

# Multimedia Tools and Applications

## Current Status and Trends of Technology, Methods, and Applications of Human-computer Intelligent Interaction (HCII): A Bibliometric Research

--Manuscript Draft--

<b>Manuscript Number:</b>	MTAP-D-23-03291R1							
<b>Full Title:</b>	Current Status and Trends of Technology, Methods, and Applications of Human-computer Intelligent Interaction (HCII): A Bibliometric Research							
<b>Article Type:</b>	Track 1: General Multimedia Topics							
<b>Keywords:</b>	Human-Computer Intelligent Interaction, Human-Computer Interaction, Artificial intelligence, Bibliometric Research, CiteSpace							
<b>Corresponding Author:</b>	Yan Gan Huazhong University of Science and Technology - Main Campus: Huazhong University of Science and Technology Wuhan, CHINA							
<b>Corresponding Author Secondary Information:</b>								
<b>Corresponding Author's Institution:</b>	Huazhong University of Science and Technology - Main Campus: Huazhong University of Science and Technology							
<b>Corresponding Author's Secondary Institution:</b>								
<b>First Author:</b>	Zi jie Ding							
<b>First Author Secondary Information:</b>								
<b>Order of Authors:</b>	Zi jie Ding Ying rui Ji Yan Gan Yu wen Wang Yu kun Xia							
<b>Order of Authors Secondary Information:</b>								
<b>Funding Information:</b>	<table border="1"><tr><td>National Laboratory of Mechanical Systems and Vibrations of China (MSV202014)</td><td>Mrs. Yan Gan</td></tr><tr><td>Funda-mental Research Funds for the Central University of China (HUST:2020kfyXJJS014)</td><td>Mrs. Yan Gan</td></tr><tr><td>the Major Philosophy and Social Science Research Project for Hubei Province Colleges and Universities of China (20ZD002)</td><td>Mrs. Yan Gan</td></tr></table>		National Laboratory of Mechanical Systems and Vibrations of China (MSV202014)	Mrs. Yan Gan	Funda-mental Research Funds for the Central University of China (HUST:2020kfyXJJS014)	Mrs. Yan Gan	the Major Philosophy and Social Science Research Project for Hubei Province Colleges and Universities of China (20ZD002)	Mrs. Yan Gan
National Laboratory of Mechanical Systems and Vibrations of China (MSV202014)	Mrs. Yan Gan							
Funda-mental Research Funds for the Central University of China (HUST:2020kfyXJJS014)	Mrs. Yan Gan							
the Major Philosophy and Social Science Research Project for Hubei Province Colleges and Universities of China (20ZD002)	Mrs. Yan Gan							
<b>Abstract:</b>	<p>This study delves into Human-Computer Intelligent Interaction (HCII), a burgeoning interdisciplinary field that builds upon traditional Human-Computer Interaction (HCI) by integrating advanced technologies like Natural Language Processing (NLP) and Machine Learning (ML). In this paper, we scrutinize 5,781 HCII papers published between 2000 and 2023, narrowing our focus to 803 most relevant articles to construct co-citation and interdisciplinary networks based on the CiteSpace Software. Our findings reveal that the publications of the United States and China are relatively high with 558 and 616 publications respectively. Furthermore, we found that machine learning and deep learning have emerged as the prevalent methodologies in HCII, which currently emphasizes multimodal emotion recognition, facial expression recognition, and NLP. We predict that HCII will be integrated into advanced applications such as neural-based interactive games and multi-sensory environments. In sum, our analysis underscores HCII's role in advancing artificial intelligence, facilitating more intuitive and efficient human-computer interactions, and its prospective societal</p>							

impact. We hope that our review and analysis may guide the efforts of researchers aiming to contribute to HCI and develop more powerful and intelligent methods, tools, and applications.

**Revisions and Response to Reviewers**  
**Current Status and Trends of Technology, Methods, and Applications of**  
**Human-computer Intelligent Interaction (HCII):**  
**A Bibliometric Research**  
**Paper Number: MTAP-D-23-03291**

*Dear Editor and dear reviewers,*

We would like to thank you for your comments and suggestions, which gave us the opportunity to improve the paper. In the revised manuscript, we hope to solve all the issues raised. In this document we answer to all the questions raised by the reviewers. Comments are shown in black font, followed by our answer/comment in red font. The major corrections/changes in the manuscript are displayed in red font.

*Response to reviewers' comments:*

*Reviewer #1*

1 In the introduction I expect more motivation for this review, a clear list of research questions and the contributions in an itemize. Then, some sections can be renamed based on the research questions they aim to answer with graphics and discussion therein.

**[Reply]** Thanks for reviewer's positive evaluation on our manuscript. Your insightful comments have greatly helped us improve the work. The point-by-point responses are as below. We have added a paragraph to list of research questions and the contributions in the first paragraph.

**The added/revised content:**

( 1 ) Research questions : Recent advancements in Human-Computer Intelligent Interaction (HCII) have yielded a plethora of applications in diverse domains, including emotion and gesture recognition methods [33-41], as well as the integration of Augmented Reality (AR) in healthcare, education, and industrial settings [42-45]. These technologies will enable users to have more natural and immersive interactions with computers. Users will be able to operate computers or other digital devices using natural gestures, and they can also connect human thoughts to computers through brain-computer interfaces, facilitating a more direct and efficient mode of interaction. These technologies will significantly make our lives more convenient and intelligent. Despite these strides, a conflation often exists between HCII and its predecessor, Human-

Computer Interaction (HCI), resulting in a vague understanding of their distinct characteristics among the general public and even within the academic community.

To bridge this gap and shed light on the evolutionary trajectory and future directions of HCII, this study employs CiteSpace to conduct a comprehensive analysis spanning literature of two decades (2000-2023). The review and analysis are structured around the following three aspects: ① Understanding the relationships and differences among HCII, AI, and HCI, particularly focusing on how AI-driven methodologies have transformed HCII from traditional HCI paradigms; ② Mapping the geographical and institutional contributions to HCII, providing a nuanced understanding of research interests and academic engagements across nations and academic communities; ③ Predicting potential important areas for HCII applications and highlight emerging avenues for innovation. By addressing these aspects, this study aims to provide a rigorous academic framework that can guide and inform future efforts in HCII research and applications.

( 2 ) Research contributions: This study represents a novel endeavor in the Human-Computer Intelligent Interaction (HCII) research landscape, leveraging CiteSpace software for the analysis and visualization of literature. This innovative methodology enables an unprecedented, nuanced understanding of HCII's evolving research themes, hotspots, and international scholarly collaborations. The analysis addresses existing challenges of correlation and clustering in conventional HCII research, thereby offering a robust analytical lens to capture the dynamism and trends shaping this rapidly growing field. The contributions of this study are three-fold: ① This paper rigorously defines HCII, elucidates its symbiotic relationship with Artificial Intelligence (AI) technologies, and outlines the architectural paradigms and technical methodologies underpinning human-computer intelligent interactions; ② Utilizing the CiteSpace Software, this paper visualizes and scrutinizes HCII-related literature spanning two decades. The analysis yields critical insights into emergent research hotspots, thematic clusters, and longitudinal trends in the HCII domain; ③ This paper prognosticates future trajectories for HCII, identifying prospective intelligent interaction paradigms and methodologies, thereby informing researchers with strategic directions and conceptual scaffolding for future research efforts. In summary, this study aims to act as a foundational reference, providing both a thorough overview of current HCII research and a forward-thinking viewpoint to steer future academic and practical efforts.

2 "[...] The following encapsulates prominent artificial intelligence algorithms [...] The field of human-computer interaction [...]". In this part of the text there is some formatting problem, and turn into a hard to follow piece. Please add an itemize.

**[Reply]** Thanks for this suggestion. We have made formatting adjustment , and described these contents in section 2.2 and 2.3, respectively.

**The added/ revised content:**

2.2 AI Algorithms in Human-Computer Intelligent Interaction

2.3 Evolution and Diversification of HCII: From HCI to Multimodal Interactions

3 WoS only? Why not Scopus or IEEEXplore? Please discuss all the biases (including DB-related ones) introduced in the review process.

**[Reply]** Thanks for your valuable feedback and for pointing out the potential sources for the literature search. We appreciate your suggestion to consider Scopus and IEEE. Our choice of using only WoS was based on its comprehensive coverage and our prior experience with its robustness. As you correctly pointed out, this approach may introduce some bias. We are already aware of this limitation. And in future research, we will expand the database to include Scopus, IEEE, and other larger databases. Thank you again for the valuable suggestion.

4 Please provide, in a dedicated folder online, all the references considered for graphics etc.

**[Reply]** According to your advice, we have provided reference materials for all graphics and other materials in the online dedicated folder

5 I hope this work will be followed by a systematic review where to grasp the actual contributions of the articles included in this bibliometric survey.

**[Reply]** This is a good point. We have added sentences to present the the actual contributions in Chapters 5 and 6.

**The added/ revised content:**

(1) Chapters 5: Based on our analysis, the three main centers of contemporary research in Human-Computer Intelligent Interaction (HCII) are as follows: Multimodal and Emotion Recognition Technology: The future of HCII lies in developing more diverse databases and achieving higher accuracy and better user experiences. The objective is to make human-computer dialogues and collaborations more natural and akin to "human-to-human" interactions. New Interaction Media and Scenarios: Innovations in interaction mediums like VR glasses and smart medical care are on the horizon. These mediums will likely make interactions more efficient and versatile.

Advancements in AI Algorithms: Natural language processing has already become mainstream in AI, and the continuous improvements in deep learning algorithms are making computerized natural language more robust, accurate, and diverse.

(2) Chapters 6: There are our main findings: Interaction Modes: The paper identifies and analyzes the five primary interaction modes in HCII—voice interaction, visual interaction, gesture interaction, multimodal interaction, and emotional interaction—contrasting them with traditional HCI to offer a roadmap for future technological advancements in HCII. Research Hotspots: Utilizing CiteSpace for bibliometric analysis, the paper pinpoints the key research areas in HCII over the past five years, including data processing, brain-computer interfaces, multimodal emotion recognition, action recognition, and language processing. Future Trends: The paper predicts broader application scenarios and greater research value for HCII in the future, particularly in the realms of multimodal interaction and emotion recognition. The development of more accurate algorithms and models, complex databases, and intelligent user interfaces is imminent. Innovations like neural interaction games, multi-sensory interaction, and de-physicalized interfaces are on the horizon. Complex and Natural Interactions: The use of advanced algorithms will likely make HCII systems more intuitive, allowing for more complex and natural interactions between humans and computers. Multimodal Emotion Recognition: Advances in this area will enable more nuanced and effective human-computer interactions, further narrowing the gap between human-to-human and human-to-computer communication. Societal Contributions: As HCII continues to evolve, it is set to offer more intelligent, intuitive, and efficient methods of interaction, thereby contributing significantly to the advancement of various sectors of human society. In summary, HCII is at a pivotal point in its development, with the potential for transformative impact on how humans interact with technology. As it continues to integrate advancements from both HCI and AI, HCII is poised to revolutionize various sectors, from healthcare to industrial manufacturing, and beyond.

### **Reviewer #2**

1 Not respecting the empty space at the end of the sentences, the number of the references does not match the order of their presence in the text, and the first letter of the keywords is not capitalized.

**[Reply]** We are appreciating the reviewer very much for positive recommendation and constructive comments to improve our work. We have modified the use of empty space , and we have corrected the writing problem of the text and checked the capitalization of

the first letter of the keyword, also the order of the references has been reorganized in the correct sequence.

2 The plan in Figure 1 is not very accurate and the relationship between the two areas HCI and AI needs to be analyzed and completed.

[Reply] This is a very good point. We have added and revised the caption for Figure 1, and the content of Figure 1 has been redrawn in chapter 2.1.

The added/revised content:

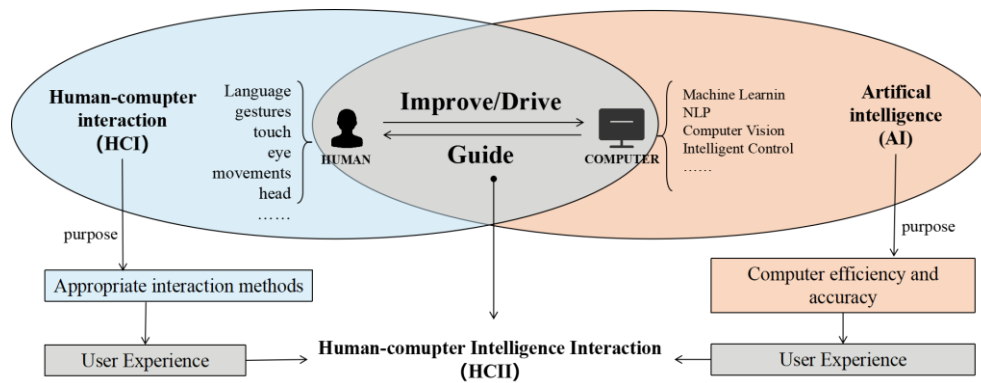


Figure 1. The relationship between AI and HCI in human-computer intelligence interaction

The symbiotic relationship between AI and HCI gives rise to the HCII domain, as illustrated in Figure 1. Specifically, AI-driven techniques such as ML and NLP can be harnessed to refine HCI elements, thereby fostering smarter user interfaces, task automation, and personalized interaction experiences. This synergy further catalyzes the development and application of avant-garde technologies, including autonomous vehicular systems.

3 Only one database (WoS) has been addressed and, for example, Scopus can also complete it.

[Reply] Thank you for your valuable feedback and for pointing out the potential sources for the literature search. We appreciate your suggestion to consider Scopus and IEEE. Our choice of using only WoS was based on its comprehensive coverage and our prior experience with its robustness. As you correctly pointed out, this approach may introduce some bias. We are already aware of this limitation. And in future research, we will expand

the database to include Scopus, IEEE, and other larger databases. Thank you again for the valuable suggestion.

4 The final orientation of the analysis is not very valuable: for example, what is the scientific value of the country or which authors have focused on this field? Why didn't you analyze how was the focus of the methods and techniques and what challenges have been given less attention.

**[Reply ]** Thank you for your valuable suggestions. We have streamlined the country and authors sections. We added an explanation of the methods and technologies predominantly utilized in the current HCII field in Chapter 4. and we also add some discussions on challenges in Chapter 5

**The added/revised content:**

(1) Chapter 4: Brain-Computer Interface: Cluster 5 (Brain-Computer Interface): This cluster is an emerging hot topic that bridges the HCII and biomedical fields. It primarily relies on EEG signals to control external devices, incorporating signal processing, classification techniques, and control theories to manage complex environments or execute fine motions. Applications often target individuals with paralysis and include Deep Learning (DL) technology-based assistive rehabilitation devices with IoT modules, as well as assistive mobility devices for disabled patients. The technology's success has also fueled the growth of the neurogaming industry. In this context, EEG signals form datasets, Support Vector Machines (SVMs) and Long Short-Term Memory networks (LSTMs) train learning algorithms for in-game decision-making, and sensor devices host the game engine for user testing. Research in brain-computer interaction games primarily targets the enhancement of human-computer interaction and user experience and is mainly applied to board games like mahjong and backgammon.

(2) Chapter 5: In alignment with the Industry 4.0 paradigm, HCII in this sector is rapidly incorporating digital twin and IoT technologies. Decision tree algorithms, a classical AI technique, are also being applied to optimize manufacturing processes and decision-making. HCII is increasingly prevalent in healthcare, particularly in psychological diagnosis, psychotherapy, and emotional well-being. Brain-computer interface technology is being integrated into HCII and is expected to find applications in rehabilitation therapy, neuro-gaming, and telemedicine. research in smart homes spans a broad and complex range, from foundational studies to practical applications, aimed at making homes more interactive, secure, and user-friendly. In this sector, the focus is primarily on human body recognition technologies, like head and facial



gesture and action recognition. These technologies aim to improve the robustness and accuracy of HCII systems, leading to more intelligent and natural advanced vehicle systems. This also becomes a challenge for future HCII.

Aerospace: HCII technology could be extremely useful in hazardous flying missions in complex high-altitude environments. AI technology can simulate spatial orientation obstacles and provide enhanced navigation functions. The applications of HCII are not limited to the above-mentioned fields. Future research could extend into space and deep-sea explorations, indicating the limitless potential and versatile applications of HCII. The field promises to continuously evolve, integrating with other emerging technologies to address increasingly complex human-computer interaction challenges across a myriad of scenarios.

### **Reviewer #3**

1 Authors cannot conclude that China has contributed the most to the advancement of the field as much of the fundamental work has been done in US, especially by the industries of silicon valley; I am not american but one cannot propose such an abnormal alteration of the reality, i suggest the author to rephrase this concept and to puit it into a more reasonable framework

**[Reply ]** Thank you for your valuable suggestion. The bias has been revised in abstract and in Chapter 4.2.3.

#### **The added/revised content:**

( 1 ) Abstract: This study delves into Human-Computer Intelligent Interaction (HCII), a burgeoning interdisciplinary field that builds upon traditional Human-Computer Interaction (HCI) by integrating advanced technologies like Natural Language Processing (NLP) and Machine Learning (ML). In this paper, we scrutinize 5,781 HCII papers published between 2000 and 2023, narrowing our focus to 803 most relevant articles to construct co-citation and interdisciplinary networks based on the CiteSpace Software. Our findings reveal that the publications of the United States and China are relatively high with 558 and 616 publications respectively. Furthermore, we found that machine learning and deep learning have emerged as the prevalent methodologies in HCII, which currently emphasizes multimodal emotion recognition, facial expression recognition, and NLP. We predict that HCII will be integrated into advanced applications such as neural-based interactive games and multi-sensory environments. In sum, our analysis underscores HCII's role in advancing artificial intelligence, facilitating more intuitive and efficient human-computer

interactions, and its prospective societal impact. We hope that our review and analysis may guide the efforts of researchers aiming to contribute to HCII and develop more powerful and intelligent methods, tools, and applications.

( 2 ) Chapter 4.2.3 : China (616 publications) and the United States (558 publications) lead the pack with substantial contributions exceeding 500 publications each. Betweenness centrality values indicate that the U.S. (0.75) has a slightly higher influence compared to China (0.55) in the field. While Germany and France have fewer publications, their higher centrality scores indicate that their research contributions are of high quality and serve as pivotal nodes in the global HCII research network.

2 Almost all examined papers which are a lot disregard a vast class of application of intelligent HCI that cannot be ignored, otherwise the paper becomes again biased. I cite as an example the medical application or among many the following one picked in a very specialistic but relevant field as archeology. The authors should provide a comprehensive reflection on this point:

**[Reply ]** Thanks for the reviewer's suggest very much. The HCI method and AI technology used in archaeological, cultural heritage, flight missions, and other literature have been supplemented in Chapter 5.

#### **The added/revised content:**

Aerospace: HCII technology could be extremely useful in hazardous flying missions in complex high-altitude environments. AI technology can simulate spatial orientation obstacles and provide enhanced navigation functions. The applications of HCII are not limited to the above-mentioned fields. Future research could extend into space and deep-sea explorations, indicating the limitless potential and versatile applications of HCII. The field promises to continuously evolve, integrating with other emerging technologies to address increasingly complex human-computer interaction challenges across a myriad of scenarios.

#### **Reviewer #4**

1 find the paper interesting and worthy of publication, I would encourage the authors to refer to and, if possible, apply well-known international guidelines, such as PRISMA (<http://www.prisma-statement.org>). Doing so would strengthen the research and prevent the authors from overlooking important and relevant references.

**[Reply ]** We are appreciating the reviewer very much for positive recommendation , The structure and content have been revised according to the international guidelines on PRISMA.

**Reviewer #7**

1The abstract does not highlight the specifics of the research or findings but contains too much background information. Some details of the research would be nice for example numbers addressing the sample, data, percentage improvement, etc.. Remove some of the background material and add some details of the research. Moreover, it is good to provide some specifics (e.g., sample size, dataset size, numbers from results, etc.).

**[Reply]** Thanks for your kind suggestion. We have modified the abstract with more research details.

**The added/revised content:**

This study delves into Human-Computer Intelligent Interaction (HCII), a burgeoning interdisciplinary field that builds upon traditional Human-Computer Interaction (HCI) by integrating advanced technologies like Natural Language Processing (NLP) and Machine Learning (ML). In this paper, we scrutinize 5,781 HCII papers published between 2000 and 2023, narrowing our focus to 803 most relevant articles to construct co-citation and interdisciplinary networks based on the CiteSpace Software. Our findings reveal that the publications of the United States and China are relatively high with 558 and 616 publications respectively. Furthermore, we found that machine learning and deep learning have emerged as the prevalent methodologies in HCII, which currently emphasizes multimodal emotion recognition, facial expression recognition, and NLP. We predict that HCII will be integrated into advanced applications such as neural-based interactive games and multi-sensory environments. In sum, our analysis underscores HCII's role in advancing artificial intelligence, facilitating more intuitive and efficient human-computer interactions, and its prospective societal impact. We hope that our review and analysis may guide the efforts of researchers aiming to contribute to HCII and develop more powerful and intelligent methods, tools, and applications.

