et al. (Amin et al. 2020b) put forth an automated approach that facilitated the discrimination between cancerous and non-cancerous brain MRI scans. Various methods were employed for the segmentation of potential lesions. Following segmentation, a feature set was selected for each identified lesion, incorporating factors such as shape, texture, and intensity. Subsequently, an SVM classifier was applied, employing diverse cross-validation strategies on the feature set to gauge the efficacy of the proposed framework.

#### 4.5 Multi-dimensional data analysis for cancer diagnostics

Evidenced by the increased attention received by the topic of *cancer prognosis through multi-dimensional data analysis* (Fig. 11) and the frequently used term “cancer” (Table 11), we formed the fourth theme as “multi-dimensional data analysis for cancer diagnostics”.

As a complex and heterogeneous disease, cancer diagnosis needs data integration from multiple sources and dimensions. AI-driven multi-dimensional data analysis and fusion contribute to the effectiveness of cancer diagnostics. Specifically, combining information from various dimensions offers a more holistic overview of the disease, contributing to a better understanding of cancer’s molecular, cellular, and clinical characteristics. Using AI for processing/analyzing large-scale, multi-dimensional datasets can promote identifying meaningful but subtle patterns and associations that might otherwise be challenging for human experts to discern. This enables earlier detection of cancer and potentially leads to improved patient outcomes through timely interventions. Furthermore, cancer is highly individualized, with variations in tumor subtypes, genetic mutations, and treatment responses among patients. AI’s ability to analyze multi-dimensional data allows for the identification of specific molecular and genetic profiles associated with different cancer types. This personalized approach enables tailored treatment plans, optimizing therapeutic effectiveness while reducing potential side effects. In addition, AI models can utilize multi-dimensional data to build predictive models in the context of cancer advancement, treatment reaction, and patient results, these models have the potential to support healthcare professionals in making well-informed decisions about treatment options and design patient-specific care plans based on predicted disease trajectories.

Researchers are increasingly harnessing the power of AI to quickly process and interpret complex and diverse data sources, to assist clinicians in making timely and evidence-based decisions for cancer diagnostics and treatment planning. For example, (Shao et al. 2019) developed an ordinal multimodal feature selector with the ability to concurrently extract significant features from pathological images and multimodal genomic data to identify cancer patients. (Wang et al. 2019) unveiled a fresh approach known as bacterial colony optimization with a multi-dimensional population (BCO-MDP) for feature selection in classification tasks. The BCO-MDP method demonstrated its superiority over binary models regarding feature size and efficiency and maintaining lower computational complexity.

#### 4.6 Advanced signal processing for sleep and gait disorder diagnosis

The fifth theme “advanced signal processing for sleep and gait disorder diagnosis” is formed considering the increased interest in the topic of *advanced signal processing*

for sleep and gait disorders (Fig. 11) and the frequent and emerging terms “signal” and “diagnosis” (Table 11).

Advanced signal processing driven by AI promotes automated analysis of sleep and gait disorders by extracting meaningful patterns and features from the raw data in several ways. First, in sleep and gait analysis, the raw data collected from various sensors can be complex and contain a vast amount of information. Deep learning and machine learning, possess the capability to autonomously recognize pertinent attributes within the unprocessed data. For example, in sleep analysis, AI can extract sleep stages and identify sleep interruptions based on patterns in EEG or actigraphy data. Similarly, in gait analysis, AI can extract gait parameters like gait speed, stride length, and cadence, from gait sensor data. Second, AI algorithms excel at pattern recognition and can identify subtle anomalies or deviations that may not be easily recognizable by human observers. For instance, AI can detect abnormal sleep architecture indicative of sleep disorders or identify abnormal gait patterns associated with neurological conditions. In addition, AI-driven advanced signal processing is capable of amalgamating information sourced from various sensors and modalities, leading to a more holistic comprehension of a patient’s sleep and gait patterns. This data fusion allows AI models to consider contextual information, such as environmental factors or patient-specific characteristics, enhancing the accuracy and specificity of the analysis.

Researchers have been interested in leveraging AI’s pattern recognition capabilities and adaptability with advanced signal processing techniques for accurate and automated analysis, ultimately improving the diagnosis and management of sleep and gait disorders in smart healthcare settings. For example, (Sun et al. 2020) described a method for recognizing identities based on gait, which is applied to manage access to wearable healthcare devices targeted at the elderly. This method addresses the challenge of variations in gait within the same individual, resulting in a notable enhancement in recognition accuracy compared to existing approaches. (Chakraborty and Nandy 2020) presented a DNN framework that employs discrete wavelet decomposition to represent data for the detection of irregular gait patterns using inertial sensors. (Lin et al. 2020) devised a hybrid architecture for a body sensor network that integrated multiple sensors through multi-sensor fusion. This architecture facilitated advanced smart medical services by amalgamating various technologies such as sensors, communication, robots, and data processing. (Yang et al. 2021) suggested an IoT-based module for fusing sleep data, known as sleep data fusion networks, utilizing a Bluetooth network in a star topology. This network integrated sleep-aware applications’ data and deployed machine learning to recognize sleep events using audio signals.

#### 4.7 IoT and mobile-assisted health management and monitoring

Evidenced by the increased attention received by the topics of *IoT-enabled sensory monitoring for health management, neurological health monitoring via mobile technologies, AI-assisted diagnostics and personalization in healthcare, and AI for post-stroke assessment* (Fig. 11) as well as the frequently used terms “health”, “diagnostic”, and emerging phrase “human–computer interaction” (Table 11 and Table 13), we formed the sixth theme as “IoT and mobile-assisted health management and monitoring”.

AI plays a crucial role in supporting IoT-enabled sensory monitoring for health management for the following reasons. First, IoT devices gather an extensive volume of health-related information through diverse sensors, including wearable gadgets, smart wristwatches, and fitness monitors. AI can efficiently integrate and process



this heterogeneous data from different sources, enabling real-time processing and analysis even in large-scale deployments. Furthermore, AI algorithms can perform real-time analysis on the continuously streaming data from IoT devices. This enables immediate identification of anomalies or critical health events, providing timely alerts and interventions to healthcare providers and patients. In addition, AI can recognize individual patterns and tailor health management strategies based on each person's unique health profile, and predict potential health risks or changes in health conditions based on historical data from IoT devices, allowing for proactive health management and personalized interventions, reducing the risk of adverse health outcomes.