2 Starting from the previous works, I suggest introducing a table to summarize the most recent works and to highlight the novelty of the proposed work.

**[Reply]** This is a very good point. The latest research of 2023 has been added to Chapter 1 to highlight the novelty of the proposed work. And 13 references of 2023 have been updated.

### **The added/revised content:**

Recent advancements in Human-Computer Intelligent Interaction (HCII) have yielded a plethora of applications in diverse domains, including emotion and gesture recognition methods [33-41], as well as the integration of Augmented Reality (AR) in healthcare, education, and industrial settings [42-45]. These technologies will enable users to have more natural and immersive interactions with computers. Users will be able to operate computers or other digital devices using natural gestures, and they can also connect human thoughts to computers through brain-computer interfaces, facilitating a more direct and efficient mode of interaction. These technologies will significantly make our lives more convenient and intelligent.

3-4 Please add more recent references. Certainly, there has been more recent (within the last two years) research on this topic published in information science and/or computer science outlets. An academic search on the topic (using keywords from the manuscript's title) shows that there is recent work in this area. Therefore, authors must update their literature review. There are only few papers published in 2023. Please add new updated references

**[Reply]** Thanks for your kind suggestion , We have added the latest representative research in the text and references and they have been appropriately cited in the references.

### **The added/revised content:**

#### **References:**

- (32) García-Méndez S, Arriba-Pérez FD, González-Castaño FJ et al (2021) Entertainment Chatbot for the Digital Inclusion of Elderly People Without Abstraction Capabilities. *IEEE Access*,pp 75878-75891
- (33) Lee W, Son G (2023) Investigation of human state classification via EEG signals elicited by emotional audio-visual stimulation. *Multimed Tools Appl*
- (34) Razzaq MA, Hussain J, Bang J (2023) Hybrid Multimodal Emotion Recognition Framework for UX Evaluation Using Generalized Mixture Functions. *Sensors* 23:4373
- (35) Jo AH, Kwak KC (2023) Speech Emotion Recognition Based on Two-Stream Deep Learning Model Using Korean Audio Information. *Appl. Sci* 13:2167
- (36) Aleisa HN, Alrowais FM, Negm N et al (2023)Henry Gas Solubility Optimization With Deep Learning Based Facial Emotion Recognition for Human Computer Interface. *IEEE Access* 11:62233-62241.
- (37) Gagliardi G, Alfeo AL, Catrambone V et al (2023) Improving Emotion Recognition Systems by Exploiting the Spatial Information of EEG Sensors. *IEEE Access*,pp 39544-39554

- (38) Eswaran KCA, Srivastava AP, Gayathri M (2023) Hand Gesture Recognition for Human-Computer Interaction Using Computer Vision. Communications in Computer and Information Science, Springer
- (39) Ansar H, Mudawi NA, Alotaibi SS et al (2023) Hand Gesture Recognition for Characters Understanding Using Convex Hull Landmarks and Geometric Features. IEEE Access 11:82065-82078
- (40) Kothadiya DR, Bhatt CM, Rehman A et al (2023) SignExplainer: An Explainable AI-Enabled Framework for Sign Language Recognition With Ensemble Learning. IEEE Access, pp 47410-47419
- (41) Sartaj AS, Ali Z, Hiroki T (2023) Cascaded deep graphical convolutional neural network for 2D hand pose estimation. International Workshop on Advanced Imaging Technology (IWAIT) ,1259215
- (42) Yue L, Pengcheng A, Yage X (2023) Eggly: Designing Mobile Augmented Reality Neurofeedback Training Games for Children with Autism Spectrum Disorder. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol
- (43) Maarten VM, Rachel CS, Marie-Monique S et al (2023) Marianne Graves Petersen, and Ole Sejer Iversen. Emerging Technologies in K–12 Education: A Future HCI Research Agenda. ACM Trans
- (44) Ometto M (2023) An innovative approach to plant and process supervision. Danieli Intelligent Plant
- (45) Matheus N, Jorge J, Vissoci J et al (2023) Exploring affordances for AR in laparoscopy. Ithaca: Cornell University Library, arXiv.org
- (46) Rawat KS, Sood SK (2021) Knowledge mapping of computer applications in education using CiteSpace. Comput Appl Eng Educ 29: 1324– 1339

5 It would be interesting if the authors proposed a new categorization of SOTA articles according to their characteristics.

**[Reply ]** Thank you for your great suggestion. In future research, we will try to use SOTA model and method to conduct more efficient research.

6 The quality of the Figures needs to be improved

**[Reply ]** Thanks for your suggestion. We have replaced the Figures with high pixel and uploaded them as an attachment

7 There needs to be an explicit research objective(s) and/or research question(s) stated, preferably as a separate section. This helps readers find out what the research is trying to address.

**[Reply]** Thank you for your good suggestion. We added research objectives and research questions in Chapter 1.

**The added/revised content:**

research objectives and questions : Recent advancements in Human-Computer Intelligent Interaction (HCII) have yielded a plethora of applications in diverse domains, including emotion and gesture recognition methods [33-41], as well as the integration of Augmented Reality (AR) in healthcare, education, and industrial settings [42-45]. These technologies will enable users to have more natural and immersive interactions with computers. Users will be able to operate computers or other digital devices using natural gestures, and they can also connect human thoughts to computers through brain-computer interfaces, facilitating a more direct and efficient mode of interaction. These technologies will significantly make our lives more convenient and intelligent. Despite these strides, a conflation often exists between HCII and its predecessor, Human-Computer Interaction (HCI), resulting in a vague understanding of their distinct characteristics among the general public and even within the academic community.

To bridge this gap and shed light on the evolutionary trajectory and future directions of HCII, this study employs CiteSpace [46] to conduct a comprehensive analysis spanning literature of two decades (2000-2023). The review and analysis are structured around the following three aspects: ① Understanding the relationships and differences among HCII, AI, and HCI, particularly focusing on how AI-driven methodologies have transformed HCII from traditional HCI paradigms; ② Mapping the geographical and institutional contributions to HCII, providing a nuanced understanding of research interests and academic engagements across nations and academic communities; ③ Predicting potential important areas for HCII applications and highlight emerging avenues for innovation. By addressing these aspects, this study aims to provide a rigorous academic framework that can guide and inform future efforts in HCII research and applications.

8 The reference list needs tidying up, as there are references missing items or formatting issues. Please be consistent with the formatting and use some standard formatting style.

**[Reply ]** Thank you for your good suggestion. We have checked the format and citation position of all references, and unified the reference format as standard (author+year+literature name+journal).

Thanks again for all your insightful comments and valuable suggestions.

# Current Status and Trends of Technology, Methods, and Applications of Human-computer Intelligent Interaction (HCII): A Bibliometric Research

Zijie Ding<sup>1</sup>, Yingrui Ji<sup>2,3</sup>, Yan Gan<sup>1\*</sup>, Yuwen Wang<sup>1</sup>, Yukun Xia<sup>1</sup>

<sup>1</sup> School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan, 430074, China

<sup>2</sup> Aerospace Information Research Institute, Chinese Academy of Science, Beijing, 100094, China

<sup>3</sup> School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China

## Abstract

This study delves into Human-Computer Intelligent Interaction (HCII), a burgeoning interdisciplinary field that builds upon traditional Human-Computer Interaction (HCI) by integrating advanced technologies like Natural Language Processing (NLP) and Machine Learning (ML). In this paper, we scrutinize 5,781 HCII papers published between 2000 and 2023, narrowing our focus to 803 most relevant articles to construct co-citation and interdisciplinary networks based on the CiteSpace Software. Our findings reveal that the publications of the United States and China are relatively high with 558 and 616 publications respectively. Furthermore, we found that machine learning and deep learning have emerged as the prevalent methodologies in HCII, which currently emphasizes multimodal emotion recognition, facial expression recognition, and NLP. We predict that HCII will be integrated into advanced applications such as neural-based interactive games and multi-sensory environments. In sum, our analysis underscores HCII's role in advancing artificial intelligence, facilitating more intuitive and efficient human-computer interactions, and its prospective societal impact. We hope that our review and analysis may guide the efforts of researchers aiming to contribute to HCII and develop more powerful and intelligent methods, tools, and applications.

**Keywords:** Human-Computer Intelligent Interaction, Human-Computer Interaction, Artificial intelligence, Bibliometric Research, CiteSpace

## 1. Introduction

With the rise of artificial intelligence (AI), the field of Human-Computer Interaction (HCI) has undergone significant transformations, leading to more intuitive and efficient human-computer collaborations. Notably, Tesla's Optimus Prime humanoid robot enhances human productivity and operational efficiency, while Apple's Apple Vision Pro mixed-reality headset provides a multisensory, immersive user experience.[1-2]. The advancements in Natural Language Processing (NLP) and Computer Vision are instrumental in these developments, presenting both opportunities and challenges in HCI [3-5]. At the same time, a new paradigm known as Human-Computer Intelligent Interaction (HCII) has arisen, driven by the combined efforts of AI and HCI

---

\* Corresponding author.

E-mail address: ganyan@hust.edu.cn (Y.Gan).



1 research.[6].Unlike traditional HCI,HCII focuses on the naturalization and humanization of  
2 interactions, particularly in complex and specialized application domains. The application of HCII,  
3 augmented by AI-centric methodologies, has demonstrated innovative solutions and feasibility  
4 across various sectors. These include healthcare [7-12],smart home systems [13-14],transportation  
5 [15-21],intelligent manufacturing [22-25],educational frameworks [26-27],and interactive gaming  
6 [28-30].The evolution in HCI and the progress of HCII have captured the attention of  
7 interdisciplinary researchers, revealing the critical role of AI in shaping the future landscape of  
8 human-computer interactions. Recent advancements in Human-Computer Intelligent Interaction  
9 (HCII) have yielded a plethora of applications in diverse domains, including emotion and gesture  
10 recognition methods [33-41],as well as the integration of Augmented Reality (AR) in healthcare,  
11 education, and industrial settings [42-45].These technologies will enable users to have more natural  
12 and immersive interactions with computers. Users will be able to operate computers or other  
13 digital devices using natural gestures, and they can also connect human thoughts to computers  
14 through brain-computer interfaces, facilitating a more direct and efficient mode of interaction.  
15 These technologies will significantly make our lives more convenient and intelligent. Despite these  
16 strides, a conflation often exists between HCII and its predecessor, Human-Computer Interaction  
17 (HCI),resulting in a vague understanding of their distinct characteristics among the general public  
18 and even within the academic community.

19 To bridge this gap and shed light on the evolutionary trajectory and future directions of HCII,  
20 this study employs CiteSpace [46] to conduct a comprehensive analysis spanning literature of two  
21 decades (2000-2023).The review and analysis are structured around the following three aspects: (1)  
22 Understanding the relationships and differences among HCII, AI,and HCI, particularly focusing on  
23 how AI-driven methodologies have transformed HCII from traditional HCI paradigms; (2) Mapping  
24 the geographical and institutional contributions to HCII, providing a nuanced understanding of  
25 research interests and academic engagements across nations and academic communities; (3)  
26 Predicting potential important areas for HCII applications and highlight emerging avenues for  
27 innovation. By addressing these aspects, this study aims to provide a rigorous academic framework  
28 that can guide and inform future efforts in HCII research and applications.

29 This study represents a novel endeavor in the Human-Computer Intelligent Interaction (HCII)  
30 research landscape, leveraging CiteSpace software for the analysis and visualization of literature.  
31 This innovative methodology enables an unprecedented, nuanced understanding of HCII's evolving  
32 research themes, hotspots, and international scholarly collaborations. The analysis addresses  
33 existing challenges of correlation and clustering in conventional HCII research, thereby offering a  
34 robust analytical lens to capture the dynamism and trends shaping this rapidly growing field. The  
35 contributions of this study are three-fold: (1) This paper rigorously defines HCII, elucidates its  
36 symbiotic relationship with Artificial Intelligence (AI) technologies, and outlines the architectural  
37 paradigms and technical methodologies underpinning human-computer intelligent interactions; (2)  
38 Utilizing the CiteSpace Software, this paper visualizes and scrutinizes HCII-related literature  
39 spanning two decades.The analysis yields critical insights into emergent research hotspots, thematic  
40 clusters, and longitudinal trends in the HCII domain; (3) This paper prognosticates future  
41 trajectories for HCII, identifying prospective intelligent interaction paradigms and methodologies,  
42 thereby informing researchers with strategic directions and conceptual scaffolding for future  
43 research efforts. In summary, this study aims to act as a foundational reference, providing both a  
44 thorough overview of current HCII research and a forward-thinking viewpoint to steer future  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65



academic and practical efforts.

The paper proceeds as follows. Section 2 provides a critical review of existing literature on HCII, Artificial Intelligence (AI) algorithms, and the modifications and applications that have evolved in HCII. Section 3 elaborates on the methodology employed in the study, focusing on the data screening processes, database creation, and the utilization of CiteSpace as the analytical tool. Section 4 presents the study's results, featuring analyses of literature growth trends, collaborative networks, co-citation structures, and disciplinary intersections within HCII research. Section 5 discusses the empirical findings, drawing connections to existing literature and theorizing on their implications for the field of HCII. Finally, Section 6 concludes the paper, offering recommendations for future research directions in HCII.

## 2. Background and related work

### 2.1 Intelligent human-computer interaction

Human-Computer Intelligent Interaction (HCII) represents an interdisciplinary confluence of Artificial Intelligence (AI) and Human-Computer Interaction (HCI) technologies [4]. At its core, HCII aims to enhance computational system performance through the integration of AI methodologies [47]. The overarching objective shared by both HCII and HCI is the optimization of the user experience [48]. AI technology encompasses a suite of algorithms and computational models designed to emulate human cognition and decision-making processes [49]. Prominent among these are machine learning (ML), natural language processing (NLP), and computer vision (CV). Conversely, HCI focuses on the design principles and implementation strategies that facilitate effective and efficient interaction between humans and computational systems [50]. These interactions may employ a myriad of modalities, including but not limited to, visual cues, auditory signals, haptic feedback, and gestural inputs [51,52]. The symbiotic relationship between AI and HCI gives rise to the HCII domain, as illustrated in Figure 1. Specifically, AI-driven techniques such as ML and NLP can be harnessed to refine HCI elements, thereby fostering smarter user interfaces, task automation, and personalized interaction experiences [53]. This synergy further catalyzes the development and application of avant-garde technologies, including autonomous vehicular systems.

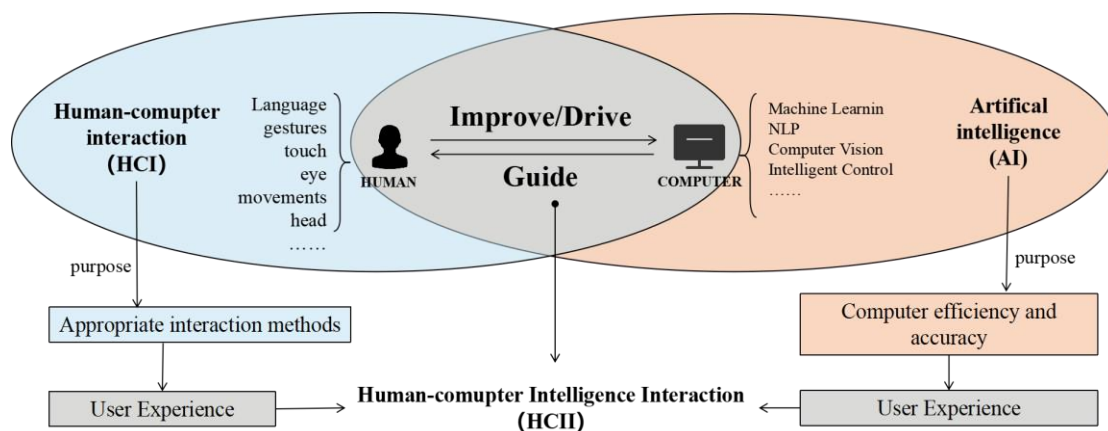


Fig.1 The relationship between AI and HCI in human-computer intelligence interaction

Intelligent User Interfaces (IUI) represent a specialized facet of Human-Computer Intelligent Interaction (HCII). These interfaces serve as platforms not just for information exchange but also as

1 adaptive, user-centric operational environments. Through advancements in smart HCI technologies,  
2 IUI has given rise to multimodal forms, revolutionizing the realm of user interfaces [54]. These  
3 environments dynamically adjust layout, content, and interaction pathways based on user needs and  
4 behaviors, thus optimizing efficiency, efficacy, and naturalness [55,56]. Moreover, IUIs employ  
5 sophisticated modeling techniques, encompassing user, domain, task, conversation, and media  
6 models, to facilitate nuanced forms of interaction [58-59].  
7

8 Contrastingly, Human-Computer Dialogue Systems extend the scope of intelligent interactions  
9 by focusing on natural language-based communication [60]. Unlike HCII, which primarily aims for  
10 efficient and effective interactions, these systems deploy Natural Language Processing (NLP)  
11 technologies to simulate natural conversational dynamics. These dialogue systems find extensive  
12 application in sectors like voice-activated assistants [11], intelligent customer service [61], and  
13 chatbots [27].  
14

15 Human-computer dialogue systems deploy a multi-tiered architectural framework to manage  
16 the complexity of information sensing, processing, decision-making, and control, ultimately  
17 culminating in application-specific functionalities [62].  
18 **Sensor Layer:** Constituting the foundational tier, the sensor layer captures a variety of human inputs, including vocal cues, visual data, and  
19 gestural signals, through specialized hardware such as microphones, cameras, and other sensors.  
20 **Signal Processing Layer:** As the subsequent layer, this module is tasked with the computational  
21 analysis of data amassed by the sensor layer. It employs algorithms for speech, image, and gesture  
22 recognition, serving as the bridge between raw data and actionable insights.  
23 **Intelligent Decision-Making Layer:** Leveraging the processed signals, this layer utilizes a range of machine learning,  
24 deep learning, and natural language processing algorithms to facilitate autonomous decision-making  
25 and control mechanisms.  
26 **Interactive Interface Layer:** Situated above the decision-making layer, this tier is responsible for presenting the processed information to the user. It accommodates various  
27 forms of user feedback through multiple interfaces, including graphical, vocal, and gestural  
28 pathways.  
29 **Application Layer:** Representing the apex of this architectural model, the application  
30 layer executes specific functionalities based on synthesized information and user interactions,  
31 thereby realizing the system's ultimate objectives.  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

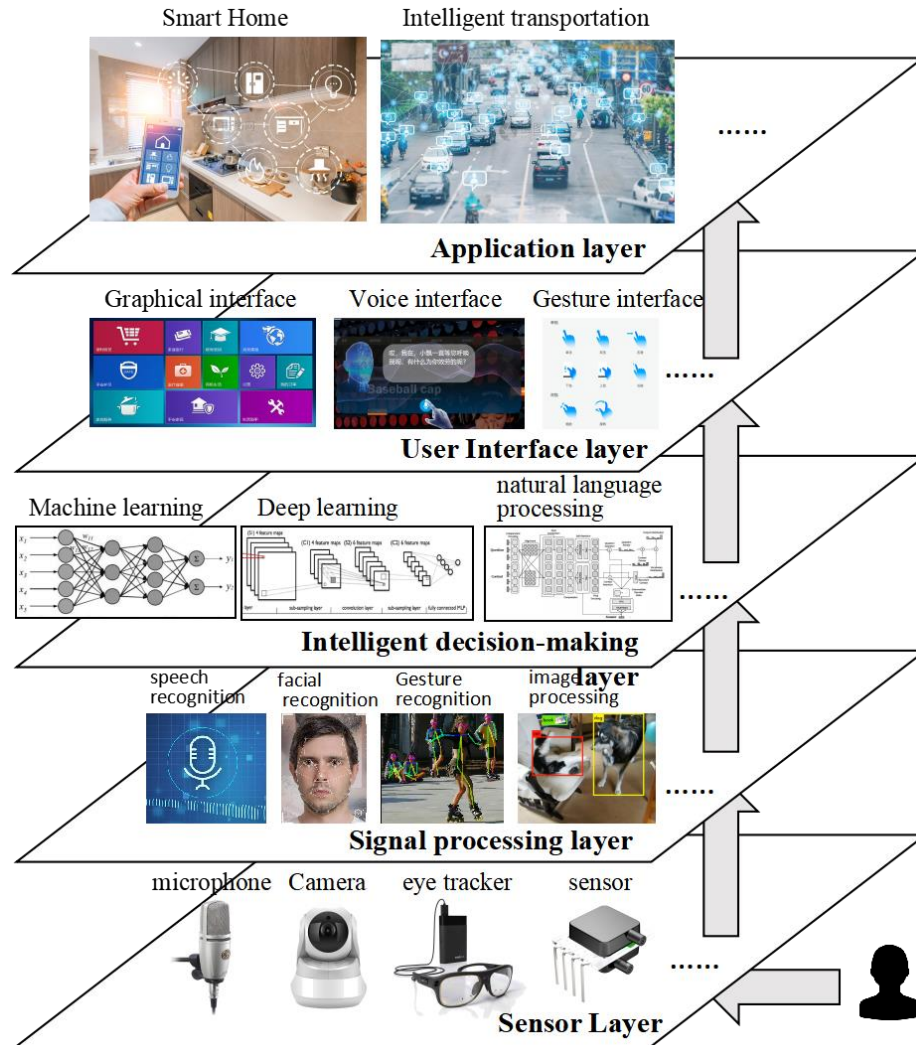


Fig.2 HCI's five levels of application architecture

## 2.2 AI Algorithms in Human-Computer Intelligent Interaction

Machine Learning Algorithms (ML) [63]: Encompassing supervised, unsupervised, and semi-supervised learning paradigms, these algorithms find applications in tasks such as classification, clustering, and predictive modeling.