AI also supports neurological health monitoring via mobile technologies in the following ways. First, AI-powered mobile technologies facilitate remote neurological health monitoring by evaluating cognitive functions, motor skills, and other neurological parameters, enabling telemedicine services, reducing the burden on healthcare facilities, and providing valuable data for healthcare professionals, especially for people with limited access to neurological care. Furthermore, AI enables continuous monitoring of neurological health through mobile devices, capturing real-life data, providing an inclusive understanding of an individual's neurological status, and tracking changes in neurological health over time, thus providing insights into disease progression and treatment effectiveness. In addition, AI algorithms can automatically analyze data from mobile neurological monitoring devices, such as smartphone-based neurocognitive tests or wearable brain-computer interfaces. AI can identify abnormal patterns or changes in neurological function, assisting in the early detection of neurological disorders.

The combination of AI and wearable sensors for post-stroke assessment has also been an important and promising topic for several reasons. First, wearable sensors can capture gait patterns and mobility-related metrics, crucial for assessing a stroke survivor's walking ability and balance in real time. The captured data, processed and analyzed by AI algorithms allow for immediate assessment of a stroke survivor's health status and detect any sudden changes or potential complications. This AI-driven analysis delivers data-driven insights for informed decision-making by clinicians and caregivers, enhancing the precision and efficacy of post-stroke rehabilitation. Furthermore, wearable sensors integrated with AI-based speech recognition can assist in assessing post-stroke speech and language impairments. AI algorithms can analyze speech patterns, detecting dysarthria or aphasia, which are common after a stroke, facilitating targeted speech therapy. AI-powered cognitive assessment tools can also analyze cognitive function through wearable sensors and other data inputs. This comprehensive evaluation allows clinicians to monitor cognitive improvements or identify cognitive challenges that may require further intervention.

Researchers have been interested in leveraging AI's abilities to process large amounts of data, conduct real-time analysis, provide predictive insights, and personalize healthcare interventions to support IoT-enabled sensory monitoring for health management, neurological health monitoring via mobile technologies, and post-stroke assessment using wearable sensors. For example, (Wu et al. 2021b) described an IoT-driven real-time health monitoring system utilizing deep learning techniques. This system employed wearable medical devices for vital sign measurements and employed diverse DNN algorithms to extract meaningful insights. Ali et al. (Ali et al. 2021) suggested an innovative healthcare monitoring framework that operated within a cloud environment. This framework incorporated big data analytics through data mining methods, ontologies, and bidirectional long short-term memory. The goal was to accurately store and examine healthcare data. Razfar et al. (Razfar et al. 2023) put forth an effective diagnostic system for post-stroke

conditions, leveraging multi-level ensemble learning that combined heterogeneous or homogeneous baseline classifiers based on Xsens wearable sensors.

#### 4.8 Mental health support based on multimedia content analysis

The seventh theme “mental health support based on multimedia content analysis” is formed considering the increased interest in the topics of *multimedia content analysis for mental health support* and *AI for emotion recognition* (Fig. 11).

The use of multimedia content like audio, video, text, and images, has become prevalent in mental health applications. AI techniques offer several advantages in this domain, allowing for more effective and personalized mental health support. In particular, AI-driven NLP algorithms such as sentiment analysis can analyze and extract valuable insights from text data, including social media posts, online forums, and chat conversations. The detected sentiment, emotions, and linguistic patterns might indicate mental health concerns. This enables real-time monitoring of individuals at scale and facilitates timely interventions. Also, AI-based computer vision techniques and Speech recognition algorithms can analyze images and videos to detect facial expressions, body language, and behavioral cues to identify emotions and affective states that indicate mental health issues. In addition, AI can integrate data from various multimedia sources along with other patient data (e.g., EHRs and wearables) to offer a comprehensive evaluation of individuals’ mental health status.

Researchers are increasingly leveraging AI’s capabilities in analyzing multimedia content to provide valuable insights and personalized support for mental health. For example, (Sawhney et al. 2020) introduced STATENet, a transformer-based model with a time-aware approach designed to conduct initial screening of suicidal risk in social media content. This was achieved by enhancing linguistic models with historical context. (Ghosh et al. 2022) presented a deep multitask framework incorporating a knowledge component that integrated external knowledge-specific features into learning, utilizing SenticNet’s IsaCore and AffectiveSpace vector spaces. The framework concurrently handled tasks such as emotion recognition, depression detection, and sentiment classification.

#### 4.9 Latest trends in AI-powered smart healthcare based on multimodal data analysis

Evidenced by the emerging phrases such as “generative adversarial network”, “contrastive learning”, “spatio-temporal”, “contextual information”, “attention model”, “noninvasive technique” (Table 13), as well as frequent phrases that show an increase in usage such as “neural network” and “deep learning” (Table 12), we highlight several latest topics in the field.

To begin, a surge of interest is evident in the application of generative adversarial networks (GANs) for both multimodal medical image fusion and synthesis. For one thing, while convolutional methods for image fusion excel at extracting local features, they often struggle to capture broader global information. This shortfall can lead to output images with reduced clarity and increased noise. To address this challenge, several researchers have incorporated GANs into image fusion. (Tang et al. 2022) developed a multiscale adaptive transformer fusion approach to enhance the preservation of global contextual information when fusing MRI and single photon emission computed tomography images. (Rao et al. 2023) developed TGFuse employing a lightweight transformer component and adversarial learning to enhance the integration of shallow features, refining fusion relations

both spatially and across channels. (Liu et al. 2023) put forth a GAN tailored for the multimodal MRI fusion of brain tumors. The approach employs a generator characterized by a nested U-shaped structure along with residual U-blocks, thereby enhancing multi-scale feature extraction.

For another thing, the exploration of GAN-based approaches extends to the realm of image synthesis. Nema, for instance, introduced the concept of residual cyclic unpaired encoder-decoder networks aimed at segmenting brain tumors within MRI data (Nema et al. 2020). (Huang et al. 2022b) put forward a 3D common-feature learning-powered context-aware GAN that adopted an encoder-decoder structure to facilitate input modality mapping into shared feature spaces. Qin et al. (Qin et al. 2022) integrated style transfer into the architecture of conditional GANs resulting in the hierarchical feature mapping and fusion model. This approach tackled the challenge of cross-modality synthesis for MR images. (Mi et al. 2022) described a medical image-fusion method that leveraged a straightforward network obtained through knowledge distillation to extract features from computer tomography and magnetic resonance modes, thereby reducing the data volume required for training complex networks. (Gao et al. 2021) developed a sophisticated DNN method by combining task-induced pyramids and attention GANs for classifying multimodal brain images.

Second, contrastive learning is a powerful self-supervised learning technique that has shown great promise in medical image fusion. Specifically, Contrastive learning helps the model learn meaningful representations of medical images. By contrasting positive with negative pairs, the model can capture important features and patterns specific to the medical imaging domain. This holds significant importance in the context of image fusion, as it enables the model to identify relevant information in multiple images for more effective fusion. Also, Medical image fusion often requires extracting and combining high-level image features effectively. Contrastive learning enables the model to recognize and extract salient features from each image independently. This leads to a more robust and informative feature space that can capture essential information from medical images. Furthermore, Contrastive learning is a self-supervised technique that leverages unlabeled data. In medical imaging, gaining labeled data for supervised learning is difficult and time-consuming because of the need for expert annotations. Contrastive learning can utilize abundant unlabeled medical images to learn powerful representations, making it a more data-efficient approach. As a result, by learning rich representations and informative features, contrastive learning can improve fused medical image quality. The fusion process benefits from a more comprehensive understanding of the underlying data, leading to better visual clarity, enhanced structural details, and improved information preservation in the fused image. For example, (Zhang et al. 2023) drew inspiration from contrastive learning and constructed pairs of positive and negative outcomes, introducing a unique contrastive loss within an auto-encoder framework. This contrastive auto-encoding, coupled with information exchange through convolutions, was employed for multimodal medical fusion. Taking cues from autoencoder networks and contrastive learning, a multi-branch encoder was established with contrastive constraints to grasp the shared and distinct attributes of paired images (Luo et al. 2021).

Thirdly, neuroimaging data, encompassing images generated through functional and metabolic assessments (such as fMRI, functional near-infrared spectroscopy, or PET), along with structural evaluations (like computed tomography, T1-, T2-, PD-, or diffusion-weighted MRI), contains an abundance of spatiotemporally detailed insights into the brain of each patient. In recent times, researchers have directed their focus toward incorporating spatiotemporal constraints to enhance the efficacy of solving the EEG inverse problem.

They've also been exploring the synergistic spatiotemporal resolutions offered by fMRI and EEG in data-centric approaches. For instance, (Liu et al. 2019b) introduced a solution that addressed inverse problems through matrix factorization within empirical Bayesian frameworks by concurrently estimating the present sources and unknown temporal basis functions through data-guided means. (Liu et al. 2021) employed "covariance components (CCs) derived from clusters defined by fMRI and EEG signals as spatial priors within the empirical Bayesian framework" (p. 14) to facilitate EEG/fMRI fusion.

#### 4.10 Challenges and prospects

This section examines the motivations, merits (contributions), and drawbacks (challenges) of AI's integration in the healthcare sector, addressing concerns such as data scarcity and biases. We further propose future research directions. These directions highlight the significance of employing various fusion methods and developing trustworthy and explainable AI (XAI) to tackle these challenges and advance AI-driven smart healthcare based on multimodal data analysis.

The motivations driving the implementation of AI applications in healthcare are centered on enhancing patient outcomes, refining diagnostic accuracy, optimizing treatment plans, reducing healthcare expenses, enabling personalized medicine, and facilitating improved resource allocation and management (Ahmed et al. 2020). The strengths of AI applications in healthcare primarily revolve around enhancing patient outcomes and overall healthcare system efficiency (Müller et al. 2020). The challenges and weaknesses related to AI applications in the healthcare sector encompass the interpretability of AI models, regulatory hurdles, integration issues with existing healthcare systems, potential biases in algorithms, ethical concerns regarding privacy and consent, and the continuous need for validation and adaptation to ensure clinical relevance and safety.

The challenge of data scarcity in training AI and DNNs poses a significant obstacle for healthcare systems, causing limitations in AI and deep learning utilization. The scarcity, especially concerning diverse multimodal healthcare data, arises due to privacy concerns, data silos, and limited access. Addressing this scarcity is vital to unlock AI's full potential in improving bone classification accuracy and aiding clinical decision-making (Alzubaidi et al. 2023).

To tackle these challenges, potential solutions are proposed. Firstly, addressing data scarcity involves initiating model training with sizable and diverse datasets. This approach enhances a model's capability in learning and recognizing patterns, ensuring its generalizability to new instances. Strategies like feature fusion in information fusion techniques integrate information from multiple sources (e.g., MRI scans, patient records, genetic data, and wearable devices) to create more informative datasets for predictive modeling or diagnosis (Alammar et al. 2023; Al-Timemy et al. 2023; Jebur et al. 2023), providing healthcare providers with detailed insights into patient conditions (Ali et al. 2020). For instance, (Alammar et al. 2023) described an approach for detecting abnormalities in X-ray images based on deep transfer learning with advanced feature fusion, showcasing superior results in humerus and wrist classification compared to prior methods.

Another solution involves utilizing federated learning to train models across decentralized data sources without sharing raw data. While this approach allows collaborative training to improve a shared global deep learning model (Rodríguez-Barroso et al. 2023), data fusion technology introduces challenges such as heterogeneous and

multi-source data fusion. Addressing these challenges involves improving data and model utilization, removing repetitive information, and consolidating various data origins to acquire valuable understanding. Future concerns include preserving user confidentiality, establishing models applicable across the board, data augmentation techniques, cross-institutional collaborations for privacy-compliant data sharing, and guaranteeing data fusion result stability across domains (Alzubaidi et al. 2023).