Deep Learning Algorithms (DL) [64]: Incorporating architectures like convolutional neural networks (CNNs), recurrent neural networks (RNNs), and generative adversarial networks (GANs), these algorithms are instrumental in image recognition and natural language processing tasks.

Decision Tree Algorithms: These algorithms facilitate classification and predictive analytics by recursively partitioning feature spaces and constructing decision trees.

Support Vector Machines (SVM): Employed for classification and regression tasks, SVM algorithms map data into high-dimensional spaces to identify optimal hyperplanes for classification.

Clustering Algorithms: Algorithms such as K-means and hierarchical clustering serve to uncover the intrinsic structures within datasets by grouping them into clusters.

Reinforcement Learning Algorithms: These algorithms enable agents to learn optimal

1 decision-making strategies through a system of rewards and penalties.

## 2 3 4 **2.3 Evolution and Diversification of HCII: From HCI to Multimodal Interactions**

5  
6 Traditional HCI primarily emphasized visual interactions via interface operating systems to meet  
7 user requirements [70].The subsequent advent of voice interaction technology marked a pivotal  
8 junction, fostering collaboration between HCI and AI disciplines [71].As the field evolved into HCII,  
9 researchers have increasingly harnessed AI techniques to explore a plethora of interaction  
10 modalities. Firstly, HCII accentuates natural user interfaces, facilitating interactions through speech,  
11 gestures, and facial expressions rather than conventional devices like mice and keyboards [65-  
12 66].This naturalistic interface design aligns more closely with human behavior, thereby enriching  
13 user experience and interaction efficiency.

14  
15 Secondly, HCII leverages AI-driven technologies capable of intelligent recognition and  
16 physiological signal analysis [51-52],thereby achieving nuanced understanding of user intentions  
17 and requirements. This level of understanding enables the provision of highly personalized and  
18 intelligent services.

19  
20 Moreover, HCII integrates multimodal interactions [67,58],such as simultaneous voice and  
21 gesture recognition, amplifying the flexibility and efficacy of the interactive experience.

22  
23 Lastly, emotional interaction capabilities have been incorporated [68-69],allowing the system  
24 to recognize and respond to users' emotional states through vocal and facial cues, thereby delivering  
25 a more empathetic user experience. **The outcomes have been transformative, diversifying HCII  
26 implementations into the following categories: Voice Interaction: Facilitates verbal communication  
27 between users and systems; Visual Interaction: Involves graphical or image-based interfaces;  
28 Gesture Interaction: Utilizes human gestures for non-verbal communication; Multimodal  
29 Interaction: Incorporates multiple interaction channels simultaneously, such as voice and gesture  
30 [67,58]; Emotional Interaction: Employs sentiment analysis algorithms to interpret user emotions  
31 for more empathetic user experiences [68-69].**

32  
33 In terms of cross-domain applications, HCII technologies are gaining traction across a wide  
34 range of sectors, including healthcare, smart homes, transportation, intelligent manufacturing,  
35 education, and interactive gaming. In these domains, HCII is principally employed for control,  
36 monitoring, administration, and service delivery. Core methodologies include: Recognition  
37 Technologies: Used for identifying voice, image, and gesture inputs; Machine Learning and Natural  
38 Language Processing: Employed for decision-making and language interpretation; Data Analytics:  
39 Utilized for extracting actionable insights from large datasets; Virtual Reality Technologies:  
40 Employed for immersive user experiences; Sensor Technologies: Used for capturing a wide range  
41 of physical data; User Interface Design: Focuses on creating intuitive and user-friendly interfaces.

42  
43 (1)Voice Interaction Originating from spoken dialogue systems, voice interaction leverages  
44 technologies such as speech recognition, natural language comprehension, dialogue  
45 management,and speech synthesis [72].The ubiquity and convenience of speech have propelled  
46 voice interaction into mainstream application technologies. Voice-controlled interfaces have found  
47 significant applications in vehicle control [13,17],autonomous driving [73],and domestic settings  
48 [11],enhancing naturalness and precision in user interfaces [74].Emotional cues in voice serve as  
49 invaluable data for emotion recognition algorithms [75-76].

50  
51 (2) Visual Interaction: Visual interaction is integrally related to Intelligent User Interfaces  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

(IUI),relying on cameras, sensors, and other hardware to facilitate human-machine visual communication. Innovations like virtual reality [29-31],augmented reality [30,77],and digital twin technology [78-81] have expanded the scope of visual interaction. Its application areas include medical imaging [41],robot control [20-21], and vehicle safety [14-16].

(3) Gesture Interaction: Gesture interaction encompasses a variety of non-verbal cues such as movements and signs [51].Due to the inherent complexity in gesture recognition [82],AI algorithms coupled with machine learning techniques are employed to enhance recognition accuracy. Gesture-based interactions are particularly relevant in robotics, somatosensory games, virtual reality, and sign language recognition [68,83].Research in this domain often focuses on pose estimation [84],motion capture [85-86],and gesture classification.

(4) Multimodal and Emotional Interaction: Multimodal interaction fuses multiple interaction modalities to offer a more natural and efficacious user experience. Emotional interactions often leverage multimodal recognition algorithms, utilizing facial and speech recognition to model emotional states [69,87,88-90].Extensive research is underway in innovative applications such as personality-based financial advisory systems [91],personalized smart clothing [92],interactive cultural heritage communities [93],and Alzheimer's disease recognition through multimodal fusion [94].

## 2.4 Related work

Similar studies from the past ten years to this paper's topic are [4][95-96]. First, umak,B.et al.[95] identified and analyzed the research on HCII and Intelligent User Interface (IUI) between 2010 and 2021,concluding that deep learning and instance-based AI methods and algorithms are the most commonly used by researchers. Support Vector Machines (SVMs) are the most popular algorithms for recognising emotions, facial expressions, and gestures, whereas Convolutional Neural Networks (CNNs) are the most popular deep learning algorithms for recognition techniques. Quantitative studies have mapped the extant sensors, techniques, and algorithms used in HCII, but there are fewer quantitative studies and forecasts of future trends, which have not yet been scientifically summarised. The combination of CNN and Long Short-Term Memory Network (LSTM) can significantly enhance the accuracy and precision of action recognition. With the proliferation of interactive devices such as AR/VR, they predict that diverse sensors will be combined to store and control more complicated processes. They imply that smarter human-computer interaction may be used in some public areas, and that virtual reality glasses will be used in entertainment, gaming, and industrial manufacturing, but they have not yet specified existing application scenarios with corresponding technical approaches. Karpov,A.A.,et al.[96] conducted a study on the future of cognitive interfaces and the most recent HCI technologies. Similar to [95],they simultaneously analyzed and reviewed HCII and IUI; however, the former focuses more on investigating the technology of human-computer interaction, whereas the latter focuses on the user interface. Examining the progression of user interfaces from command text to graphics to intelligent unimodal and multimodal user interfaces. It also analyses the development of multimodal interfaces between countries and derives some conclusions about HCI systems for people with disabilities and the trend towards "human-to-human" interaction.

In addition, there are a number of related and comparable review articles, such as the systematic review of HCII and AI in healthcare [97],the research progress between human-computer interaction and intelligent robots [98],and human-computer interaction and evaluation in AI conversational

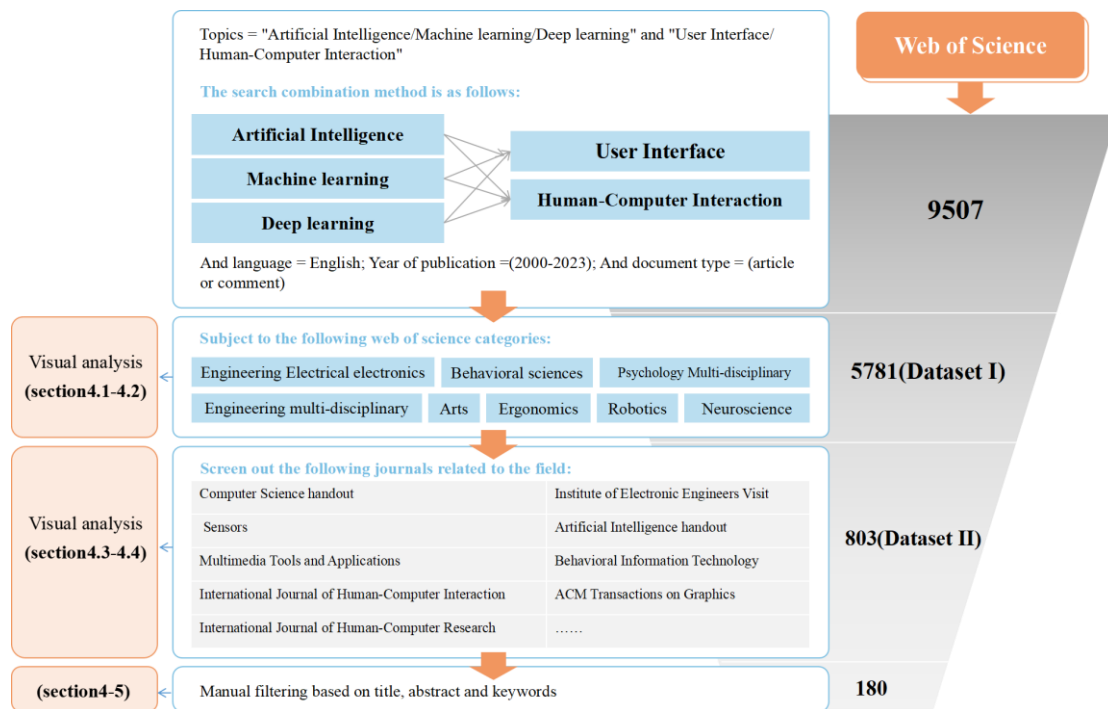
agents [99-100]. There is no research that summarises the overall development of human-computer intelligent interaction and related application areas and popular trends. This paper seeks to comprehend the existing hot technologies and application areas in the field of human-computer intelligent interaction, to create a series of hot knowledge maps using bibliometric techniques, and to investigate the emerging methods and technologies of human-computer intelligent interaction as well as the future application scenarios.

### 3. Methodology

#### 3.1 Database Selection and Search Strategy

The Web of Science (WoS) Core Collection was elected as the primary database for this research, owing to its comprehensive and reliable citation indexing for scholarly literature across scientific and technical domains [63]. This database was favored for its high-quality content and its capability for detailed categorization of retrieved data, encompassing attributes such as title, author, institution, and publication type. The search query was formulated to capture articles pertinent to "Human-Computer Intelligent Interaction." Given the overlapping thematic scopes of "Artificial Intelligence" and "Human-Computer Interaction," these terms were included in the query. Further specificity was afforded by incorporating "Machine Learning" and "Deep Learning" into the search criteria. The final query string was thus constructed as follows: Initial search outcomes were refined by applying category-based filters, such as "Computer Science," "Electrical and Electronic Engineering," "Human Engineering," "Psychology," "Multidisciplinary Engineering," among others. This step was implemented to exclude non-relevant literature in disparate fields like chemistry. This exercise resulted in an initial pool of 5,781 articles. Subsequently, articles appearing in journals not directly aligned with the thematic focus of HCII were excluded. Examples of such journals include but are not limited to "Lecture Notes in Computer Science," "Institute of Electrical Engineers Access," and "Sensors," yielding a final corpus of 803 papers. A meticulous manual screening process was conducted, wherein articles were evaluated based on their title, abstract, and keywords. The screening was narrowed to cover HCII-specific tasks such as data mining, signal processing, feature recognition, and recommendation models, within applicative domains like healthcare, smart homes, transportation, and education. A comprehensive analysis was subsequently carried out on a curated subset of 180 scholarly articles, aimed at encapsulating the existing paradigms in user-interface interactions and future directions in the field.





**Fig.3 Literature Screening and Utilization Process**

### 3.2 Bibliometric Analysis and Knowledge Mapping

This study adopts a bibliometric approach combined with knowledge mapping to elucidate the current state, emergent trends, and future trajectories in the domain of Human-Computer Intelligent Interaction (HCII). Knowledge mapping is an interdisciplinary methodology designed to visualize, analyze, and interpret patterns in scientific literature, thereby contributing to the broader field of scientometrics [101,102]. We employ CiteSpace, a software application designed for the interactive and exploratory analysis of scientific fields. It is particularly useful for generating visual co-citation networks and conducting collaborative analyses across authors, institutions, and nations [105]. The software also facilitates co-occurrence analyses of terms, keywords, and subject categories, as well as co-citation analyses of referenced literature, authors, and journals [46]. CiteSpace offers Betweenness Centrality (BC) as a metric for identifying pivot nodes in a scientific network. The BC value is calculated based on the shortest paths connecting the node of interest to all other nodes in the network [106]. The time-zone graph feature in CiteSpace provides a temporal perspective on the dataset, enabling the identification of citation bursts and emergent research trends [106]. This tool has been widely applied in research fields such as sustainability and urban studies [107-109]. To achieve the research objectives, several parameters in CiteSpace — including time slices, node types, links, selection criteria, and pruning methods—were carefully configured during the analytical process [110]. The study uses CiteSpace to create various types of visual networks, including collaborative networks among countries, authors, and institutions; co-citation networks that highlight clusters of cited literature; and disciplinary category symbiosis networks that reveal the interconnections among key knowledge domains.

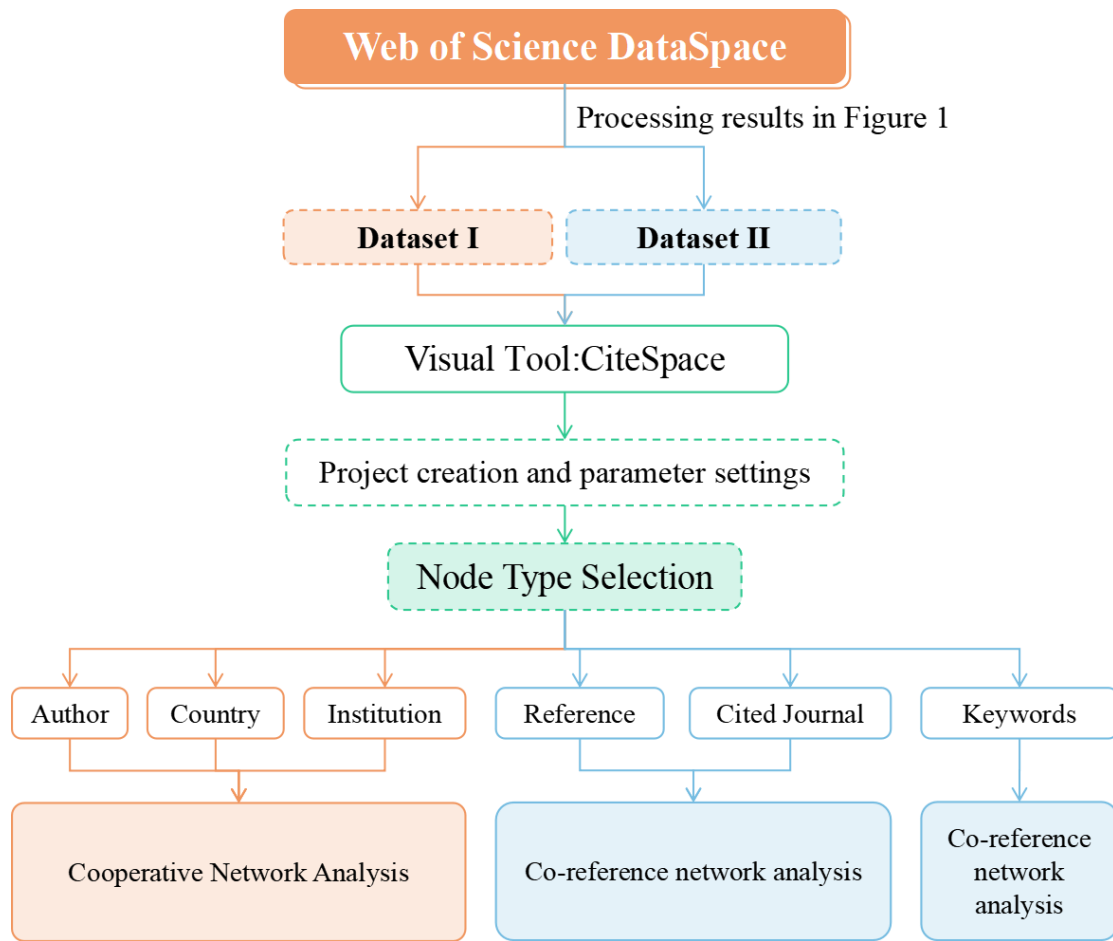


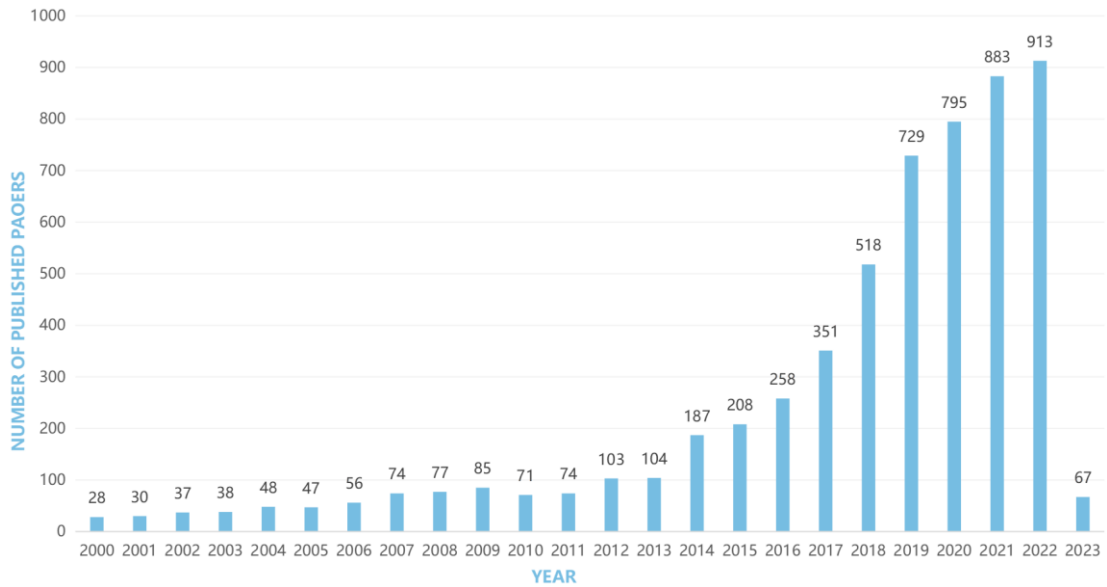
Fig.4 CiteSpace visualization and analysis process

## 4. Results

### 4.1 Trends in literature growth

Figure 5 offers a comprehensive statistical analysis based on the initial screening of 5,781 articles. The scholarly output in the Human-Computer Intelligent Interaction (HCII) domain appears to have evolved through three distinct developmental phases: an exploratory phase from 2000 to 2013, an initial development phase from 2014 to 2016, and a rapid growth phase from 2017 to 2022. During the exploratory phase, a total of 872 papers were published, accounting for approximately 15% of the total literature, indicating the field was in its nascent stage. The initial development phase was relatively short-lived, signaling a swift uptick in research interest. The rapid growth phase, stretching from 2017 to 2022, saw a substantial increase in applied research related to HCII. The average annual growth rate of the literature was around 100 articles during this period. Specifically, the years 2020, 2021, and 2022 witnessed the publication of approximately 800, 883, and 913 papers, respectively. Publications in this stage make up 72.4% of the entire dataset. Given the current trajectory, it is hypothesized that scholarly contributions to the HCII field will likely continue to proliferate at an accelerated rate.





**Fig.5** Research and Development Trends in HCII Domain Data

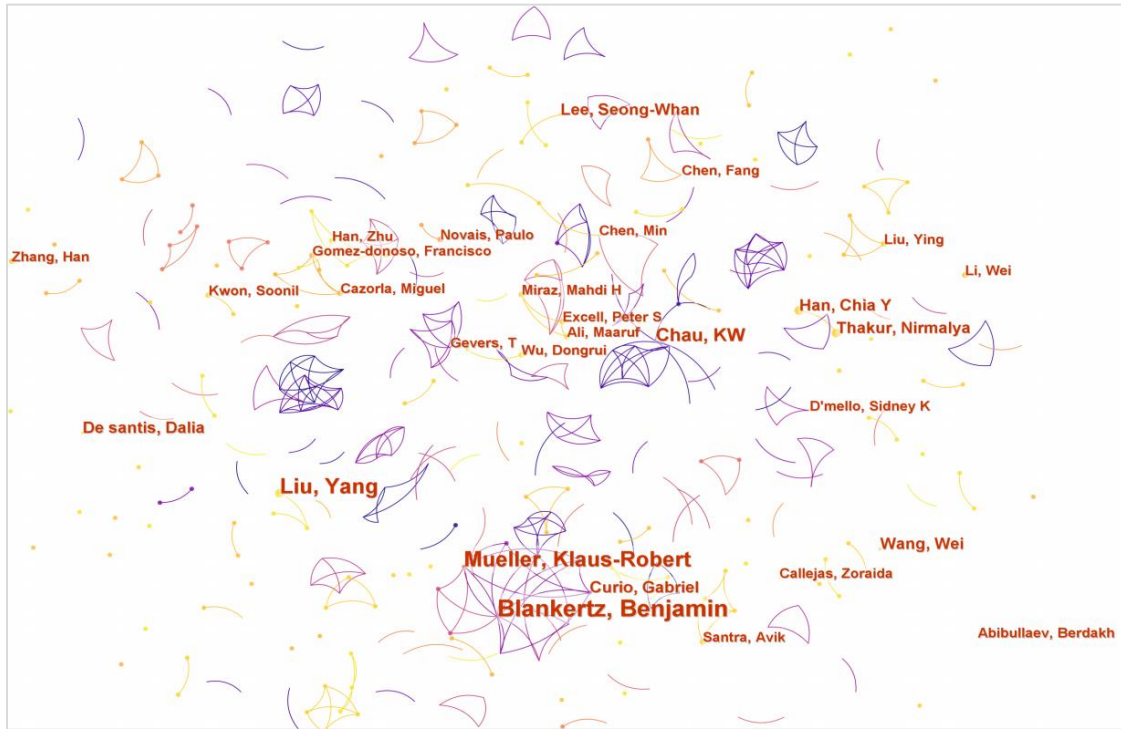
## 4.2 Analysis of cooperation networks

A network analysis of author collaborations was undertaken, based on the initial screening of 5,781 articles, to identify patterns of cooperation among authors in the HCII domain.

### 4.2.1 Study author association analysis

Figure 6 presents a visual network graph generated using CiteSpace, where each node symbolizes an individual author. The graph reveals that 713 authors have engaged in HCII-related research, resulting in 410 collaborative instances. The density of this collaboration network is measured at 0.0016, suggesting that scholarly cooperation in this field is relatively limited. Most contributions appear to be independent efforts, without the formation of extensive, collaborative research teams.

Table 1 lists the top 10 prolific authors, ranked by their publication count. The most active author in this realm has contributed to nine relevant papers, and all authors in the top 10 have published at least four articles. This analysis underscores the current state of collaboration within the HCII field, indicating that although the field is growing, it has yet to develop extensive networks of academic cooperation.



**Fig.6** Network diagram of research authors in the HCII domain of the WoS core database

**Table 1** Top 10 research authors in the HCII domain of the WoS core database

Serial No.	Author Name	Year of Debut	Number of Articles
1	Blankertz Benjamin	2006	9
2	Liu, Yang	2021	7
3	Mueller, Klaus-Robert	2007	7
4	Chau, KW	2001	5
5	Thakur, Nirmalya	2021	4
6	Wang, Wei	2006	4
7	Han, Chia Y	2020	4
8	Curio, Gabriel	2020	4
9	De santis, Dalia	2020	4
10	Lee, Seong-Whan	2019	4

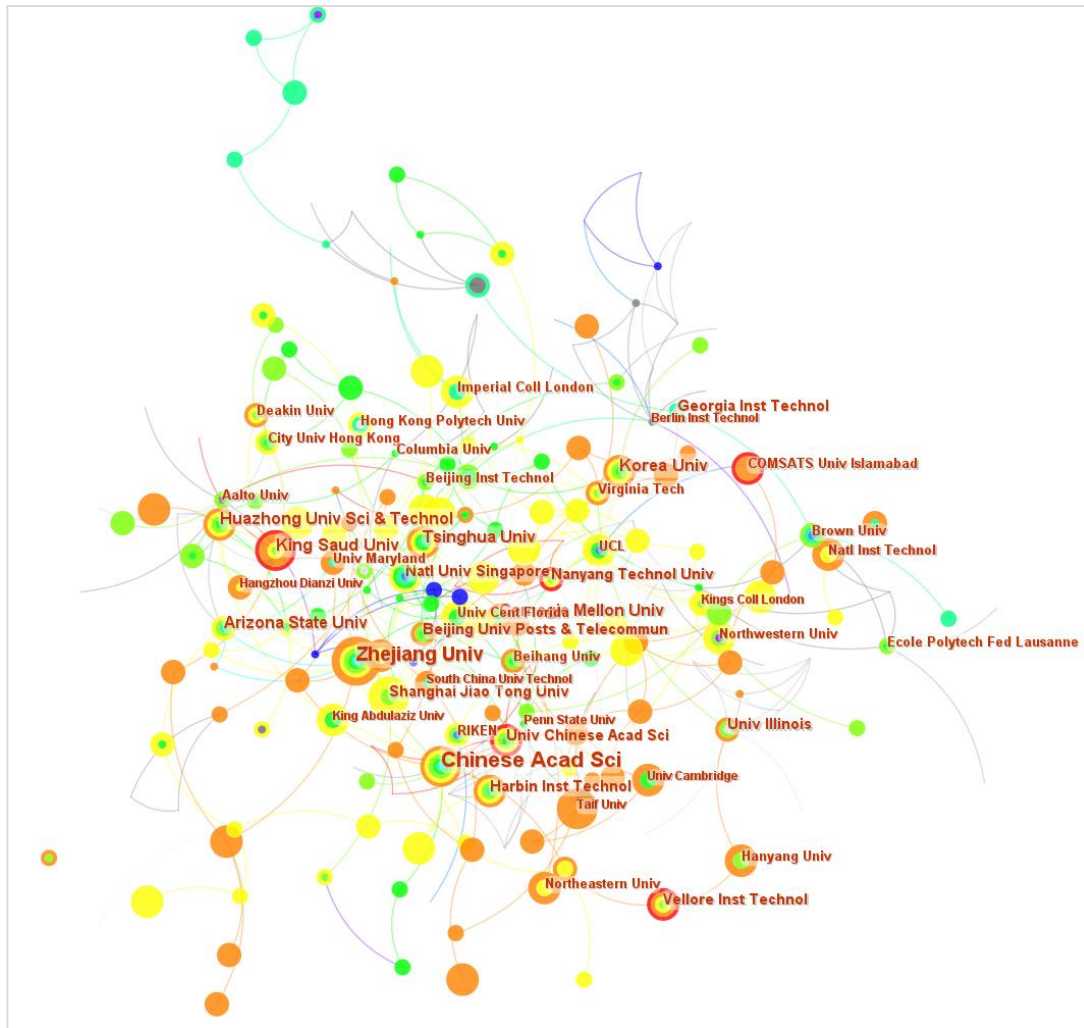
#### 4.2.2 Collaborative linkage analysis of research institutions

In this section, the focus shifts to the network of collaborations among research institutions engaged in HCII studies. Figure 7 provides a visual depiction of this network, where each node represents an institution. The size of each node is proportional to the institution's publication count, while the links between nodes signify collaborative endeavors. An analysis reveals that 602 institutions have participated in HCII research, resulting in 466 instances of inter-institutional collaboration. The density of this network is calculated to be 0.0026, which is notably higher than the author collaboration network, indicating tighter institutional linkages.

Table 2 lists the top 10 research institutions based on their publication output. Among them, 80 key research institutions have published at least five articles in the HCII domain. The Chinese

1 Academy of Sciences leads the pack with 36 publications over the past two decades. It is followed  
 2 by Korea University, Zhejiang University, King Saud University, Tsinghua University, Carnegie  
 3 Mellon University, Huazhong University of Science and Technology, Arizona State  
 4 University, Vellore Institute of Technology, and Shanghai Jiao Tong University, each contributing  
 5 more than 10 articles.  
 6

7 The data suggests that while a multitude of institutions are active in HCII research, their efforts  
 8 are relatively dispersed. Despite the high level of interest and attention from these institutions, there  
 9 is room for enhancing collaboration and synergies. Currently, there are few institutions that have  
 10 emerged as definitive leaders in the field.  
 11



12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47 **Fig.7** Collaborative network of research institutions in the HCII area of the WoS core database.

48 **Table 2** Top 10 cooperating organizations in the HCII area of the WoS core database

Serial No.	Name of Institution	Year of Debut	Number of Articles
1	Chinese Acad Sci	2005	36
2	Korea Univ	2017	21
3	Zhejiang Univ	2005	17
4	King Saud Univ	2020	16
5	Tsinghua Univ	2018	15

6	Carnegie Mellon Univ	2003	14
7	Huazhong Univ Sci & Technol	2006	13
8	Arizona State Univ	2008	13
9	Vellore Inst Technol	2020	12
10	Shanghai Jiao Tong Univ	2020	11

In the last two years, the Chinese Academy of Sciences (CAS) has concentrated its efforts on several key areas within the realm of Human-Computer Intelligent Interaction (HCII). These focus areas include: CAS has been exploring the fusion of various interaction modalities to create a more natural and effective user experience. The institute has been particularly active in this domain. Researchers have employed Convolutional Neural Networks (CNNs) to extract and classify implicit features from normalized facial images. Techniques such as facial cropping and rotation have been deployed to optimize CNN architectures, thereby enhancing the efficiency of facial expression recognition systems [111]. Tang Y et al. introduced a frequency neural network (FreNet) as a pioneering approach in frequency-based facial expression recognition, achieving superior performance [89,114]. CAS has also delved into the creation of human models, particularly in the context of gesture detection [113] and 3D spatio-temporal skeleton rendering [7]. Research at CAS has resulted in innovative medical applications based on deep learning models. These include: Skeleton-based human body recognition, facilitating 3D rendering of spatio-temporal skeletons [7]; Surgical robots controllable via eye movements [8-9]; Pulse cumulative image mapping techniques using CNNs to accurately measure real heart rates [10,11]. Promising results have been achieved using recurrent neural networks for the recognition and classification of EEG signals, contributing to the field of affective computing.

Overall, the extensive research portfolio of CAS manifests its leadership in the HCII domain, particularly in applying machine learning and deep learning techniques to tackle complex, real-world problems.

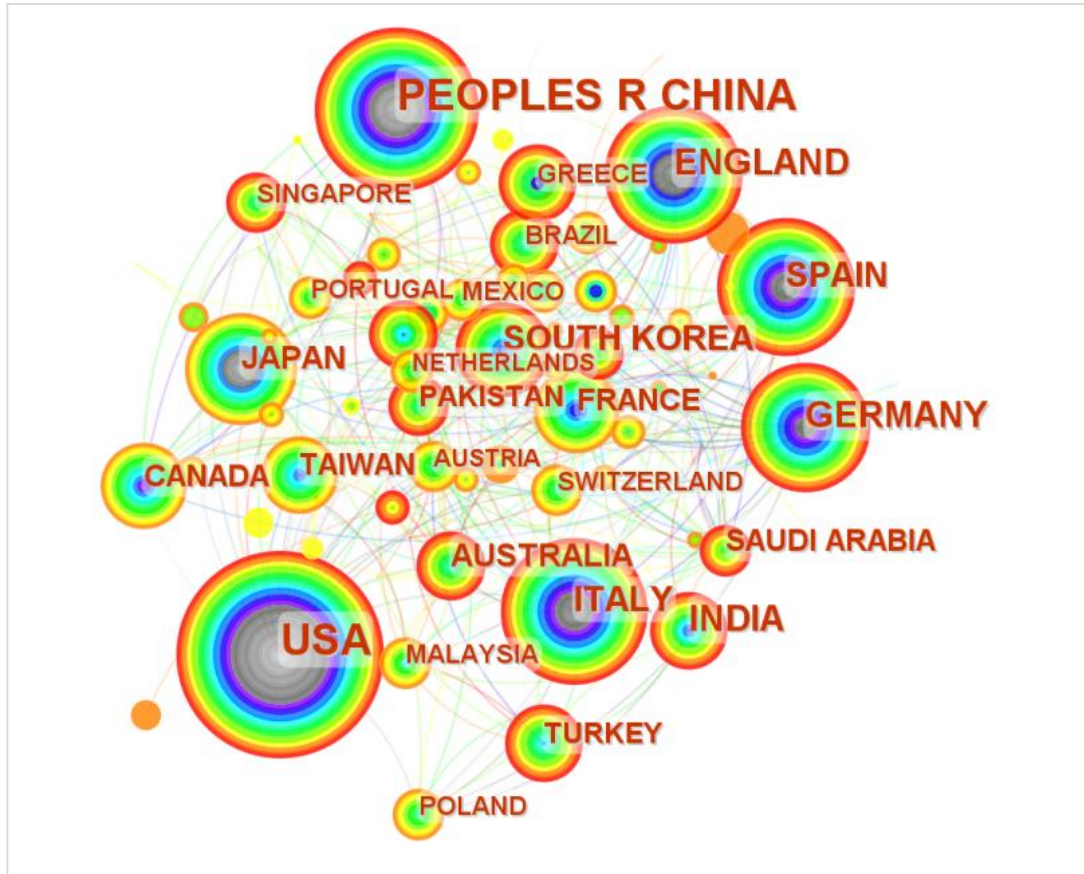
The second major contributor to the field of Human-Computer Intelligent Interaction (HCII) is a Korean university. Their research is largely application-driven, targeting the improvement of human-computer interaction and user interfaces in smart sensing devices and virtual systems. Below are some of the key focus areas: Virtual Reality Sports [116]: Development of immersive virtual reality experiences tailored for sports. Gaming Fitness [26]: Leveraging gaming elements to enhance physical fitness training. Interactive Sports Training Strategy [117]: Utilizing sensor data and machine learning algorithms to develop personalized and effective sports training strategies. AI Agent System for Multiple Vehicles [17]: Deployment of intelligent agents to manage fleets of vehicles. Real-time Bus Management System [18]: Utilizing data mining technologies for real-time management and optimization of bus services. Self-driving Car Takeover Performance [19]: Evaluating the efficiency and reliability of human takeover in self-driving cars through machine learning techniques. Facial Recognition [118]: Research into advanced facial recognition technologies for various medical applications. Gesture Recognition [119]: Active investigation into gesture-based control and monitoring systems for healthcare applications.

Overall, the research emanating from this Korean university demonstrates a strong orientation towards applied HCII. The primary focus is on the recognition, extraction, and classification of human behaviors. This research is not only innovative but also finds practical applications in critical sectors like healthcare and transportation. The work contributes to the broader goal of making

human-computer interaction more natural, efficient, and beneficial across various domains.

#### 4.2.3 Global Distribution of Research Contributions by Country

The geographic distribution of research contributions serves as an indicator of a country's significance and influence in a specific field. Moreover, international collaboration is instrumental in fostering innovation and making strides in research. In this analysis, countries are represented as nodes, and metrics such as the number of publications and betweenness centrality are computed. Figure 8 presents a symbiotic network diagram that outlines the global distribution of HCII research contributions. The network revealed that: A total of 101 countries have engaged in HCII research (N=101). These countries have collaborated 374 times (E=374). The collaboration density is calculated to be 0.741, signifying a high level of international collaboration. China, the United States, and France emerged as the most interconnected nations in this domain, each contributing 7 articles. This attests to a vibrant academic environment in HCII within these countries. Table 3 extracts the top 15 nations based on their publication counts: **China (616 publications) and the United States (558 publications) lead the pack with substantial contributions exceeding 500 publications each. Betweenness centrality values indicate that the U.S.(0.75) has a slightly higher influence compared to China (0.55) in the field. While Germany and France have fewer publications, their higher centrality scores indicate that their research contributions are of high quality and serve as pivotal nodes in the global HCII research network.** In summary, the HCII field enjoys a rich tapestry of international contributions, with China being the most prolific contributor both in terms of publications and international collaborations. However, there is room for more active academic cooperation and integration to advance the field. From a global standpoint, the research output from the United States exerts a greater influence, whereas China's contributions are more abundant. For the future, there is a pressing need to encourage more diversified and intelligent development in HCII research across the globe.



**Fig.8** Global distribution of research countries in the HCII domain of the WoS core database

**Table 3** Top 15 countries in the HCII area of the WoS core database

Serial No.	Country or Area	Year of Debut	Number of Articles	Centrality
1	PEOPLES R CHINA	2000	616	0.55
2	USA	2000	558	0.22
3	INDIA	2012	185	0.00
4	ENGLAND	2000	182	0.23
5	GERMANY	2000	179	0.44
6	SOUTH KOREA	2000	153	0.04
7	ITALY	2003	138	0.12
8	SPAIN	2004	135	0.11
9	AUSTRALIA	2000	96	0.19
10	JAPAN	2000	92	0.24
11	CANADA	2000	88	0.00
12	SAUDI ARABIA	2002	86	0.18
13	FRANCE	2014	84	0.49
14	TAIWAN	2000	72	0.26
15	PAKISTAN	2000	51	0.27



### 4.3 Co-citation network analysis

Co-citation analysis offers a lens through which the academic landscape of a field can be examined in detail. Based on the 803 articles screened during the second phase, this section conducts both publication co-citation and literature co-citation analyses in the field of Human-Computer Intelligent Interaction (HCII).

#### 4.3.1 Publication Co-citation Analysis

The co-citation of publications serves as a valuable metric for identifying influential work within a scientific domain. Academic publications constitute the building blocks of any research field, and analyzing them can guide both researchers and practitioners toward seminal work in the area. As illustrated in Figure 9, nodes in this analysis represent cited journals. The network comprises: A total of 748 nodes ( $N = 748$ ), 1,939 links ( $E = 1,939$ ), A collaborative density bit of 0.0069. The size of each node reflects the journal's co-citation frequency. Journals with more than 60 co-citations are prominently displayed in the network. Notably, "Lecture Notes in Computer Science" leads the pack with 226 publications, closely followed by "IEEE Access" with 222 citations and "Sensors" with 181 citations. The high co-citation frequency of these core journals suggests that they publish research that is both in-depth and exhaustive. Their high co-citation count is indicative of their significant contributions to the field and their value as reliable sources of high-impact research.