Secondly, mitigating bias risks and ensuring AI and deep learning trustworthiness necessitates the development of trustworthy and explainable AI. XAI aims to elucidate AI models' inner workings and decision-making processes, making AI-powered applications more ethical, private, secure, trustworthy, confident, and safe (Albahri et al. 2023). The integration of trustworthy AI into healthcare systems offers various advantages, including aiding disease diagnosis, promoting patient care, enhancing privacy, and reducing treatment costs and durations. Numerous studies discuss XAI's explainability in healthcare applications (e.g., (Lucieri et al. 2022; Martínez-Agüero et al. 2022; Deperlioglu et al. 2022; Arrieta et al. 2020; Yang et al. 2022; El-Sappagh et al. 2021; Rahman et al. 2021), using methods such as textual explanations, Shapley values, Class activation mapping, and heatmap, across tasks like melanoma classification, disease diagnosis, and evaluating sustainability features in healthcare applications. For example, in Martínez-Agüero et al. (2022), Shapley values were employed to create reliable intelligent systems to forecast antimicrobial multidrug resistance early.

Additionally, integrating modern information fusion techniques with XAI ensures responsible and ethical utilization of sensitive data (Abdar et al. 2023). Studies like (Yang et al. 2022) explored XAI's advancements in healthcare, proposing XAI solutions that leverage multimodal and multi-center data fusion for clinical use. Utilizing information fusion can revolutionize solutions bridging research and practical applications in trustworthy healthcare AI (Holzinger et al. 2021). By amalgamating multiple information sources like graph analysis and feature visualization, data fusion enhances AI systems' accuracy. Yet, AI explainability faces challenges in handling diverse feature representation spaces, necessitating robust AI systems capable of responsibly managing complex and diverse data (Oprescu et al. 2022). These difficulties emphasize the critical need for AI applications capable of managing varied data in a socially responsible way.

#### 4.11 Limitations, reflections, and future work

This section discusses the limitations of this study. Initially, we only considered journal articles with SSCI and SCI indices in the WoS that have undergone rigorous peer review. As a result, not all papers related to AI-driven smart healthcare through multimodal data analysis were included. During the data search, we initially explored other databases like Springer and Scopus. Our search across multiple databases yielded over ten thousand records. However, a preliminary check of five hundred randomly selected records revealed significant noise and low relevance within this extensive dataset. Conversely, when we focused on SSCI and SCI-indexed articles, most were highly relevant, covering the primary concerns in this field. Considering efficiency, reliability, and resource costs, we chose to utilize SSCI and SCI-indexed journals. The data analysis confirmed this decision's efficacy, identifying key issues. Yet, for future endeavors, we propose to develop a strategy to streamline screening within vast datasets from multiple databases, enhancing comprehensive results.

Regarding methodology, we employed topic modeling-based bibliometric analysis, distinct from systematic analysis focusing on a limited number of articles based on predefined dimensions (e.g., the weakness, strengths, and challenges of each application of AI or datasets and achieved results on them). Topic modeling-based bibliometrics offer rapid, automatic large-scale data analysis, extracting crucial topics not confined to AI applications or datasets used. This methodology, notably popular in various fields related to computer science [e.g., (Chen et al. 2023a; Cui et al. 2023b; Jimma 2023)], aids in comprehending research progress, technology development, and novel ideas. Nonetheless, it may lack the depth of manual coding and meta-analysis studies.

While our results largely cover major issues in AI-powered multimodal data fusion in smart healthcare research, overlapped words and conceptual ambiguities might lead to some issues not being detected. Future work could involve systematic, qualitative analyses to achieve more comprehensive results. This entails a quality assessment to filter papers meeting predefined criteria.

Despite limitations, our study fulfills the aim of presenting an overview of the status, tendencies, and thematic structure of AI-powered multimodal data fusion in smart healthcare research. Future efforts could merge topic models with manual techniques for a more robust understanding of the field.

## 5 Conclusion, contributions, and significance

To unravel the themes and their progression in the realm of AI-powered smart healthcare based on multimodal data analysis, this paper engaged in an exploration of scientific literature via the avenues of topic modeling and bibliometric analysis. Beyond merely identifying the cutting-edge areas of research, this study delved into the ebb and flow of topics via a non-parametric trend assessment. An examination of the annual scholarly output within this interdisciplinary domain exhibited mounting enthusiasm among researchers. Noteworthy were the cross-disciplinary publishing outlets that bridge Healthcare and medical research with Information technology and AI, demonstrating active involvement in this sphere. Leading nations like China, the USA, and India stood out in terms of prolific contributions, accounting for more than 54% of the dataset. The foremost academic entities encompassed Chinese Academy of Sciences, Shanghai Jiao Tong University, and Fudan University. International collaborative endeavors were found to catalyze improved scholarly achievements and expedited progress. Commonly recurring terms in the scrutinized literature encompassed disease, brain, detection, system, and imaging. The topics that surfaced frequently encompassed *adaptive transformations for enhanced visual data analysis*, *neurodegenerative disease prediction with multi-task diagnostics*, *cross-modality MRI for brain tumor analysis*, *AI-assisted imaging for cancer detection*, and *cancer prognosis through multi-dimensional data analysis*. Eleven research topics that experienced significantly increased interest include: *neurodegenerative disease prediction with multi-task diagnostics*, *cross-modality MRI for brain tumor analysis*, *cancer prognosis through multi-dimensional data analysis*, *IoT-enabled sensory monitoring for health management*, *neuroimaging for cognitive impairment detection*, *AI for emotion recognition and post-stroke assessment*, *AI for neuroimaging-based brain disorder prediction*, *multimedia content analysis for mental health support*, *AI-assisted diagnostics and personalization in healthcare*, *advanced signal processing for sleep and gait disorders*, and *neurological health monitoring via mobile technologies*.

This study offers valuable advantages for future researchers in various aspects. Firstly, it furnishes researchers with a thorough grasp of the current state and progress of AI-driven smart healthcare using multimodal data analysis. Secondly, insights into influential authors, institutions, countries/regions, and journals assist researchers in pinpointing key figures and resources to learn from and potentially collaborate with in this research domain. Findings facilitate the exploration of scientific partnerships and the advancement of AI-powered smart healthcare via collective expertise. Thirdly, the examination of evolving research themes and their significance over time equips upcoming researchers to stay abreast of the most crucial and emerging areas within AI-powered smart healthcare. This aids them in making informed decisions on research directions and aligning their studies with community interests. It helps them swiftly grasp the essence of related topics. Fourthly, researchers can pinpoint the most impactful journals dealing with AI-driven smart healthcare via multimodal data analysis, especially those with international reach. Moreover, the comprehensive insights from this study empower researchers, policymakers, and practitioners to make well-informed decisions regarding involvement in AI and multimodal data fusion in smart healthcare endeavors. Furthermore, the utilization of STM and bibliometric analysis, leveraging large-scale scientific data, contributes methodologically, offering a systematic framework for analyzing underlying topics and developmental trends in academic or practical fields. This methodology could be replicated or enhanced by future researchers for their investigations.

This study contributes to the cross-disciplinary exploration centered on AI-powered smart healthcare grounded in the amalgamation of multimodal data. It offers insightful and valuable implications, which can serve as a guiding light for scholars, policymakers, and practitioners aiming to grasp the panoramic view and structure of this ever-important domain. The identified productive entities also hold potential as exemplars and potential collaborators for researchers. Moreover, there is room to enhance and deepen academic partnerships, leading to a more comprehensive exploration of the benefits and hurdles of AI, particularly those rooted in deep learning, to multimodal data fusion in order to optimize decision-making.

The findings furnished bestow actionable insights onto researchers, illuminating the decision-making path by unraveling critical themes within the literature. Varied AI methods have permeated the landscape of medical and multimodal data fusion, showing considerable promise for future advancement. Future endeavors should transcend the inquiry into the mere feasibility of adopting deep learning technologies to facilitate multimodal health/medical data fusion, expanding into the realm of orchestrating the most effective integration of diverse technologies to facilitate potent multimodal data fusion.

Attention should also be dedicated to cutting-edge technologies like contrastive learning and GANs, and their synergies with different fusion strategies (such as multi-task, multi-sensor, multi-dimensional, and multimedia content). This amalgamation could foster efficient automated disease diagnosis, bolster computer-assisted prognosis/prediction, and construct intelligent health/medical systems, ultimately driving the evolution of and advancement in smart healthcare.

**Acknowledgements** The research has been supported by the National Natural Science Foundation of China (No. 62307010), Shenzhen University-Lingnan University Joint Research Programme (SZU-LU006/2122), the research grant entitled "Medical Text Feature Representations based on Pre-trained Language Models" (871238), the Direct Grant (DR23B2), and the Faculty Research Grants (DB23A3 and DB23B2) of Lingnan University, Hong Kong.

**Author contributions** X. Chen and H. Xie wrote the main manuscript text and X. Chen, X. Tao and F. Wang conducted the data analysis, X. Chen, M. Leng and B. Lei conducted data visualization and identified the potential issues and directions from the results.

## Declarations

**Competing interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abdar M, Salari S, Qahremani S, Lam H-K, Karray F, Hussain S, Khosravi A, Acharya UR, Makarenkov V, Nahavandi S (2023) UncertaintyFuseNet: robust uncertainty-aware hierarchical feature fusion model with ensemble Monte Carlo dropout for COVID-19 detection. *Inf Fusion* 90:364–381
- Ahmed Z, Mohamed K, Zeeshan S, Dong X (2020) Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. Database. <https://doi.org/10.1093/database/baaa010>
- Airoldi EM, Bischof JM (2016) Improving and evaluating topic models and other models of text. *J Am Stat Assoc* 111:1381–1403
- Alammar Z, Alzubaidi L, Zhang J, Li Y, Lafta W, Gu Y (2023) Deep transfer learning with enhanced feature fusion for detection of abnormalities in x-ray images. *Cancers (basel)* 15:4007
- Albahri AS, Duhaim AM, Fadhel MA, Alnoor A, Baqer NS, Alzubaidi L, Albahri OS, Alamoodi AH, Bai J, Salhi A (2023) A systematic review of trustworthy and explainable artificial intelligence in healthcare: assessment of quality, bias risk, and data fusion. *Inf Fusion*. <https://doi.org/10.1016/j.inffus.2023.03.008>
- Ali F, El-Sappagh S, Islam SMR, Kwak D, Ali A, Imran M, Kwak K-S (2020) A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. *Inf Fusion* 63:208–222
- Ali F, El-Sappagh S, Islam SMR, Ali A, Attique M, Imran M, Kwak K-S (2021) An intelligent healthcare monitoring framework using wearable sensors and social networking data. *Futur Gener Comput Syst* 114:23–43
- Al-Timemy AH, Alzubaidi L, Mosa ZM, Abdelmotaal H, Ghaeb NH, Lavric A, Hazarbasanov RM, Takahashi H, Gu Y, Yousefi S (2023) A deep feature fusion of improved suspected keratoconus detection with deep learning. *Diagnostics* 13:1689
- Alzubaidi L, Bai J, Al-Sabaawi A, Santamaría J, Albahri AS, Al-dabbagh BSN, Fadhel MA, Manoufali M, Zhang J, Al-Timemy AH (2023) A survey on deep learning tools dealing with data scarcity: definitions, challenges, solutions, tips, and applications. *J Big Data* 10:46
- Amin J, Sharif M, Gul N, Yasmin M, Shad SA (2020a) Brain tumor classification based on DWT fusion of MRI sequences using convolutional neural network. *Pattern Recognit Lett* 129:115–122
- Amin J, Sharif M, Yasmin M, Fernandes SL (2020b) A distinctive approach in brain tumor detection and classification using MRI. *Pattern Recognit Lett* 139:118–127
- Arrieta AB, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, García S, Gil-López S, Molina D, Benjamins R (2020) Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf Fusion* 58:82–115
- Bastian M, Heymann S, Jacomy M (2009) Gephi: an open source software for exploring and manipulating networks Third Int. AAAI Conf. Weblogs Soc. Media. *ICWSM*. 3(1):361–362
- Cai Q, Wang H, Li Z, Liu X (2019) A survey on multimodal data-driven smart healthcare systems: approaches and applications. *IEEE Access* 7:133583–133599