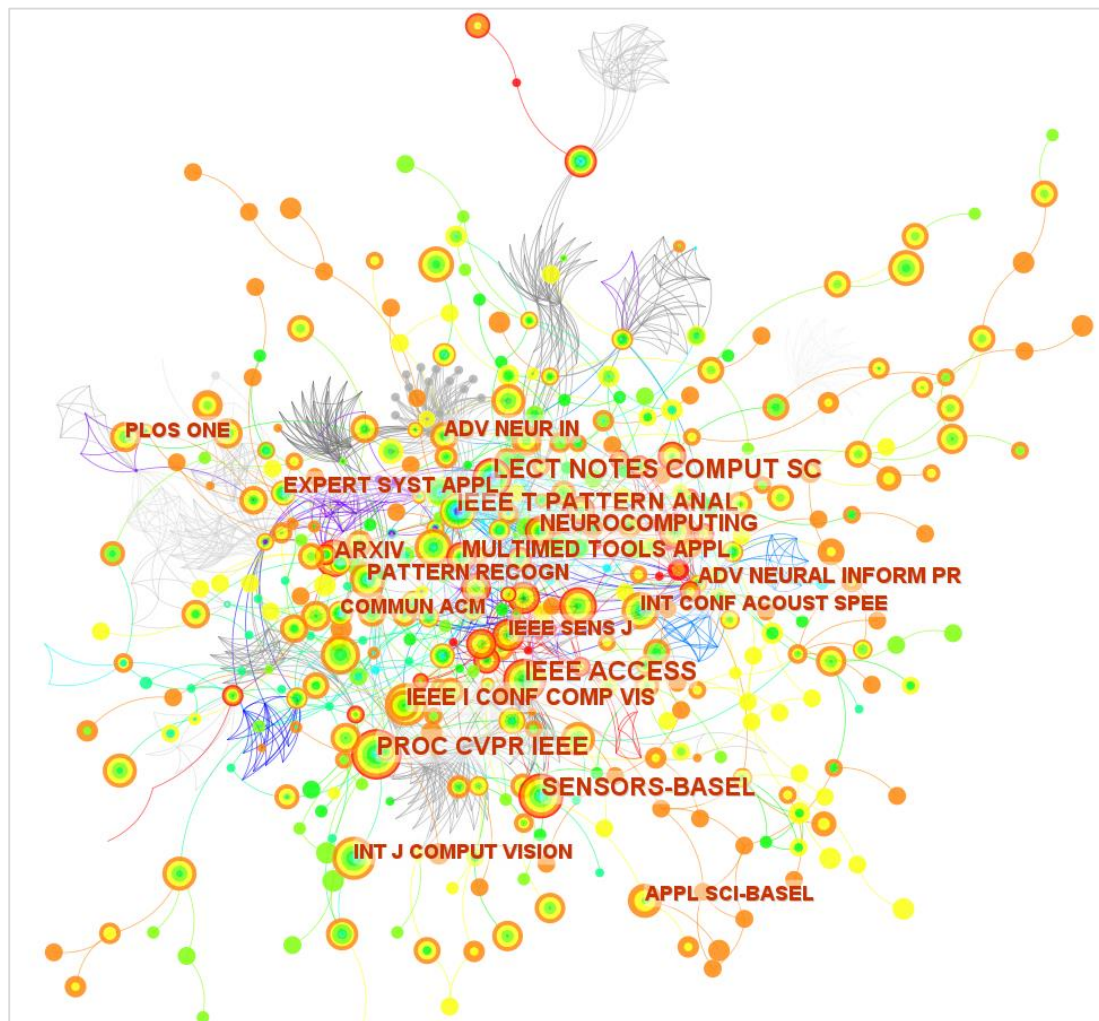


Fig.9 Co-citation network map of publications in the HCII domain of the WoS core database

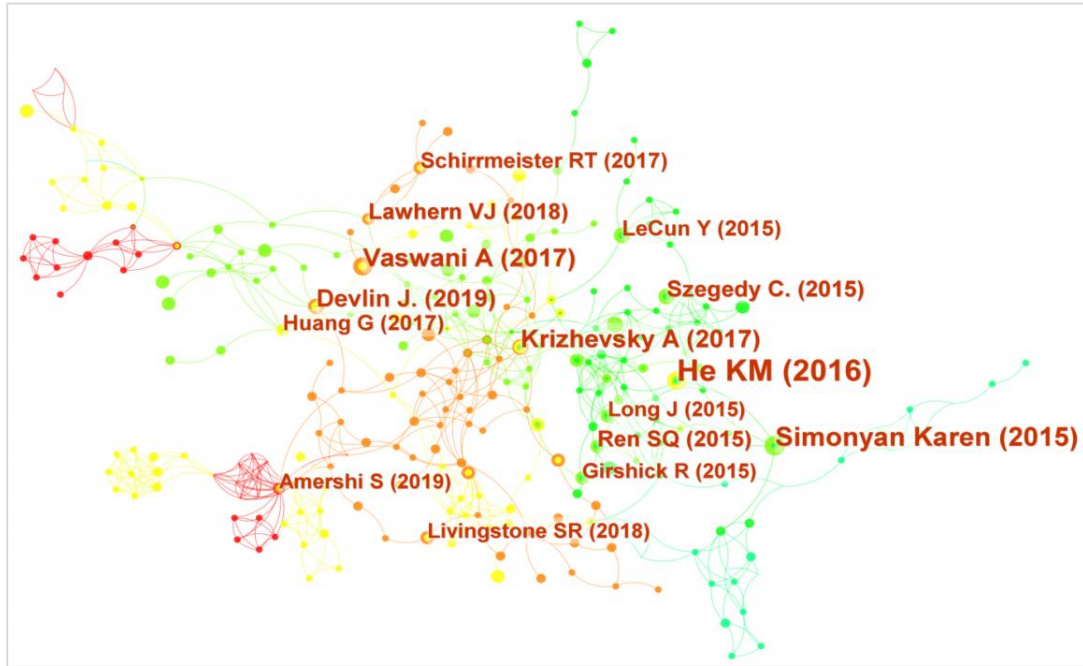
Table 4 Top 10 publications in the HCII area of the WoS core database

Serial No.	Cited Journals	Year of Debut	Citations Count
1	LECT NOTES COMPUT SCI	2006	226
2	IEEE ACCESS	2018	222
3	SENSORS	2014	181
4	PROC CVPR IEEE	2016	177
5	IEEE T PATTERN ANAL	2001	136
6	ARXIV	2018	133
7	MULTIMED TOOLS APPL	2015	108
8	IEEEI CONF COMP VIS	2016	94
9	NEUROCOMPUTING	2015	94
10	EXPERT SYST APPL	2002	85

### 4.3.2 Literature co-citation analysis

Literature co-citation analysis serves as an invaluable tool for understanding the intellectual structure of a scientific field. By examining the co-citation relationships between academic articles, this analytical approach uncovers the seminal works that have shaped the field, identifies influential technologies, and elucidates current trends. The references within the HCII field are conceptualized as nodes to form a co-citation network, as illustrated in Figure 10. The network exhibits the following characteristics: 706 nodes ( $N = 706$ ), 1,870 links ( $E = 1,870$ ), a collaborative density of 0.0075. Overall, the network reveals a robust pattern of citation relationships, highlighting an environment of interdisciplinary learning and cross-citation. The most frequently cited literature in the HCII field includes: He, Kaiming, et al. (2016) with 26 citations, Simonyan, K. with 17 citations, Vaswani, A. with 15 citations. These works predominantly focus on foundational artificial intelligence algorithms and architectures, especially those based on machine learning and deep learning. These algorithms are often deployed for tasks involving image and language recognition and classification. Specifically, they frequently employ convolutional neural networks as the basis for more advanced algorithms [120-121]. The prominence of these works suggests that they have contributed significantly to the development and advancement of the HCII field. By identifying these cornerstone publications, we can gain a comprehensive understanding of the state-of-the-art technologies and methods that currently shape HCII.





**Fig.10** Co-citation network of literature in the HCII domain of the WoS core database

**Table 5** Top 10 cited documents in the HCII domain of the WoS Core Database

No.	Cited Literature	Count	Year	DOI
1	He, Kaiming et al.	26	2016	DOI 10.1109/cvpr.2016.90_
2	Simonyan K et al.	17	2015	DOI 10.48550/arXiv.1409.1556
3	Vaswani A et al.	15	2017	DOI 10.48550/arXiv.1706.03762
4	Krizhevsky A et al.	14	2017	<a href="https://doi.org/10.1145/3065386">DOI 10.1145/3065386</a>
5	Devlin, J et al.	13	2019	DOI 10.18653/v1/N19-1423_
6	Szegedy C et al.	11	2015	DOI 10.1109/CVPR.2015.7298594_
7	Lawhern VJ et al.	10	2018	DOI 10.1088/1741-2552/aace8c
8	Ren SQ et al.	10	2015	DOI 10.1109/TPAMI.2016.2577031
9	LeCun Y et al.	9	2015	DOI 10.1038/nature14539
10	Huang G et al.	9	2017	DOI 10.1109/CVPR.2017.243

## 4.4 Analysis of disciplinary symbiotic networks

### 4.4.1 Keyword Co-occurrence Analysis

Keywords serve as the intellectual core of an academic article, encapsulating the paper's primary focus, themes, methodologies, and research scope. They provide a snapshot of the research domain, methodologies, and specific topics covered. By analyzing keyword co-occurrence, we can construct a network graph that reflects the thematic concentration of the field. In Figure 11, the size of each node represents the frequency of keyword occurrence. Nodes with purple edges indicate higher centrality values. The larger the purple circle, the more central the keyword is to the field. The keyword co-occurrence network consists of: 379 Keywords (N = 379), 850 Links (E = 850), A density of 0.0119.

This indicates a strong interrelation between the keywords, revealing a cohesive and robust correlation network for the HCII field. Table 6 lists the top 15 most frequently occurring keywords, their first year of appearance, their frequency, and their centrality measures. The most frequently occurring keywords are "Deep Learning" and "Machine Learning," with the latter having the highest mediational centrality. This suggests that machine learning serves as a foundational keyword within the field. Deep learning is commonly applied in areas such as multimodal human-computer interaction [4] and user interface evaluation [123]. The next most frequent keywords are: Human-Computer Interaction, Feature Extraction, Artificial Intelligence, Neural Network, Convolutional Neural Network, Model, System, Emotion Recognition, Gesture Recognition.

These keywords suggest that the primary focus of HCII research is on the "human" aspect of human-computer interaction. Specifically, there is a strong emphasis on recognizing various human senses, postures, and emotions. This is enabled by technologies such as artificial intelligence, and the most frequently used models aim to optimize the system for enhanced user interface usability and increased HCI effectiveness in a range of tasks. The keyword co-occurrence analysis provides valuable insights into the overarching themes and focal points of HCII research. The analysis reveals a strong emphasis on machine learning technologies, with a particular focus on deep learning, and highlights the role of AI in enhancing human-computer interactions.



**Fig.11** keyword co-occurrence network diagram for the HCII domain of the WoS core database

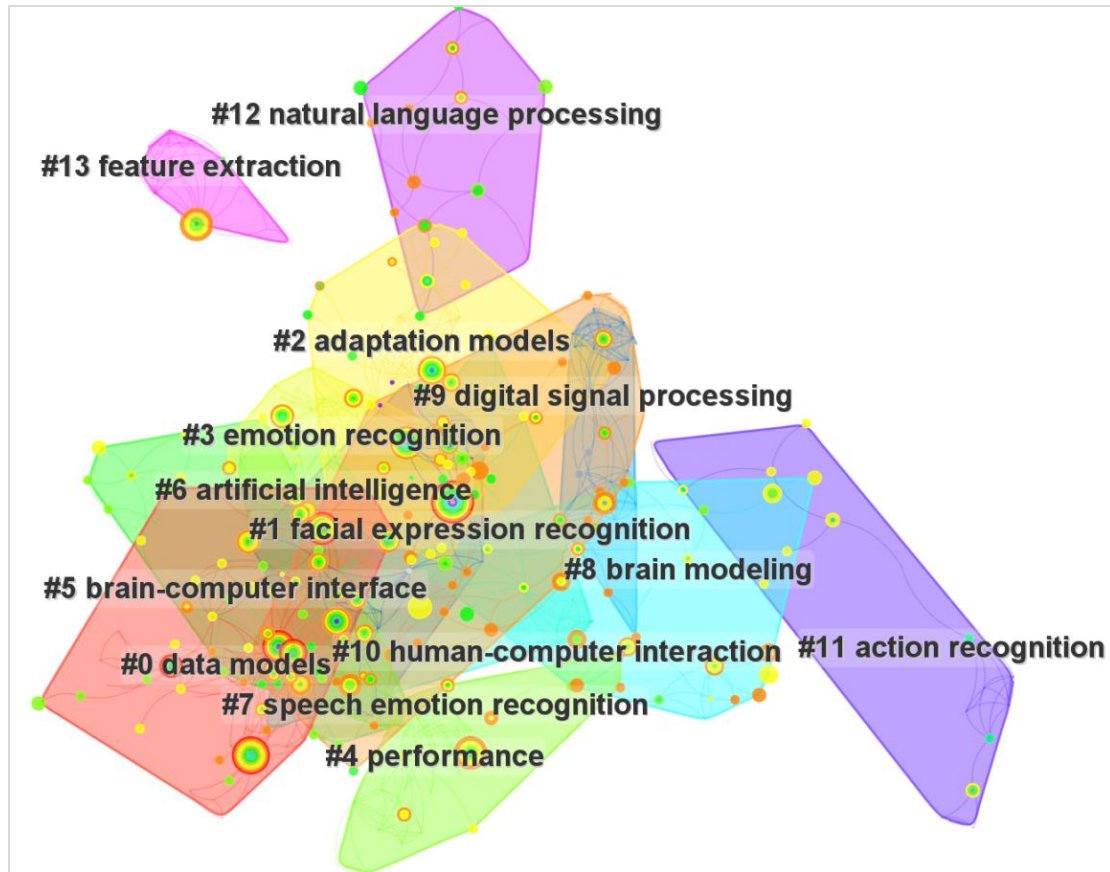
**Table 6** .top 15 high-frequency keywords in the HCII domain of the WoS core database

Serial No.	Keywords	Year of Debut	Frequency	Centrality
1	Deep learning	2015	127	0.06
2	Machine learning	2007	80	0.25
3	Human-computer interaction	2007	72	0.14
4	Feature Extraction	2017	54	0.07
5	Artificial Intelligence	2003	49	0.18
6	Neural Network	2001	43	0.00
7	Convolutional Neural Network	2016	41	0.11
8	Model	2003	38	0.04
9	System	2001	35	0.35
10	Emotion Recognition	2012	34	0.10
11	Gesture Recognition	2016	31	0.01
12	Recognition	2014	31	0.02
13	Classification	2017	30	0.08
14	Design	2014	21	0.09
15	Network	2004	20	0.09

#### 4.4.2 keyword clustering

Keyword clustering is an advanced technique used to group related keywords together, providing a lens through which to view various research hotspots within the domain. This kind of clustering can reveal the knowledge structure of a given field, offering insights into its core components and emergent trends. In this analysis, keywords were clustered using the Log-Likelihood Ratio (LLR) algorithm. The parameter for the "K clusters showing the largest in the cluster" was set to 14 to isolate the top 14 clusters within this research area. This resulted in the keyword clustering map for the HCII research domain shown in Figure 12. Modularity Value = 0.791 > 0.3, Q=0.791 > 0.3: Indicates that the clusters have a significant structure. Average Silhouette Value = 0.9056 > 0.5, S=0.9056 > 0.5: Suggests that the clusters are well-defined and significant. The most prominent clusters identified were: Data Models (#0), Facial Expression Recognition (#1), Adaptation Models (#2). These clusters indicate that extensive research has been conducted in these specific areas.

The complete list of clusters, in order of prominence, includes: Data Models (#0), Facial Expression Recognition (#1), Adaptation Models (#2), Performance (#4), Brain-Computer Interface (#5), Artificial Intelligence (#6), Speech Emotion Recognition (#7), Brain Modeling (#8), Digital Signal Processing (#9), Human-Computer Interaction (#10), Action Recognition (#11), Natural Language Processing (#12), Feature Extraction (#13). The keyword clustering analysis reveals a well-defined and structured field with several key areas of focus. This provides a comprehensive overview of the current state of HCII research, highlighting major themes and pointing to potential avenues for future exploration.



**Fig.12** .network diagram of keyword clustering mapping for HCII domains in the WoS core database

A preliminary analysis indicates that current research in the Human-Computer Intelligent Interaction (HCII) field predominantly lies at the intersection of three disciplines: Human Behavior (Cluster 4), Artificial Intelligence (Cluster 6), Human-Computer Interaction (HCI) (Cluster 10). The primary data source is human behavior, which is integrated with HCI and AI-related knowledge and technology to enhance the effectiveness and user experience of intelligent user interfaces (IUI) [123]. **The remaining clusters can be manually screened and reorganized into five main groups: Data Processing, Brain-Computer Interface, Multimodal Emotion Recognition, Action Recognition, Language Processing.**

(1) Data Model Processing: Cluster 0 (Data Models): This cluster focuses on the concept of data models, which are tools used in database design to abstract real-world scenarios for better data description [124]. Data models are the core foundation of any database system [125]. In the context of HCII, researchers predominantly use visual data, such as video and still images, or rely on sensor data to acquire target data [126]. Wearable sensors and computer vision technology are significant tools for acquiring human data models [127-128]. Additionally, 3D data models of digitized products often require a depth sensor for data acquisition and processing [129]. Cluster 2 (Adaptation Models): This cluster has gained popularity across various fields recently. Adaptation models serve as one of the most effective ways to improve the performance of data models. They are often used for dataset training in webpage recommendation models [130]. By recording user behavior during webpage interactions, a fuzzy inference system is constructed to evaluate and continually optimize user experience [131]. Domain adaptation, when applied to complex and large image tasks with the same training dataset, significantly improves model performance [132-133]. Within the HCII domain, the AI methods predominantly used are Deep Learning (DL) models, Convolutional Neural

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

Networks (CNN), Long Short-Term Memory networks (LSTM), and Support Vector Machines (SVM). Researchers employ various network models to acquire datasets, which are then compared and analyzed experimentally. Based on these experiments, suitable models are selected, and optimizations are proposed to improve the precision or robustness of experimental results. In summary, the HCII field is a multidisciplinary area that has been making significant strides in integrating human behavior analytics, artificial intelligence, and HCI to create more intelligent, effective, and user-friendly interfaces.

(2) Brain-Computer Interface: Cluster 5 (Brain-Computer Interface): This cluster is an emerging hot topic that bridges the HCII and biomedical fields. It primarily relies on EEG signals to control external devices, incorporating signal processing, classification techniques, and control theories to manage complex environments or execute fine motions [134]. Applications often target individuals with paralysis and include Deep Learning (DL) technology-based assistive rehabilitation devices with IoT modules [135][61], as well as assistive mobility devices for disabled patients [136]. The technology's success has also fueled the growth of the neurogaming industry. In this context, EEG signals form datasets, Support Vector Machines (SVMs) and Long Short-Term Memory networks (LSTMs) train learning algorithms for in-game decision-making [27], and sensor devices host the game engine for user testing. Research in brain-computer interaction games primarily targets the enhancement of human-computer interaction and user experience and is mainly applied to board games like mahjong [137] and backgammon [75].

(3) Multimodal Emotion Recognition: Cluster 3 (Emotion Recognition): This cluster is a key research area in HCII. Emotion detection enables more natural computer interaction and the design of human-centered user interfaces and systems. The cluster encompasses Cluster 1 (Facial Expression Recognition) and Cluster 7 (Speech Emotion Recognition), both of which are hot topics. Facial expression recognition is particularly robust; it allows for direct assessment of user emotions. Researchers have enhanced the accuracy of such recognition by refining convolutional neural network (CNN) approaches [88-89]. Speech emotion recognition, on the other hand, focuses on extracting emotional cues from discourse [76] and has shown improved classification capabilities with one-dimensional CNNs (1D CNNs). These technologies underscore the importance of speech emotion recognition in intelligent interaction and illustrate that most research in this area employs deep learning techniques to improve data model accuracy [90]. In line with the multimodal nature of emotion recognition models, existing frameworks like CNN and Gated Recurrent Units (GRU) are also being used to extract emotional features from videos, actions, and even brain activity, aiming to enhance multi-dimensional recognition accuracy [87][138]. Overall, the HCII field is making strides in integrating advanced AI methodologies with human-centered approaches across various applications, from healthcare and rehabilitation to gaming and emotion recognition. These advances not only improve the user experience but also open new avenues for research and application in real-world scenarios.

(4) Action Recognition: Cluster 11 (Action Recognition): This cluster has a broad range of applications, including human-computer interaction, healthcare, intelligent surveillance, autonomous driving, and virtual reality. The research within this cluster mainly targets human pose estimation, human action capture, and hand action recognition. Human pose estimation often begins with refining a dataset that captures various joints and key body parts, which can be applied in various studies [84]. There are a growing number of studies focusing on human motion capture and recognition, largely based on deep learning techniques. For example, local error convolutional



neural network models [85] and hybrid deep learning-based activity recognition models like QWSA-HDLAR [86] offer significant insights into action recognition.

(5) Language Processing: Cluster 12 (Natural Language Processing): NLP technologies have been successfully integrated into the HCII field, especially in dialogue systems across various sectors. These include human-like dialogue systems for industrial robots [22-23], entertainment chatbots for the elderly to promote digital inclusion [32][139], smart home systems with voice commands [12], and educational systems [24]. Additionally, there is a wealth of research in text mining and sentiment analysis. For example, semantic networks are being used for humor recognition in text [140], evaluations in VR games [28], and assessments of learning forum discussions [25]. Image caption generation is another growing area within this cluster [141].

The research in the HCII field is increasingly becoming interdisciplinary, merging advanced computational methods with a deep understanding of human behavior and needs. Whether it is through the lens of action recognition for healthcare and surveillance or natural language processing for more interactive and intuitive user interfaces, the advancements in HCII are setting the stage for a future where technology is not just a tool but an extension of human capability.

#### 4.4.3 Keyword timeline analysis

The keyword timeline serves as a visual tool to track the development, duration, and interconnections among the hotspots in HCII research. By plotting the 14 clusters in a timeline view, as demonstrated in Figure 13, several key insights emerge: **Interconnected Research:** The clusters within the HCII field are interconnected, reflecting an increasing trend in interdisciplinary studies. **Longevity and Impact:** Clusters such as #0 data models, #4 performance, and #7 speech emotion recognition have been around for a long time and have remained persistent in the field. For instance, data models have been pivotal since the early stages of HCII research, indicating their foundational role. **Emerging Focus Areas:** In recent years, areas like facial expression recognition, brain-computer interface, speech emotion recognition, brain modelling, and natural language processing have gained prominence. While clusters like #1 (facial expression recognition), #5 (brain-computer interface), and #8 (brain modelling) are relatively new, they show strong connections with other clusters, suggesting that they are extensions or innovations based on foundational research and current needs. **Leading Clusters:** The keywords 'sentiment recognition,' 'model,' 'design,' and 'user interface' frequently appear across the timeline, especially in cluster #3. This indicates the leading role and extensive, high-quality research value of cluster #3 in the HCII field. **Bridging Research:** Digital signal processing (cluster #9) may have appeared later in the timeline but shows an extended duration and links with other clusters, serving as a bridge in the overall research landscape. These insights not only offer a historical perspective but also help to identify enduring and emerging areas of interest, thereby providing a comprehensive understanding of the field's evolution. It is clear that as HCII research matures, it is becoming increasingly interdisciplinary, integrating more closely with emerging technologies and user needs.