- Chakraborty J, Nandy A (2020) Discrete wavelet transform based data representation in deep neural network for gait abnormality detection. *Biomed Signal Process Control* 62:102076
- Chen RJ, Lu MY, Wang J, Williamson DFK, Rodig SJ, Lindeman NI, Mahmood F (2020) Pathomic fusion: an integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. *IEEE Trans Med Imaging* 41:757–770
- Chen X, Xie H, Cheng G, Li Z (2022) A decade of sentic computing: topic modeling and bibliometric analysis. *Cognit Comput* 14:24–47
- Chen X, Xie H, Li Z, Cheng G, Leng M, Wang FL (2023a) Information fusion and artificial intelligence for smart healthcare: a bibliometric study. *Inf Process Manag* 60:103113
- Chen X, Zou D, Cheng G, Xie H, Jong M (2023b) Blockchain in smart education: contributors, collaborations, applications and research topics. *Educ Inf Technol* 28:4597–4627
- Cui J, Wang Z, Ho S-B, E. (2023a) Cambria, Survey on sentiment analysis: evolution of research methods and topics. *Artif Intell Rev* 56:8469–8510
- Cui W, Tang J, Zhang Z, Dai X (2023b) A bibliometric analysis on innovation convergence. *Libr Hi Tech* 41:333–354
- De la Vega Hernández IM, Urdaneta AS, Carayannis E (2023) Global bibliometric mapping of the frontier of knowledge in the field of artificial intelligence for the period 1990–2019. *Artif Intell Rev* 56:1699–1729
- Demir EK (2021) The role of social capital for teacher professional learning and student achievement: a systematic literature review. *Educ Res Rev* 33:100391
- Deperlioglu O, Kose U, Gupta D, Khanna A, Giampaolo F, Fortino G (2022) Explainable framework for Glaucoma diagnosis by image processing and convolutional neural network synergy: analysis with doctor evaluation. *Futur Gener Comput Syst* 129:152–169
- El-Sappagh S, Alonso JM, Islam SMR, Sultan AM, Kwak KS (2021) A multilayer multimodal detection and prediction model based on explainable artificial intelligence for Alzheimer's disease. *Sci Rep* 11:2660
- Flores AM, Demsas F, Leeper NJ, Ross EG (2021) Leveraging machine learning and artificial intelligence to improve peripheral artery disease detection, treatment, and outcomes. *Circ Res* 128:1833–1850
- Gao X, Shi F, Shen D, Liu M (2021) Task-induced pyramid and attention GAN for multimodal brain image imputation and classification in alzheimer's disease. *IEEE J Biomed Heal Inform* 26:36–43
- Ghosh S, Ekbal A, Bhattacharyya P (2022) A multitask framework to detect depression, sentiment and multi-label emotion from suicide notes. *Cognit Comput*. <https://doi.org/10.1038/s41598-022-08438-z>
- Goodell JW, Kumar S, Lim WM, Pattnaik D (2021) Artificial intelligence and machine learning in finance: identifying foundations, themes, and research clusters from bibliometric analysis. *J Behav Exp Financ* 32:100577
- Guo Y, Hao Z, Zhao S, Gong J, Yang F (2020) Artificial intelligence in health care: bibliometric analysis. *J Med Internet Res* 22:e18228
- Hartl D, de Luca V, Kostikova A, Laramie J, Kennedy S, Ferrero E, Siegel R, Fink M, Ahmed S, Millholland J (2021) Translational precision medicine: an industry perspective. *J Transl Med* 19:1–14
- Holzinger A, Malle B, Saranti A, Pfeifer B (2021) Towards multi-modal causability with graph neural networks enabling information fusion for explainable AI. *Inf Fusion* 71:28–37
- Holzinger A, Dehmer M, Emmert-Streib F, Cucchiara R, Augenstein I, Del Ser J, Samek W, Jurisica I, Díaz-Rodríguez N (2022) Information fusion as an integrative cross-cutting enabler to achieve robust, explainable, and trustworthy medical artificial intelligence. *Inf Fusion* 79:263–278
- Hu Y, Modat M, Gibson E, Li W, Ghavamini N, Bonmati E, Wang G, Bandula S, Moore CM, Emberton M (2018) Weakly-supervised convolutional neural networks for multimodal image registration. *Med Image Anal* 49:1–13
- Huang W, Hew KF, Fryer LK (2022a) Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *J Comput Assist Learn* 38:237–257
- Huang P, Li D, Jiao Z, Wei D, Cao B, Mo Z, Wang Q, Zhang H, Shen D (2022b) Common feature learning for brain tumor MRI synthesis by context-aware generative adversarial network. *Med Image Anal* 79:102472
- Huynh E, Hosny A, Guthrie C, Bitterman DS, Petit SF, Haas-Kogan DA, Kann B, Aerts HJWL, Mak RH (2020) Artificial intelligence in radiation oncology. *Nat Rev Clin Oncol* 17:771–781
- Jebur SA, Hussein KA, Hoomod HK, Alzubaidi L (2023) Novel deep feature fusion framework for multi-scenario violence detection. *Computers* 12:175
- Jeyaraj A, Zadeh AH (2020) Evolution of information systems research: insights from topic modeling. *Inf Manag* 57:103207
- Jimma BL (2023) Artificial intelligence in healthcare: a bibliometric analysis *Telemat. Inform Rep* 9:100041