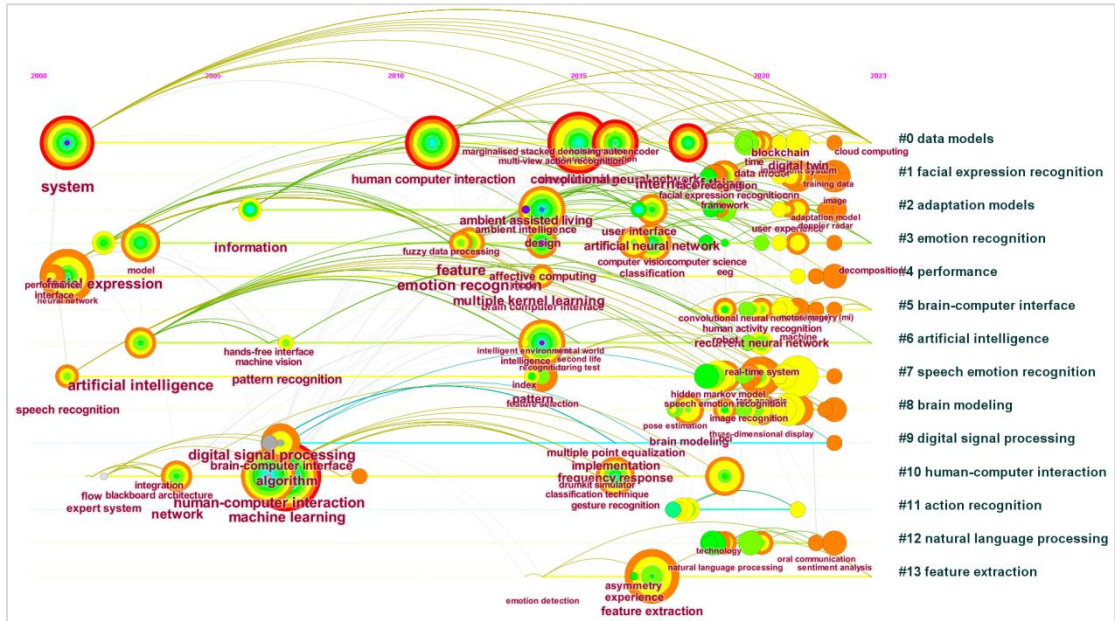
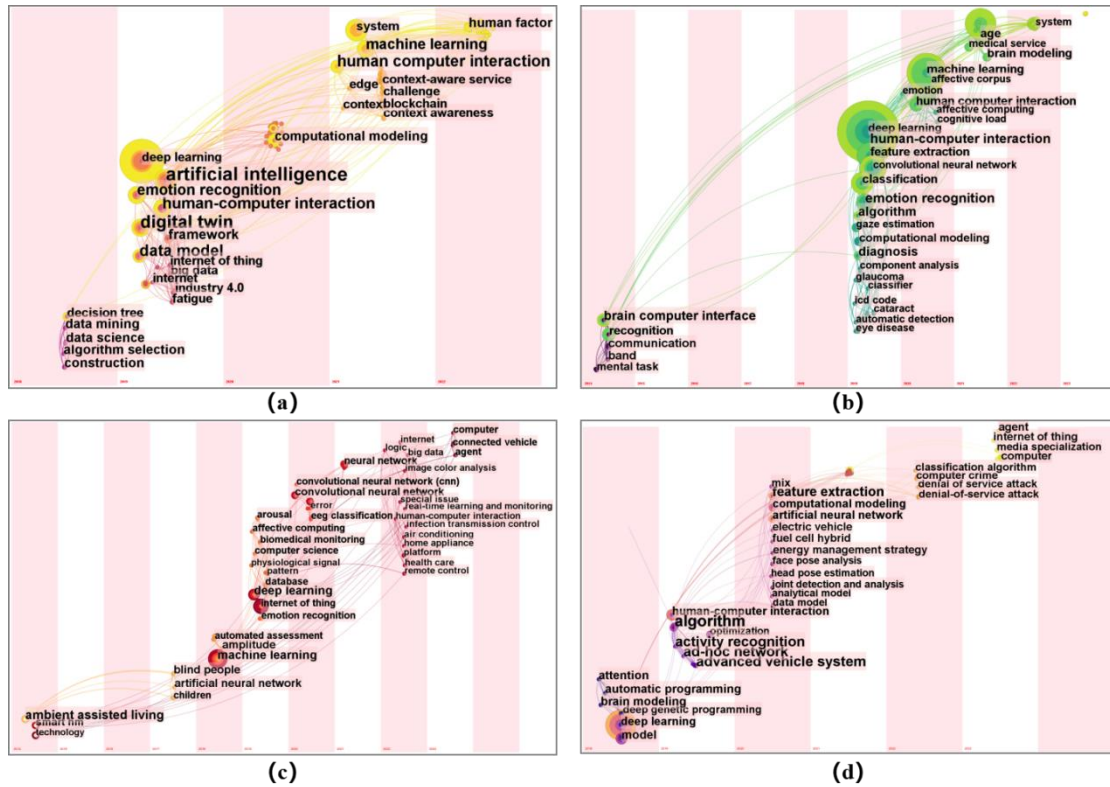


Fig.13 WoS core database HCII domain keyword clustering timeline

## 5. Analysis and discussion of results

### 5.1 HCII application area hotspot time zone map

The time-zone diagram, depicted in Figure 14, focuses on the four most prevalent application areas in HCII research: industrial manufacturing, healthcare, smart homes, and vehicles & transportation. The diagram serves multiple purposes: Identifying Common Hotspots: Keywords like "computational modelling," "programming," and "model" frequently appear across these sectors, signaling these as common areas of research focus. Unveiling Technical Methods: "ML/DL" (Machine Learning/Deep Learning) emerge as hot keywords, indicating that the future of HCII is increasingly intertwined with advancements in artificial intelligence. Computational modelling stands out as the most commonly used technique. Emotion Computing and Recognition: This has become a significant area of interest, especially in healthcare, smart homes, and industrial settings. The potential for emotion recognition technology to be integrated into a broader range of applications is high, aiming for increased accuracy and enhanced user experience. The Rise of IoT: In recent years, the Internet of Things (IoT) has gained significant attention, particularly in vehicles, transportation, and industrial applications. IoT technology combined with HCII can revolutionize various sectors: In Transportation: Smart parking, smart navigation, and traffic flow supervision are a few areas where this amalgamation can be beneficial. In Industry: Applications extend to smart warehousing, smart supply chains, and smart robotic arms, among others. In Smart Homes: Home appliance interconnection, home health, and security can be improved through IoT and HCII integration. The time-zone diagram analysis reveals that HCII is moving towards a future deeply integrated with emerging technologies like artificial intelligence and IoT. Emotion recognition, once limited to specific contexts, is finding broader applications and is expected to evolve further. These developments present both opportunities and challenges for HCII research. On one hand, there is enormous potential for breakthroughs in improving human-computer interactions in various sectors. On the other, these advancements also demand a multidisciplinary approach to address complex challenges in data accuracy, user experience, and system integration.



**Fig.14** industrial manufacturing (a),healthcare (b),smart home (c),vehicles and transportation (d) with HCII domain keyword time zone map

In alignment with the Industry 4.0 paradigm, HCII in this sector is rapidly incorporating digital twin and IoT technologies. Decision tree algorithms, a classical AI technique, are also being applied to optimize manufacturing processes and decision-making. HCII is increasingly prevalent in healthcare, particularly in psychological diagnosis, psychotherapy, and emotional well-being. Brain-computer interface technology is being integrated into HCII and is expected to find applications in rehabilitation therapy, neuro-gaming, and telemedicine. **HCII research in smart homes spans a broad and complex range, from foundational studies to practical applications, aimed at making homes more interactive, secure, and user-friendly. In this sector, the focus is primarily on human body recognition technologies, like head and facial gesture and action recognition. These technologies aim to improve the robustness and accuracy of HCII systems, leading to more intelligent and natural advanced vehicle systems. This also becomes a challenge for future HCII**

**Archaeological Site Detection:** HCII can facilitate the use of AI systems to assist archaeologists in identifying potential sites in aerial or satellite images. Semantic segmentation models can be used to draw accurate maps, enhancing operational efficiency.

**Cultural Heritage:** Digital collections and 3D visualizations are becoming new user requirements. An intelligent framework, built on AI technology, can allow museum visitors to interact virtually with cultural artifacts.

**Aerospace:** HCII technology could be extremely useful in hazardous flying missions in complex high-altitude environments. AI technology can simulate spatial orientation obstacles and provide enhanced navigation functions. The applications of HCII are not limited to the above-mentioned fields. Future research could extend into space and deep-sea explorations, indicating the limitless potential and versatile applications of HCII. The field promises to continuously evolve, integrating with other emerging technologies to address increasingly complex human-computer



interaction challenges across a myriad of scenarios.

## 5.2 Research Hot Spots and Trends in HCII

Based on our analysis, the three main centers of contemporary research in Human-Computer Intelligent Interaction (HCII) are as follows: Multimodal and Emotion Recognition Technology: The future of HCII lies in developing more diverse databases and achieving higher accuracy and better user experiences. The objective is to make human-computer dialogues and collaborations more natural and akin to "human-to-human" interactions. New Interaction Media and Scenarios: Innovations in interaction mediums like VR glasses and smart medical care are on the horizon. These mediums will likely make interactions more efficient and versatile. Advancements in AI Algorithms: Natural language processing has already become mainstream in AI, and the continuous improvements in deep learning algorithms are making computerized natural language more robust, accurate, and diverse.

Brain-Computer Interface Development: The integration of neurology and HCII is poised to allow physically handicapped patients to interact with external devices through EEG signals. This would significantly enhance their quality of life and offer applications like multi-modal neuro-interactive games and intelligent rehabilitation devices. Intelligent Human-Computer Interaction Systems: Future systems will have more complex computing capacities. Intelligent User Interfaces (IUIs) will offer a variety of sensory options and may even become de-physicalized, offering more natural and intuitive user experiences. Emerging Applications and Industries: HCII systems are expected to find applications in various emerging scenarios and industries. These could range from intelligent customer service in the financial sector, smart shopping carts in retail, intelligent control of urban facilities, to more realistic and engaging virtual reality interactions in the gaming industry. In summary, the field of HCII is rapidly evolving, with promising avenues for research and practical applications. The integration of advanced algorithms, new interaction mediums, and multidisciplinary approaches are expected to revolutionize how humans interact with computers, thereby impacting various sectors from healthcare to industrial manufacturing and beyond.

## 6. Conclusion

This paper offers a comprehensive analysis of the evolution of Human-Computer Intelligent Interaction (HCII), elucidating its core concepts, definitions, and research hotspots. It affirms that HCII is a cross-discipline that synergizes Human-Computer Interaction (HCI) and Artificial Intelligence (AI) technologies. This synergy not only amplifies the capabilities of HCI through AI but also guides the development of AI technology itself, thereby producing significant research results across various application scenarios.

There are our main findings: Interaction Modes: The paper identifies and analyzes the five primary interaction modes in HCII—voice interaction, visual interaction, gesture interaction, multimodal interaction, and emotional interaction—contrasting them with traditional HCI to offer a roadmap for future technological advancements in HCII. Research Hotspots: Utilizing CiteSpace for bibliometric analysis, the paper pinpoints the key research areas in HCII over the past five years, including data processing, brain-computer interfaces, multimodal emotion recognition, action recognition, and language processing. Future Trends: The paper predicts broader application

scenarios and greater research value for HCII in the future, particularly in the realms of multimodal interaction and emotion recognition. The development of more accurate algorithms and models, complex databases, and intelligent user interfaces is imminent. Innovations like neural interaction games, multi-sensory interaction, and de-physicalized interfaces are on the horizon. Complex and Natural Interactions: The use of advanced algorithms will likely make HCII systems more intuitive, allowing for more complex and natural interactions between humans and computers. Multimodal Emotion Recognition: Advances in this area will enable more nuanced and effective human-computer interactions, further narrowing the gap between human-to-human and human-to-computer communication. Societal Contributions: As HCII continues to evolve, it is set to offer more intelligent, intuitive, and efficient methods of interaction, thereby contributing significantly to the advancement of various sectors of human society. In summary, HCII is at a pivotal point in its development, with the potential for transformative impact on how humans interact with technology. As it continues to integrate advancements from both HCI and AI, HCII is poised to revolutionize various sectors, from healthcare to industrial manufacturing, and beyond.

## Funding

This research is funded by the State Key Laboratory of Mechanical Systems and Vibrations of China (Grant No.MSV202014); the Fundamental Research Funds for the Central Universities of China (HUST: 2020kfyXJJS014); the Major Philosophy and Social Science Research Project for Hubei Province Colleges and Universities of China (Grant No.20ZD002).

**Conflicts of Interest:** The authors declare that they have no conflict of interest.

**Data Availability Statement:** My manuscript has data included as electronic supplementary material.

## Reference

- [1] Hirschberg K, Manning CD (2023) Advances in natural language processing. *Science* 349:261-266.
- [2] Hussien RM, Al-Jubouri KQ, Gburi MA et al (2021) Computer Vision and Image Processing the Challenges and Opportunities for new technologies approach: A paper review. *Journal of Physics: Conference Series*, 1973(1):012002
- [3] Maja P, Alex P, Anton N et al (2006) Human computing and machine understanding of human behavior: a survey. *Association for Computing Machinery*:239-248
- [4] Lv Z, Poiesi F, Dong Q et al Deep (2022) Learning for Intelligent Human-Computer Interaction. *Appl. Sci* 12,11457
- [5] Wang G, Li L, Xing S et al (2018) Intelligent HMI in Orthopedic Navigation. *Advances in Experimental Medicine and Biology*, Springer
- [6] Miao Y, Jiang Y, Peng L et al (2018) Telesurgery Robot Based on 5G Tactile Internet. *Mobile Netw Appl* 23:1645-1654
- [7] R Li, H Fu, W L et al (2019) Skeleton-Based Action Recognition With Key-Segment Descriptor and Temporal Step Matrix Model. *IEEE Access*, pp 169782-169795
- [8] P Li, X Hou, X Duan et al (2019) Appearance-Based Gaze Estimator for Natural Interaction Control of Surgical Robots. *IEEE Access*, pp 25095-25110
- [9] Li P, Hou X, Wei L et al (2018) Efficient and Low-Cost Deep-Learning Based Gaze Estimator for Surgical Robot Control. *2018 IEEE International Conference on Real-time Computing and Robotics (RCAR)*
- [10] Zhang S, Song R, Cheng J et al (2019) A feasibility study of a video-based heart rate estimation method with convolutional neural networks. *2019 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*

- 1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65
- [11] A Singh,R Kabra,R Kumar et al (2021) On-Device System for Device Directed Speech Detection for Improving Human Computer Interaction. *IEEE Access*,pp 131758-131766
  - [12] Malagardi I,Alexandris C (2009) Verb Processing in Spoken Commands for Household Security and Appliances. *Lecture Notes in Computer Science*,Springer
  - [13] Hu W,Li X,Li Z (2022) Research on Auditory Performance of Vehicle Voice Interaction in Different Sound Index.In: Kurosu,M.(eds) *Human-Computer Interaction.User Experience and Behavior.HCII 2022.Lecture Notes in Computer Science*,vol 13304.Springer
  - [14] Waldron SM,Patrick J,Duggan GB et al (2008) 'Designing information fusion for the encoding of visual-spatial information'.*Ergonomics*,pp 775-797
  - [15] Li Z,Li X,Zhang J et al (2021) Research on Interactive Experience Design of Peripheral Visual Interface of Autonomous Vehicle.In: Kurosu,M.(eds) *Human-Computer Interaction.Design and User Experience Case Studies.HCII 2021.Lecture Notes in Computer Science*,12764.Springer
  - [16] Hazoor A,Terrafino A,Di Stasi LL et al (2022) How to take speed decisions consistent with the available sight distance using an intelligent speed adaptation system. *Accid Anal Prev*.Sep174:106758
  - [17] Kwan ML,Yohan M,Inyoung P (2023) Voice orientation of conversational interfaces in vehicles. *Behaviour & Information Technology*
  - [18] Shanthi N,Sathishkumar VE,Upendra BK et al (2022) Analysis on the Bus Arrival Time Prediction Model for Human-Centric Services Using Data Mining Techniques *Computational Intelligence and Neuroscience*
  - [19] Kim H,Kim W,Kim J et al (2022) Study on the Take-over Performance of Level 3 Autonomous Vehicles Based on Subjective Driving Tendency Questionnaires and Machine Learning Methods. *ETRI Journal* 45:75-92
  - [20] Zhiping Z,Aixue Y,Feng W et al (2010) Visual servo control of intelligent wheelchair mounted robotic arm.2010 8th World Congress on Intelligent Control and Automation,pp 6506-6511
  - [21] Chen L,Yang H,Liu P (2019) Intelligent Robot Arm: Vision-Based Dynamic Measurement System for Industrial Applications. *Lecture Notes in Computer Science*,Springer
  - [22] C Li,X Zhang,D Chrysostomou et al (2022) ToD4IR: A Humanised Task-Oriented Dialogue System for Industrial Robots. *IEEE Access*,pp 91631-91649
  - [23] Chen L,Jinha P,Hahyeon K et al (2021) How can I help you? An Intelligent Virtual Assistant for Industrial Robots. *Association for Computing Machinery*:220-224
  - [24] Kacalak W,Majewski M,Zurada JM (2010) Intelligent E-Learning Systems for Evaluation of User's Knowledge and Skills with Efficient Information Processing. *Lecture Notes in Computer Science*,Springer
  - [25] Cheong MLF,Chen JYC,Dai BT et al (2019) An Intelligent Platform with Automatic Assessment and Engagement Features for Active Online Discussions. *Lecture Notes in Computer Science*,Springer
  - [26] Choi Y,Jeon H,Lee S et al (2022) Seamless-walk: Novel Natural Virtual Reality Locomotion Method with a High-Resolution Tactile Sensor. *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*,pp 696-697
  - [27] Muhammad Amin,Abdallah Tubaishat,Feras Al-Obeidat et al (2022) Leveraging brain-computer interface for implementation of a bio-sensor controlled game for attention deficit people. *Comput.Electr.Eng.*102,C (Sep 2022)
  - [28] Gao Y,Chen A,Chi S et al (2022) Analysis of Emotional Tendency and Syntactic Properties of VR Game Reviews," *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*,pp 648-649
  - [29] Hu Z (2021) Eye Fixation Forecasting in Task-Oriented Virtual Reality. *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*,pp 707-708
  - [30] Kim J (2020) VIVR: Presence of Immersive Interaction for Visual Impairment Virtual Reality. *IEEE Access*,pp 196151-196159
  - [31] Krepki R,Blankertz B,Curio G et al (2007) The Berlin Brain-Computer Interface (BBCI) – towards a new communication channel for online control in gaming applications. *Multimed Tools Appl* 33:73-90
  - [32] García-Méndez S,Arriba-Pérez FD,González-Castaño FJ et al (2021) Entertainment Chatbot for the Digital Inclusion of Elderly People Without Abstraction Capabilities. *IEEE Access*,pp 75878-75891
  - [33] Lee W,Son G (2023) Investigation of human state classification via EEG signals elicited by emotional audio-visual stimulation. *Multimed Tools Appl*
  - [34] Razaq MA,Hussain J,Bang J (2023) Hybrid Multimodal Emotion Recognition Framework for UX Evaluation Using Generalized Mixture Functions. *Sensors* 23:4373
  - [35] Jo AH,Kwak KC (2023) Speech Emotion Recognition Based on Two-Stream Deep Learning Model Using Korean Audio Information. *Appl.Sci* 13:2167
  - [36] Aleisa HN,Alrowais FM,Negm N et al (2023) Henry Gas Solubility Optimization With Deep Learning Based Facial Emotion Recognition for Human Computer Interface. *IEEE Access* 11:62233-62241.
  - [37] Gagliardi G,Alfeo AL,Catrambone V et al (2023) Improving Emotion Recognition Systems by Exploiting the Spatial Information of EEG Sensors. *IEEE Access*,pp 39544-39554

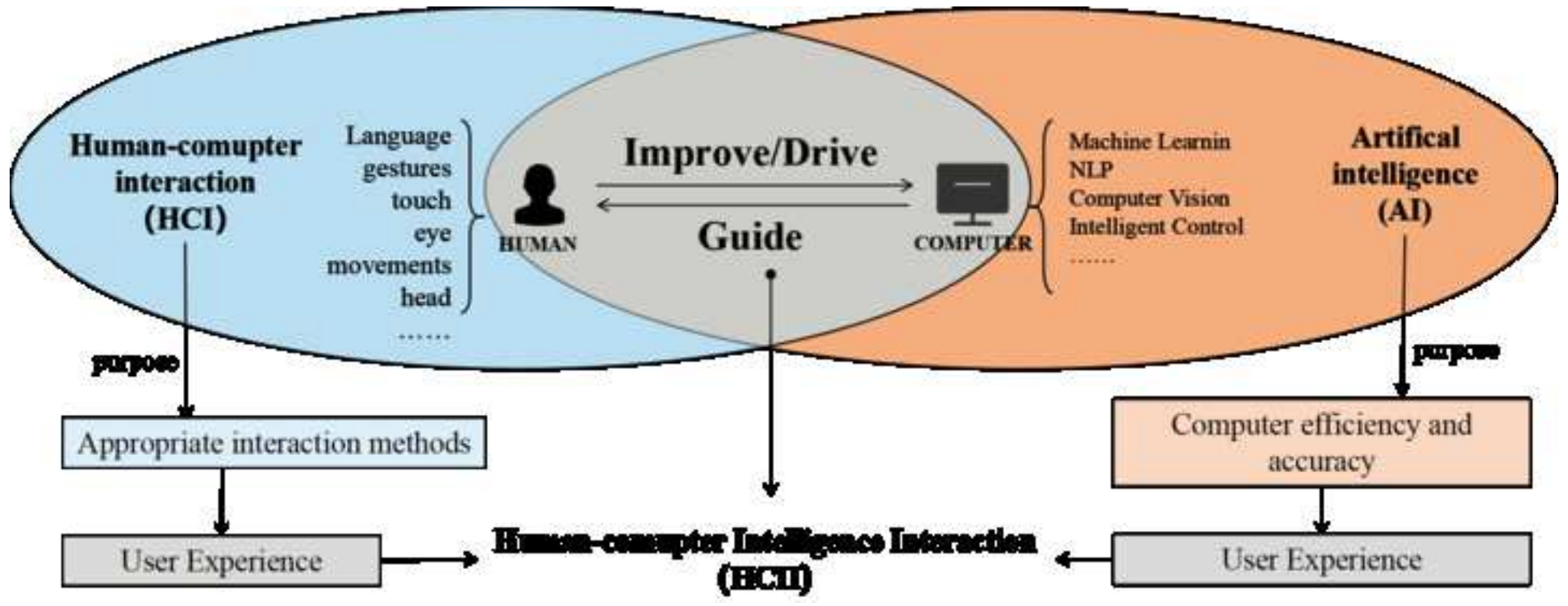
- 1 [38] Eswaran KCA,Srivastava AP,Gayathri M (2023) Hand Gesture Recognition for Human-Computer Interaction Using Computer Vision.Communications in Computer and Information Science,Springer
- 2 [39] Ansar H,Mudawi NA,Alotaibi SS et al (2023) Hand Gesture Recognition for Characters Understanding Using Convex Hull Landmarks and Geometric Features. IEEE Access 11:82065-82078
- 3 [40] Kothadiya DR,Bhatt CM,Rehman A et al (2023) SignExplainer: An Explainable AI-Enabled Framework for Sign Language Recognition With Ensemble Learning.IEEE Access,pp 47410-47419
- 4 [41] Sartaj AS,Ali Z,Hiroki T (2023) Cascaded deep graphical convolutional neural network for 2D hand pose estimation.International Workshop on Advanced Imaging Technology (IWAIT) ,1259215
- 5 [42] Yue L,Pengcheng A,Yage X (2023) Eggly: Designing Mobile Augmented Reality Neurofeedback Training Games for Children with Autism Spectrum Disorder.Proc.ACM Interact.Mob.Wearable Ubiquitous Technol
- 6 [43] Maarten VM,Rachel CS,Marie-Monique S et al (2023) Marianne Graves Petersen,and Ole Sejer Iversen.Emerging Technologies in K–12 Education: A Future HCI Research Agenda.ACM Trans
- 7 [44] Ometto M (2023) An innovative approach to plant and process supervision.Danieli Intelligent Plant
- 8 [45] Matheus N,Jorge J,Vissoci J et al (2023) Exploring affordances for AR in laparoscopy.Ithaca: Cornell University Library,arXiv.org
- 9 [46] Rawat KS, Sood SK (2021) Knowledge mapping of computer applications in education using CiteSpace. Comput Appl Eng Educ 29: 1324– 1339
- 10 [47] Grigsby SS (2018) Artificial Intelligence for Advanced Human-Machine Symbiosis.Lecture Notes in Computer Science,Springer
- 11 [48] Gomes CC,Preto S (2018) Artificial Intelligence and Interaction Design for a Positive Emotional User Experience.In: Karwowski,W.,Ahram,T.(eds) Intelligent Human Systems Integration.IHSI 2018.Advances in Intelligent Systems and Computing,Springer
- 12 [49] Zhang C,Lu Y (2021) Study on artificial intelligence: The state of the art and future prospects[J].Journal of Industrial Information Integration,2021(23-):23
- 13 [50] Ahamed MM (2017) Analysis of Human Machine Interaction Design Perspective-A Comprehensive Literature Review.International Journal of Contemporary Computer Research,1(1): 31-42
- 14 [51] Li X (2020) Human–robot interaction based on gesture and movement recognition.Signal Processing: Image Communication 81: 115686
- 15 [52] Majaranta P,Räihä KJ,Hyrskykari A et al (2019) Eye Movements and Human-Computer Interaction.Studies in Neuroscience,Psychology and Behavioral Economics,Springer
- 16 [53] Bi L,Pan C,Li J et al (2023) Discourse-based psychological intervention alleviates perioperative anxiety in patients with adolescent idiopathic scoliosis in China: a retrospective propensity score matching analysis.BMC Musculoskeletal Disorders
- 17 [54] Maybury MT (1998) Intelligent user interfaces: an introduction.Proceedings of the 4th international conference on Intelligent user interfaces,ACM
- 18 [55] Jaimes A,Sebe N (2005) Multimodal human computer interaction: A survey.Lecture notes in computer science 3766: 1
- 19 [56] Xiao LI (2004) Evolution and Trends of Intelligent User Interfaces.Computer Science
- 20 [57] Yiyi Zhao (2022) "Interaction Design System for Artificial Intelligence User Interfaces Based on UML Extension Mechanisms", Mobile Information Systems
- 21 [58] Margiené A,Ramanauskaitė S (2019) Trends and Challenges of Multimodal User Interfaces. 2019 Open Conference of Electrical,Electronic and Information Sciences,pp 1-5
- 22 [59] Maybury M,Wahlster W (1998) Intelligent user interfaces: An Introduction.In RUIU; Morgan Kaufmann: San Francisco,pp 3–4
- 23 [60] Mallios S,Bourbakis N,A survey on human machine dialogue systems.2016 7th International Conference on Information,Intelligence,Systems & Applications (IISA),Chalkidiki,Greece,pp 1-7
- 24 [61] Vinoj PG,Jacob SM,Varun G et al (2021) IoT-powered deep learning brain network for assisting quadriplegic people.Computers and Electrical Engineering
- 25 [62] Peruzzini M,Grandi F,Pellicciari M (2017) Benchmarking of Tools for User Experience Analysis in Industry 4.0.Procedia Manufacturing 11:806-813
- 26 [63] Mahesh B (2020) Machine learning algorithms-a review.International Journal of Science and Research (IJSR) 9: 381-386
- 27 [64] LeCun Y,Bengio Y,Hinton G (2015) Deep learning. Nature 521:436–444
- 28 [65] Cantoni V,Cellario M,Porta M (2004) Perspectives and challenges in e-learning: towards natural interaction paradigms.Journal of Visual Languages & Computing,15(5): 333-345
- 29 [66] Changhoon O,Jungwoo S,Jinhan C,et al (2018) You Help but Only with Enough Details: Understanding User Experience of Co-Creation with Artificial Intelligence.Association for Computing Machinery, Paper 649:1–13
- 30 [67] Turk M (2005) Multimodal Human-Computer Interaction.In: Kisačanin B,Pavlović V,Huang TS,(eds) Real-Time Vision for Human-Computer Interaction, Springer
- 31 [68] Singh SK,Chaturvedi AA (2022) reliable and efficient machine learning pipeline for american sign language gesture recognition using EMG sensors. Multimed Tools Appl
- 32
- 33
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45
- 46
- 47
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60
- 61
- 62
- 63
- 64
- 65

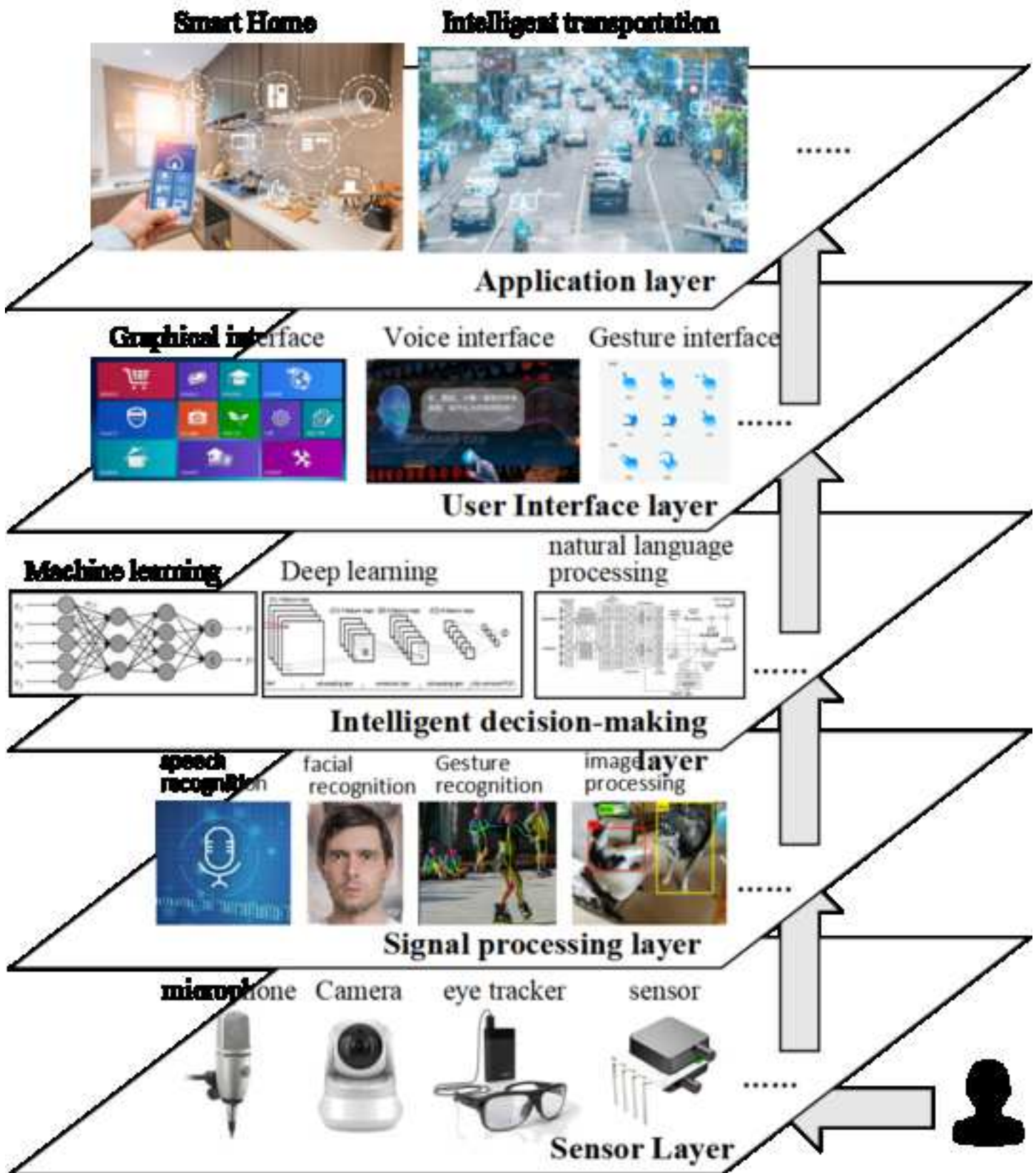
- 1 [69] Jinlei Z,Xue Q,Xiang L et al (2021) Support Vector Machine Weather Prediction Technology Based  
2 on the Improved Quantum Optimization Algorithm. Computational Intelligence and Neuroscience
- 3 [70] Giatsintov A,Mamrosenko K,Bazhenov P (2023) Architecture of the Graphics System for Embedded  
4 Real-Time Operating Systems.Tsinghua Science and Technology
- 5 [71] Lee S,Oh J,Moon WK (2023) Adopting Voice Assistants in Online Shopping: Examining the Role o  
6 f Social Presence,Performance Risk,and Machine Heuristic.International Journal of HumanComputer I  
7 nteraction
- 8 [72] Mctear MF (2002) Spoken dialogue technology - toward the conversational user interface,ACM
- 9 [73] Du Y,Qin J,Zhang S,et al (2018) Voice User Interface Interaction Design Research Based on User  
10 Mental Model in Autonomous Vehicle.In: Kurosu,M.(eds) Human-Computer Interaction.Interaction Te  
11 chnologies.HCI 2018.Lecture Notes in Computer Science,Springer
- 12 [74] Koni YJ,AI-Absi MA (2021) AI-Based Voice Assistants Technology Comparison in Term of Conver  
13 sational and Response Time.In: Singh M,Kang DK,Lee JH,Tiwary US,Singh D,Chung WY.
- 14 [75] Li M,Li F,Pan J et al (2021) The MindGomoku: An Online P300 BCI Game Based on Bayesian D  
15 eep Learning. Sensors 21,1613
- 16 [76] Abeer AA,Mohammed Z,Aseel A (2022) Human-Computer Interaction with Detection of Speaker Em  
17 otions Using Convolution Neural Networks. Computational Intelligence and Neuroscience
- 18 [77] Charissis V,Falah J,Lagoo R,et al (2021) Employing Emerging Technologies to Develop and Evaluat  
19 e In-Vehicle Intelligent Systems for Driver Support: Infotainment AR HUD Case Study. Appl.Sci 11,  
20 1397.
- 21 [78] Fan Y,Yang J,Chen J,et al (2021) A digital-twin visualized architecture for Flexible Manufacturing S  
22 ystem.Journal of Manufacturing Systems,2021(60-):60
- 23 [79] Wang T,Li J,Kong Z,et al (2020) Digital twin improved via visual question answering for vision-lan  
24 guage interactive mode in human-machine collaboration.Journal of Manufacturing Systems,58(4)
- 25 [80] Tian W,jiakun L,Yingjun D,et al (2021) Digital twin for human-machine interaction with convolutio  
26 nal neural network.International Journal of Computer Integrated Manufacturing, 34:7-8, 888-897
- 27 [81] Zhang Q,Wei Y,Liu Z,et al (2023) A Framework for Service-Oriented Digital Twin Systems for Dis  
28 crete Workshops and Its Practical Case Study. Systems 11,156
- 29 [82] El OI,Benouini R,Zenkouar K. et al (2022) RGB-D feature extraction method for hand gesture recog  
30 nition based on a new fast and accurate multi-channel cartesian Jacobi moment invariants. Multimed  
31 Tools Appl 81:12725-12757
- 32 [83] Miah ASM,Shin J,Hasan MAM et al (2022) BenSignNet: Bengali Sign Language Alphabet Recogniti  
33 on Using Concatenated Segmentation and Convolutional Neural Network. Appl.Sci 12,3933
- 34 [84] Munea TL,Jembre YZ,Weldegebriel HT et al (2020) The Progress of Human Pose Estimation: A Su  
35 rvey and Taxonomy of Models Applied in 2D Human Pose Estimation. IEEE Access,pp 133330-133  
36 348
- 37 [85] Xia H,Lei H,Yang J (2022) Human Behavior Recognition in Outdoor Sports Based on the Local Er  
38 ror Model and Convolutional Neural Network. Computational Intelligence and Neuroscience
- 39 [86] Malibari AA,Alzahrani JS,Qahmash A (2022) Quantum Water Strider Algorithm with Hybrid-Deep-L  
40 earning-Based Activity Recognition for Human-Computer Interaction. Appl.Sci 12,6848.
- 41 [87] Jia N,Zheng C,Sun W (2022) A multimodal emotion recognition model integrating speech,video and  
42 MoCAP. Multimed Tools Appl 81:32265-32286
- 43 [88] Yue D,Zhou L,Yue W (2016) Facial Expression Recognition Based on Convolution Neural Network.  
44 2nd International Conference on Computer Engineering,Information Science & Application Technolog  
45 y (ICCIA 2017)
- 46 [89] Tang Y,Zhang X,Hu X et al (2021) Facial Expression Recognition Using Frequency Neural Networ  
47 k. IEEE Transactions on Image Processing,pp 444-457
- 48 [90] Abbaschian BJ,Sierra-Sosa D,Elmaghraby A (2021) Deep Learning Techniques for Speech Emotion  
49 Recognition,from Databases to Models,Sensors 21(4)
- 50 [91] Lee CH,Yang HC,Su XQ (2022) A Multimodal Affective Sensing Model for Constructing a Persona  
51 lity-Based Financial Advisor System. Appl Sci 12,10066.
- 52 [92] Deng H,Liu M (2022) Personalized Smart Clothing Design Based on Multimodal Visual Data Detect  
53 ion. Computational Intelligence and Neuroscience
- 54 [93] Raptis GE,Kavvetsos G,Katsini C (2021) MuMIA: Multimodal Interactions to Better Understand Art  
55 Contexts. Appl.Sci 11,2695.
- 56 [94] Pandey SK,Shekhawat HS,Bhasin S et al (2022).Alzheimer's Dementia Recognition Using Multimoda  
57 l Fusion of Speech and Text Embeddings.Lecture Notes in Computer Science,Springer.
- 58 [95] Šumak B,Brdnik S,Pušnik M (2022) Sensors and Artificial Intelligence Methods and Algorithms for  
59 Human-Computer Intelligent Interaction: A Systematic Mapping Study. Sensors 22(20)
- 60 [96] Karpov AA,Yusupov RM (2018) Multimodal Interfaces of Human-Computer Interaction. Her.Russ.Ac  
61 ad.Sci. 88:67-74
- 62
- 63
- 64
- 65