- Jin K, Yan Y, Chen M, Wang J, Pan X, Liu X, Liu M, Lou L, Wang Y, Ye J (2022) Multimodal deep learning with feature level fusion for identification of choroidal neovascularization activity in age-related macular degeneration. *Acta Ophthalmol* 100:e512–e520
- Katsigiannis S, Ramzan N (2017) DREAMER: a database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices. *IEEE J Biomed Heal Inform* 22:98–107
- Khan MA, Ashraf I, Alhaisoni M, Damaševićus R, Scherer R, Rehman A, Bukhari SAC (2020) Multimodal brain tumor classification using deep learning and robust feature selection: a machine learning application for radiologists. *Diagnostics* 10:565
- Lin K, Li Y, Sun J, Zhou D, Zhang Q (2020) Multi-sensor fusion for body sensor network in medical human–robot interaction scenario. *Inf Fusion* 57:15–26
- Liu S, Liu S, Cai W, Che H, Pujol S, Kikinis R, Feng D, Fulham MJ (2014) Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer’s disease. *IEEE Trans Biomed Eng* 62:1132–1140
- Liu X, Cao P, Wang J, Kong J, Zhao D (2019a) Fused group lasso regularized multi-task feature learning and its application to the cognitive performance prediction of Alzheimer’s disease. *Neuroinformatics* 17:271–294
- Liu K, Yu ZL, Wu W, Gu Z, Zhang J, Cen L, Nagarajan S, Li Y (2019b) Bayesian electromagnetic spatio-temporal imaging of extended sources based on matrix factorization. *IEEE Trans Biomed Eng* 66:2457–2469
- Liu K, Yu ZL, Wu W, Chen X, Gu Z, Guan C (2021) fMRI-SI-STBF: an fMRI-informed Bayesian electromagnetic spatio-temporal extended source imaging. *Neurocomputing* 462:14–30
- Liu X, Chen H, Yao C, Xiang R, Zhou K, Du P, Liu W, Liu J, Yu Z (2023) BTMF-GAN: a multi-modal MRI fusion generative adversarial network for brain tumors. *Comput Biol Med* 157:106769
- Lu S, Liu M, Yin L, Yin Z, Liu X, Zheng W (2023) The multi-modal fusion in visual question answering: a review of attention mechanisms. *PeerJ Comput Sci* 9:e1400
- Lucieri A, Bajwa MN, Braun SA, Malik MI, Dengel A, Ahmed S (2022) ExAID: a multimodal explanation framework for computer-aided diagnosis of skin lesions. *Comput Methods Programs Biomed* 215:106620
- Luo X, Gao Y, Wang A, Zhang Z, Wu X-J (2021) IFSepR: a general framework for image fusion based on separate representation learning. *IEEE Trans Multimed* 25:608–623
- Ma M, Sun C, Chen X (2018) Deep coupling autoencoder for fault diagnosis with multimodal sensory data. *IEEE Trans Ind Inform* 14:1137–1145
- Mann HB (1945) Nonparametric tests against trend. *Econom J Econom Soc* 13:245–259
- Martínez-Agüero S, Soguero-Ruiz C, Alonso-Moral JM, Mora-Jiménez I, Álvarez-Rodríguez J, Marques AG (2022) Interpretable clinical time-series modeling with intelligent feature selection for early prediction of antimicrobial multidrug resistance. *Futur Gener Comput Syst* 133:68–83
- Mi J, Wang L, Liu Y, Zhang J (2022) KDE-GAN: a multimodal medical image-fusion model based on knowledge distillation and explainable AI modules. *Comput Biol Med* 151:106273
- Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 151:264–269
- Mohsen F, Ali H, El Hajj N, Shah Z (2022) Artificial intelligence-based methods for fusion of electronic health records and imaging data. *Sci Rep* 12:17981
- Muhammad G, Alshehri F, Karray F, El Saddik A, Alsulaiman M, Falk TH (2021) A comprehensive survey on multimodal medical signals fusion for smart healthcare systems. *Inf Fusion* 76:355–375
- Müller J, Stoehr M, Oeser A, Gaebel J, Streit M, Dietz A, Oeltze-Jafra S (2020) A visual approach to explainable computerized clinical decision support. *Comput Graph* 91:1–11
- Mustak M, Salminen J, Plé L, Wirtz J (2021) Artificial intelligence in marketing: topic modeling, scientometric analysis, and research agenda. *J Bus Res* 124:389–404
- Nema S, Dudhane A, Murala S, Naidu S (2020) RescueNet: an unpaired GAN for brain tumor segmentation. *Biomed Signal Process Control* 55:101641
- Nguyen DC, Pham Q-V, Pathirana PN, Ding M, Seneviratne A, Lin Z, Dobre O, Hwang W-J (2022) Federated learning for smart healthcare: a survey. *ACM Comput Surv*. <https://doi.org/10.1145/3501296>
- Noorbakhsh-Sabet N, Zand R, Zhang Y, Abedi V (2019) Artificial intelligence transforms the future of health care. *Am J Med* 132:795–801
- de Oliveira OJ, da Silva FF, Juliani F, Barbosa LCFM, Nunhes TV (2019) Bibliometric method for mapping the state-of-the-art and identifying research gaps and trends in literature: an essential instrument to support the development of scientific projects. *Sci Recent Adv, IntechOpen*.
- Opreescu AM, Miró-Amarante G, García-Díaz L, Rey VE, Chimenea-Toscano A, Martínez-Martínez R, Romero-Ternero MC (2022) Towards a data collection methodology for responsible artificial intelligence in health: a prospective and qualitative study in pregnancy. *Inf Fusion* 83:53–78

- Pian W, Chi J, Ma F (2021) The causes, impacts and countermeasures of COVID-19 “Infodemic”: a systematic review using narrative synthesis. *Inf Process Manag* 58:102713
- Principe VA, de Souza Vale RG, de Castro JBP, Carvano LM, Henriques RAP, de Almeida VJ, e Sousa Lobo, R. de Alkmim Moreira Nunes, (2022) A computational literature review of football performance analysis through probabilistic topic modeling. *Artif Intell Rev* 55:1351–1371
- Qin Z, Liu Z, Zhu P, Ling W (2022) Style transfer in conditional GANs for cross-modality synthesis of brain magnetic resonance images. *Comput Biol Med* 148:105928
- Rahman MA, Hossain MS, Showail AJ, Alrajeh NA, Alhamid MF (2021) A secure, private, and explainable IoT framework to support sustainable health monitoring in a smart city. *Sustain Cities Soc* 72:103083
- Rao D, Xu T, Wu X-J (2023) Tgfuse: an infrared and visible image fusion approach based on transformer and generative adversarial network. *IEEE Trans Image Process*. <https://doi.org/10.48550/arXiv.2201.10147>
- Razfar N, Kashef R, Mohammadi F (2023) An artificial intelligence model for smart post-stroke assessment using wearable sensors. *Decis Anal J* 7:100218
- Roberts ME, Stewart BM, Tingley D, Lucas C, Leder-Luis J, Gadarian SK, Albertson B, Rand DG (2014) Structural topic models for open-ended survey responses. *Am J Pol Sci* 58:1064–1082
- Rodríguez-Barroso N, Jiménez-López D, Luzón MV, Herrera F, Martínez-Cámara E (2023) Survey on federated learning threats: concepts, taxonomy on attacks and defences, experimental study and challenges. *Inf Fusion* 90:148–173
- Sawhney R, Joshi H, Gandhi S, Shah R (2020) A time-aware transformer based model for suicide ideation detection on social media, In: *Proc. 2020 Conf. Empir. Methods Nat. Lang. Process.*, pp. 7685–7697.
- Shaik T, Tao X, Li L, Xie H, Velásquez JD (2023) A survey of multimodal information fusion for smart healthcare: mapping the journey from data to wisdom. *Inf Fusion* 102:102040. <https://doi.org/10.1016/j.inffus.2023.102040>
- Shao W, Han Z, Cheng J, Cheng L, Wang T, Sun L, Lu Z, Zhang J, Zhang D, Huang K (2019) Integrative analysis of pathological images and multi-dimensional genomic data for early-stage cancer prognosis. *IEEE Trans Med Imaging* 39:99–110
- Shao W, Peng Y, Zu C, Wang M, Zhang D, Initiative ADN (2020) Hypergraph based multi-task feature selection for multimodal classification of Alzheimer’s disease. *Comput Med Imaging Graph* 80:101663
- Shao B, Li X, Bian G (2021) A survey of research hotspots and frontier trends of recommendation systems from the perspective of knowledge graph. *Expert Syst Appl* 165:113764
- Shi J, Zheng X, Li Y, Zhang Q, Ying S (2017) Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer’s disease. *IEEE J Biomed Heal Inform* 22:173–183
- Soni N, Sharma EK, Singh N, Kapoor A (2020) Artificial intelligence in business: from research and innovation to market deployment. *Procedia Comput Sci* 167:2200–2210
- Sujith A, Sajja GS, Mahalakshmi V, Nuhmani S, Prasanalakshmi B (2022) Systematic review of smart health monitoring using deep learning and artificial intelligence. *Neurosci Informatics* 2:100028
- Suk H-I, Lee S-W, Shen D, Initiative ADN (2014) Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *Neuroimage* 101:569–582
- Sun F, Zang W, Gravina R, Fortino G, Li Y (2020) Gait-based identification for elderly users in wearable healthcare systems. *Inf Fusion* 53:134–144
- Swacha J (2021) State of research on gamification in education: a bibliometric survey. *Educ Sci* 11:69
- Tandon A, Kaur P, Mäntymäki M, Dhir A (2021) Blockchain applications in management: a bibliometric analysis and literature review. *Technol Forecast Soc Change* 166:120649
- Tang W, He F, Liu Y, Duan Y (2022) MATR: multimodal medical image fusion via multiscale adaptive transformer. *IEEE Trans Image Process* 31:5134–5149
- Tian S, Yang W, Le Grange JM, Wang P, Huang W, Ye Z (2019) Smart healthcare: making medical care more intelligent. *Glob Heal J* 3:62–65
- Venugopalan J, Tong L, Hassanzadeh HR, Wang MD (2021) Multimodal deep learning models for early detection of Alzheimer’s disease stage. *Sci Rep* 11:3254
- Wamba SF, Bawack RE, Guthrie C, Queiroz MM, Carillo KDA (2021) Are we preparing for a good AI society? A bibliometric review and research agenda. *Technol Forecast Soc Change* 164:120482
- Wang Z, Ma Y (2008) Medical image fusion using m-PCNN. *Inf Fusion* 9:176–185
- Wang H, Tan L, Niu B (2019) Feature selection for classification of microarray gene expression cancers using bacterial colony optimization with multi-dimensional population. *Swarm Evol Comput* 48:172–181

- Wang S, Zheng K, Kong W, Huang R, Liu L, Wen G, Yu Y (2023) Multimodal data fusion based on IGERNNC algorithm for detecting pathogenic brain regions and genes in Alzheimer's disease. *Brief Bioinform* 24:515
- Wu J, Wang K, He C, Huang X, Dong K (2021a) Characterizing the patterns of China's policies against COVID-19: a bibliometric study. *Inf Process Manag* 58:102562
- Wu X, Liu C, Wang L, Bilal M (2021b) Internet of things-enabled real-time health monitoring system using deep learning. *Neural Comput Appl* 35(20):14565–14576
- Yang F, Wu Q, Hu X, Ye J, Yang Y, Rao H, Ma R, Hu B (2021) Internet-of-things-enabled data fusion method for sleep healthcare applications. *IEEE Internet Things J* 8:15892–15905
- Yang G, Ye Q, Xia J (2022) Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: a mini-review, two showcases and beyond. *Inf Fusion* 77:29–52
- Yin M, Liu X, Liu Y, Chen X (2018) Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampling shearlet transform domain. *IEEE Trans Instrum Meas* 68:49–64
- Zhang J, Yin Z, Chen P, Nichele S (2020a) Emotion recognition using multi-modal data and machine learning techniques: a tutorial and review. *Inf Fusion* 59:103–126
- Zhang Y, Liu Y, Sun P, Yan H, Zhao X, Zhang L (2020b) IFCNN: a general image fusion framework based on convolutional neural network. *Inf Fusion* 54:99–118
- Zhang L, Wang L, Gao J, Risacher SL, Yan J, Li G, Liu T, Zhu D, Initiative ADN (2021) Deep fusion of brain structure-function in mild cognitive impairment. *Med Image Anal* 72:102082
- Zhang Y, Nie R, Cao J, Ma C (2023) Self-supervised fusion for multi-modal medical images via contrastive auto-encoding and convolutional information exchange. *Ieee Comput Intell Mag* 18:68–80
- Zhu X, Suk H-I, Lee S-W, Shen D (2015) Subspace regularized sparse multitask learning for multiclass neurodegenerative disease identification. *IEEE Trans Biomed Eng* 63:607–618

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.