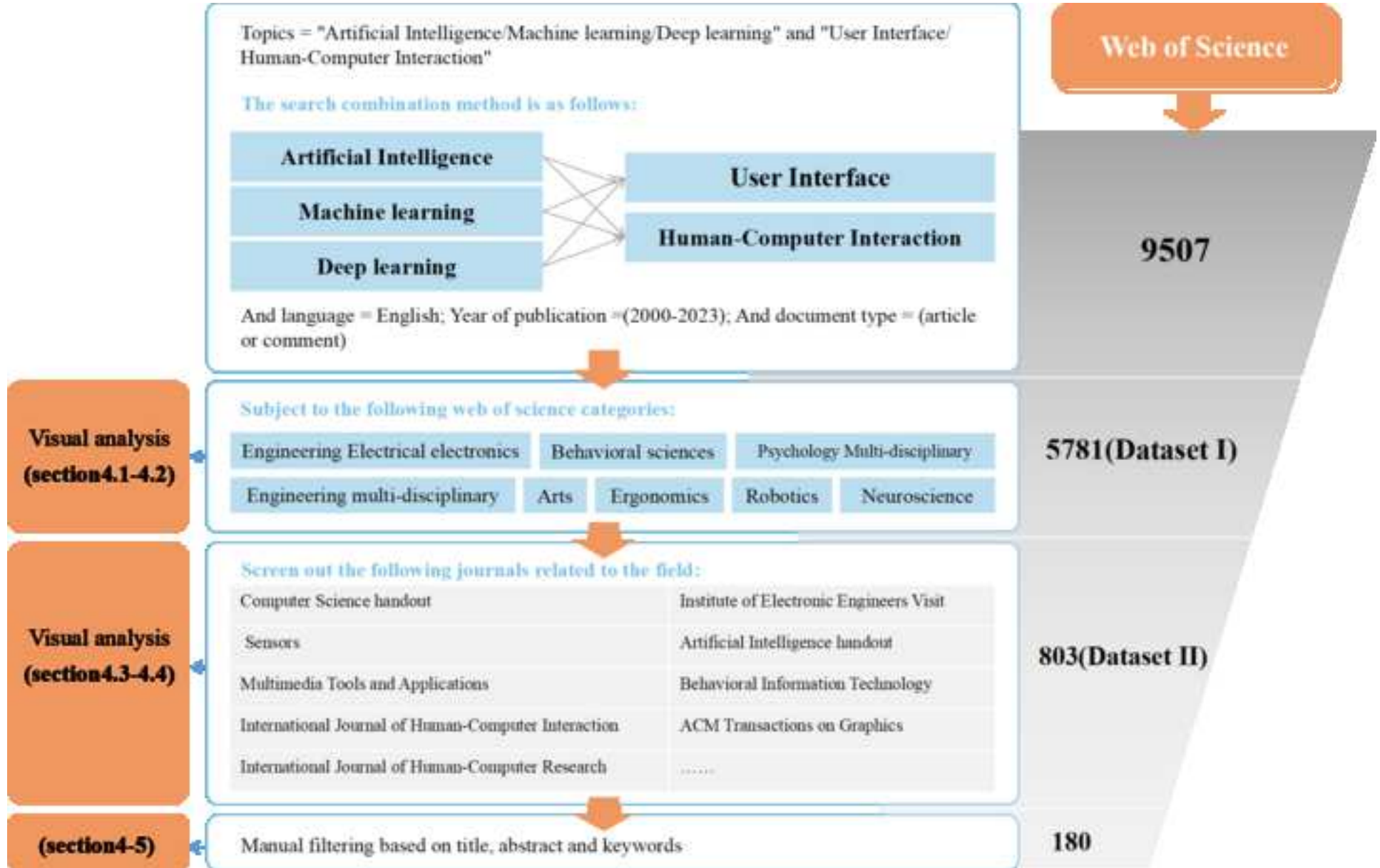
- 1 [97] Nazar M,Alam MM,Yafi E et al (2021) A Systematic Review of Human–Computer Interaction and  
2 Explainable Artificial Intelligence in Healthcare With Artificial Intelligence Techniques.IEEE Access.p  
3 p 153316-153348
- 4 [98] Diederich S,Brendel AB,Morana S et al (2022) On the Design of and Interaction with Conversational  
5 Agents: An Organizing and Assessing Review of Human-Computer Interaction Research.Journal of  
6 the Association of Information Systems,2022(1):23
- 7 [99] Kusal SS,Patil J,Choudrie K et al (2022) AI-Based Conversational Agents: A Scoping Review From  
8 Technologies to Future Directions. IEEE Access,pp 92337-92356
- 9 [100] Ren F,Bao Y (2019) A Review on Human-Computer Interaction and Intelligent Robots.International  
10 Journal of Information Technology and Decision Making 19(3)
- 11 [101] Vail EF (1999) Knowledge mapping: Getting started with knowledge management.Inf Syst Manag,pp  
12 10-23
- 13 [102] Lin S,Shen T,Guo W (2021) Evolution and Emerging Trends of Kansei Engineering: A Visual An  
14 alysis Based on CiteSpace.IEEE Access,pp 111181-111202
- 15 [103] Eck NJ,Waltman L (2010) Software survey: VOSviewer,a computer program for bibliometric mappin  
16 g. Scientometrics 84,523–538
- 17 [104] Cobo MJ,López-Herrera AG,Herrera-Viedma EE,et al (2012) SciMAT: A new science mapping analy  
18 sis software tool.Journal of the American Society for Information Science and Technology 63(8):160  
19 9-1630
- 20 [105] Chaomei C (2018) Visualizing and Exploring Scientific Literature with CiteSpace: An Introduction.In  
21 Proceedings of the 2018 Conference on Human Information Interaction & Retrieval (CHIIR '18).As  
22 sociation for Computing Machinery,369–370
- 23 [106] Wei F,Grubestic TH,Bishop BW (2015) Exploring the gis knowledge domain using citespace, Profess  
24 ional Geographer 67:374–384
- 25 [107] Zhong S,Chen R,Song F et al (2019) Knowledge Mapping of Carbon Footprint Research in a LCA  
26 Perspective: A Visual Analysis Using CiteSpace. Processes 7:818
- 27 [108] Yang J,Lu H (2022) Visualizing the Knowledge Domain in Urban Soundscape: A Scientometric Ana  
28 lysis Based on CiteSpace. International Journal of Environmental Research and Public Health,19.
- 29 [109] Chen Y,Wang Y,Zhou D (2021) Knowledge Map of Urban Morphology and Thermal Comfort: A B  
30 ibliometric Analysis Based on CiteSpace. Buildings 11:427
- 31 [110] Chen C (2017) Science mapping: A systematic review of the literature, J.Data Inform.Sci. 2 :1–40
- 32 [111] Li K,Jin Y,Akram MW et al (2020) Facial expression recognition with convolutional neural networks  
33 via a new face cropping and rotation strategy. Vis Comput 36:391–404
- 34 [112] Wang J,Su X (2019) Human Motion Modeling Based on Single Action Context. 2019 4th Internatio  
35 nal Conference on Communication and Information Systems (ICCIS),pp 89-94
- 36 [113] Zhao X,Song Z,Guo J et al (2012) Real-Time Hand Gesture Detection and Recognition by Random  
37 Forest.Communications in Computer and Information Science,Springer
- 38 [114] Zhang Y,Dong W,Hu BG et al (2018) Weakly-Supervised Deep Convolutional Neural Network Lear  
39 ning for Facial Action Unit Intensity Estimation.2018 IEEE/CVF Conference on Computer Vision an  
40 d Pattern Recognition,pp 2314-2323
- 41 [115] Guo Y,Liu W,Wei D et al (2019) Emotional Recognition Based on EEG Signals Comparing Long-t  
42 erm and Short-term Memory with Gated Recurrent Unit Using Batch Normalization
- 43 [116] Afsar MM,Saqib S,Ghadi YY et al (2022) Body worn sensors for health gaming and e-learning in  
44 virtual reality. Computers,Materials & Continua,pp 4763–4777
- 45 [117] Li X,Li Y (2022) Sports Training Strategies and Interactive Control Methods Based on Neural Netw  
46 ork Models.Computational Intelligence and Neuroscience
- 47 [118] Ahn J,Nguyen TP,Kim YJ et al (2021) Automated Analysis of Three-dimensional CBCT Images Ta  
48 ken in Natural Head Position That Combines Facial Profile Processing and Multiple Deep-learning  
49 Models
- 50 [119] Xi Y,Cho S,Fong S et al (2016) Gesture Recognition Method Using Sensing Blocks. Wireless Pers C  
51 ommun 91,1779–1797
- 52 [120] He K,Zhang X,Ren S (2015) Deep Residual Learning for Image Recognition. 2016 IEEE Conference  
53 on Computer Vision and Pattern Recognition (CVPR),770-778.
- 54 [121] Simonyan K,Zisserman A (2014) Very Deep Convolutional Networks for Large-Scale Image Recogni  
55 tion.Computer Science
- 56 [122] Keselj A,Milicevic M,Zubrinic K et al The Application of Deep Learning for the Evaluation of Use  
57 r Interfaces. Sensors 22,9336.
- 58 [123] Khalil RA,Jones E,Babar MI et al (2019) Speech Emotion Recognition Using Deep Learning Techni  
59 ques: A Review. IEEE Access 7:117327–117345.
- 60 [124] Bai J (2009) Panel Data Models With Interactive Fixed Effects.Econometrica 77(4):1229-1279.
- 61 [125] Wu F (2016) Study on Composition and Development of the Database Management System.2016 2n  
62 d International Conference on Advances in Mechanical Engineering and Industrial Informatics(AMEII  
63 2016)
- 64
- 65

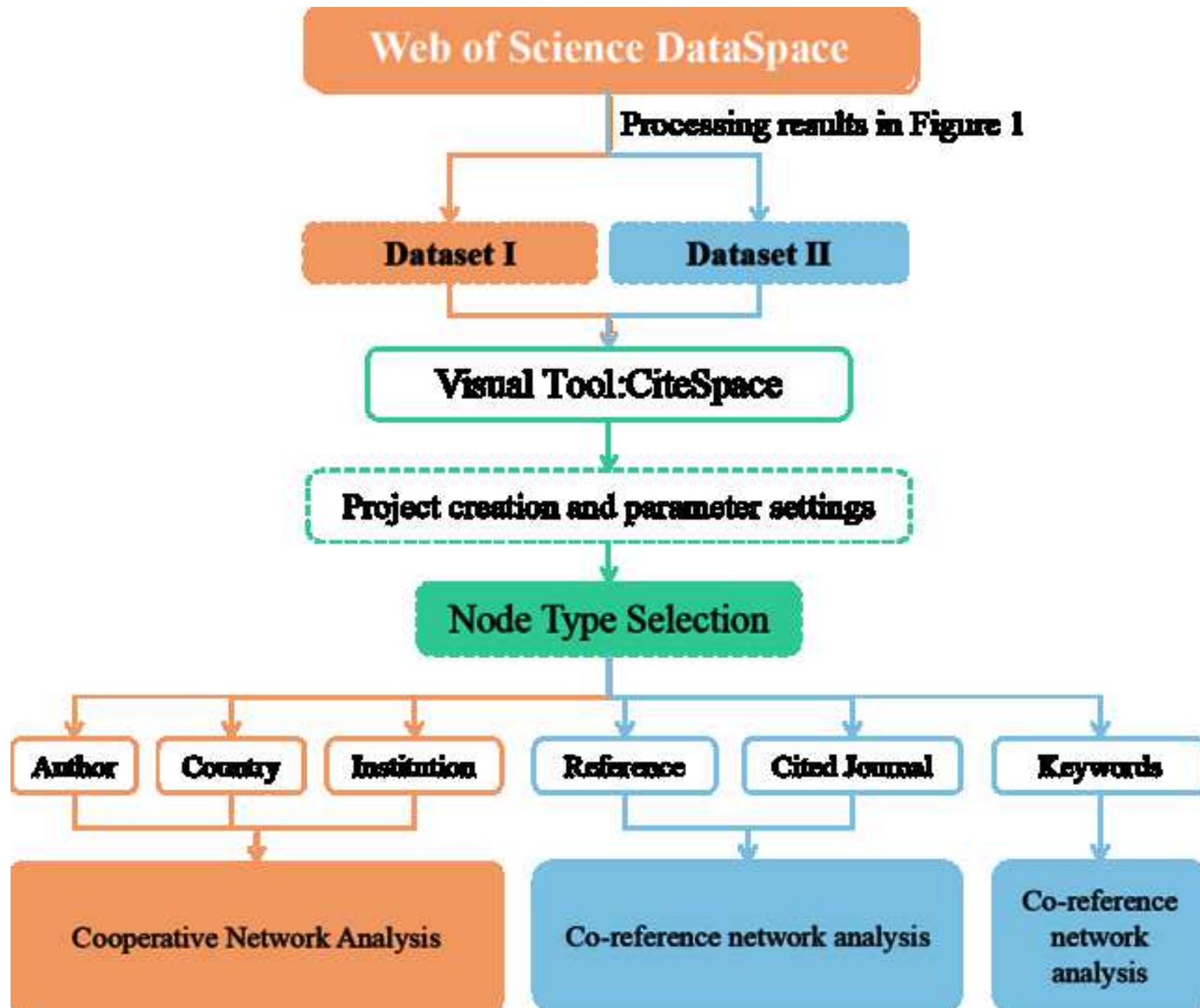
- 1 [126] Ozcan T,Basturk A (2020) Static facial expression recognition using convolutional neural networks based on transfer learning and hyperparameter optimization. *Multimed Tools Appl* 79:26587–26604
- 2 [127] Chen W,Yu C,Tu C et al (2020) A Survey on Hand Pose Estimation with Wearable Sensors and Computer-Vision-Based Methods.*Sensors* 20(4):1074.
- 3 [128] Lv T,Wang X,Jin L et al (2020) A Hybrid Network Based on Dense Connection and Weighted Feature Aggregation for Human Activity Recognition. *IEEE Access*,pp 68320-68332
- 4 [129] Belhi A,Ahmed HO,Alfaqheri T et al (2023) An integrated framework for the interaction and 3D visualization of cultural heritage. *Multimed Tools Appl*
- 5 [130] Correa JD,González BM,Hernández R et al (2014) Dynamic topic-based adaptation of language models: a comparison between different approaches
- 6 [131] Souza KE,Seruffo MC,De Mello HD (2019) User Experience Evaluation Using Mouse Tracking and Artificial Intelligence. *IEEE Access*,pp 96506-96515
- 7 [132] Basak S,Corcoran P,Khan F (2021) Learning 3D Head Pose From Synthetic Data: A Semi-Supervised Approach.*IEEE Access*,pp 37557-37573
- 8 [133] Chen YW,Pei SC (2022) Domain Adaptation for Underwater Image Enhancement via Content and Style Separation.*IEEE Access*,pp 90523-90534
- 9 [134] Padfield N,Camilleri K,Camilleri T (2022) A Comprehensive Review of Endogenous EEG-Based BCIs for Dynamic Device Control. *Sensors* 22,5802
- 10 [135] Jacob S,Alagirisamy M,Menon VG et al,An Adaptive and Flexible Brain Energized Full Body Exoskeleton With IoT Edge for Assisting the Paralyzed Patients. *IEEE Access*,pp 100721-100731
- 11 [136] Tang J,Liu Y,Jiang J et al (2019) Toward Brain-Actuated Mobile Platform. *International Journal of Human-Computer Interaction* 35(10): 846-858
- 12 [137] Wu X,Qi Y,Zhu X et al (2022) A Brain-Controlled Mahjong Game with Artificial Intelligence Augmentation.*CICAI 2022 Lecture Notes in Computer Science*,Springer
- 13 [138] Choi DY,Kim DH,Song BC (2020) Multimodal Attention Network for Continuous-Time Emotion Recognition Using Video and EEG Signals.*IEEE Access*,pp 203814-203826
- 14 [139] Pérez FD,García-Méndez S,González-Castaño FJ et al (2021) Evaluation of Abstraction Capabilities and Detection of Discomfort with a Newscaster Chatbot for Entertaining Elderly Users. *Sensors* 21:5515
- 15 [140] Xiong S,Wang R,Huang X (2022) Multidimensional Latent Semantic Networks for Text Humor Recognition. *Sensors* 22:5509
- 16 [141] Wan H,Zhang Y,Yu X (2020) An Overview of Image Caption Generation Methods", *Computational Intelligence and Neuroscience*
- 17 [142] Casini L,Marchetti N,Montanucci A et al (2023) A human-AI collaboration workflow for archaeological sites detection. *Sci Rep* 13:8699
- 18 [143] Minaskan N,Alban-Dromoy C,Pagani A et al (2022).Human Intelligent Machine Teaming in Single Pilot Operation: A Case Study.*Lecture Notes in Computer Science*,Springer,.
- 19 [144] Foucher J,Coller AC,Goff KL (2022) et al Simulation and Classification of Spatial Disorientation in a Flight Use-Case Using Vestibular Stimulation.*IEEE Access*,pp 104242-104269
- 20
- 21
- 22
- 23
- 24
- 25
- 26
- 27
- 28
- 29
- 30
- 31
- 32
- 33
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45
- 46
- 47
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60
- 61
- 62
- 63
- 64
- 65



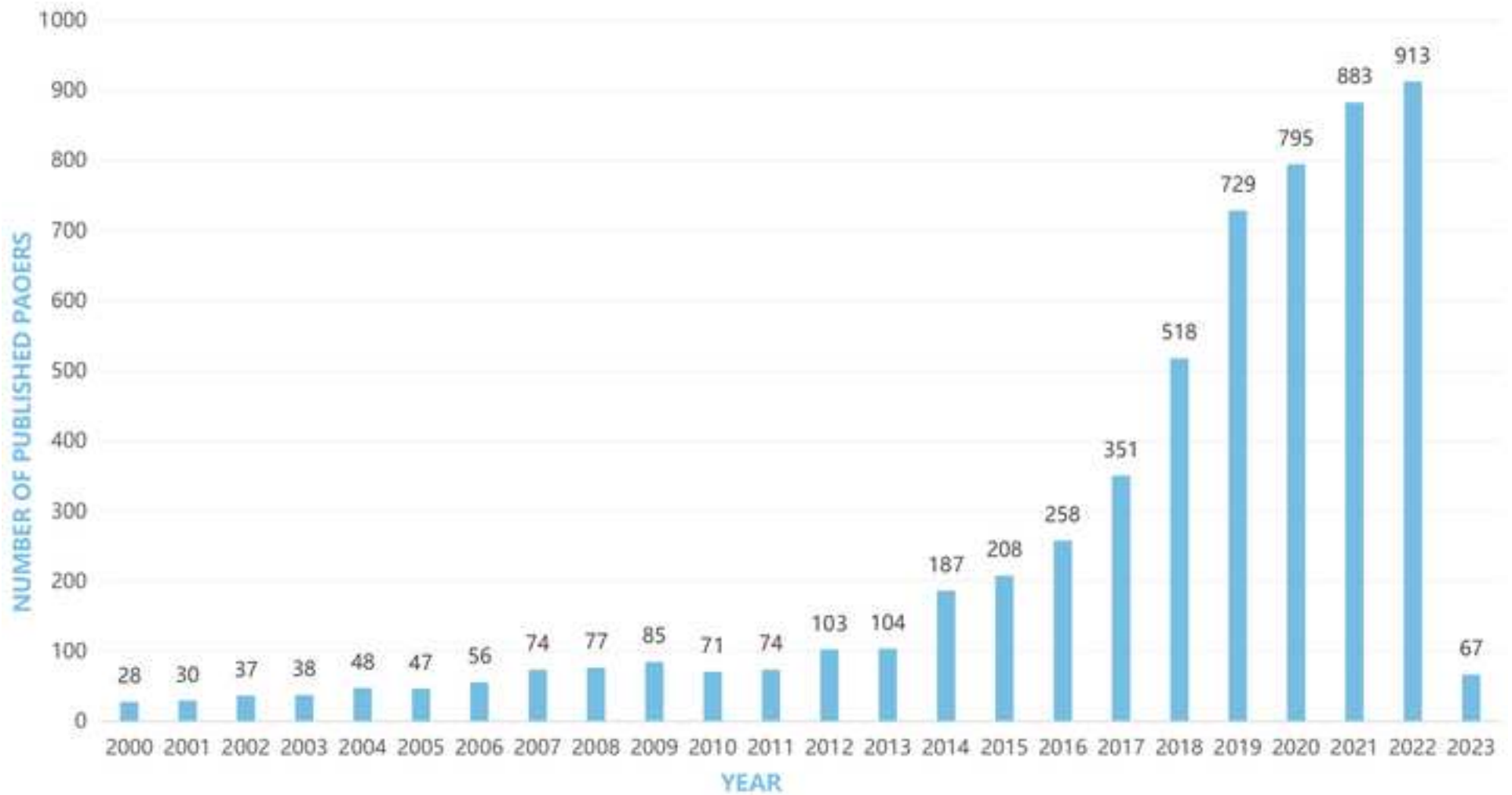


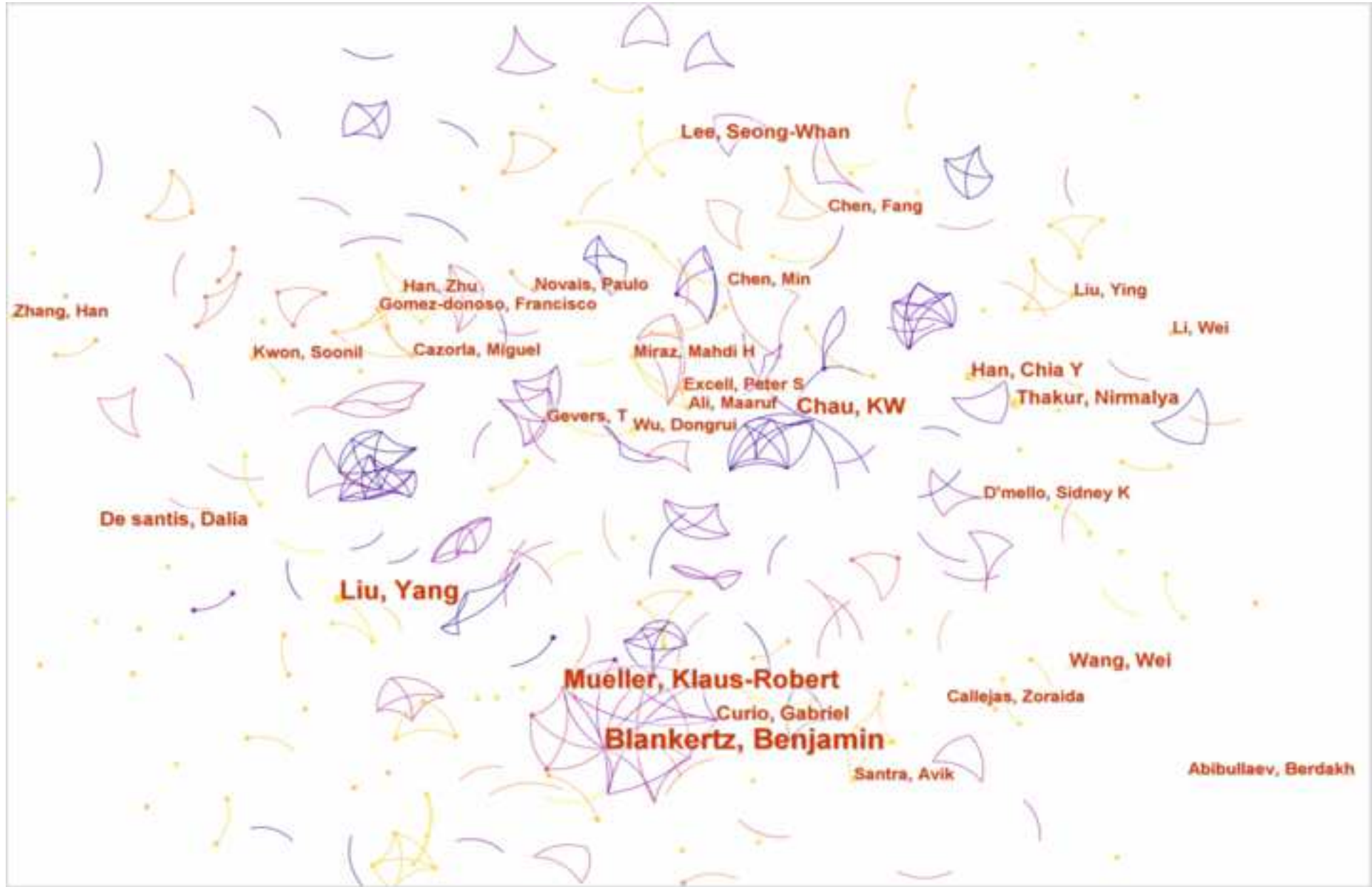


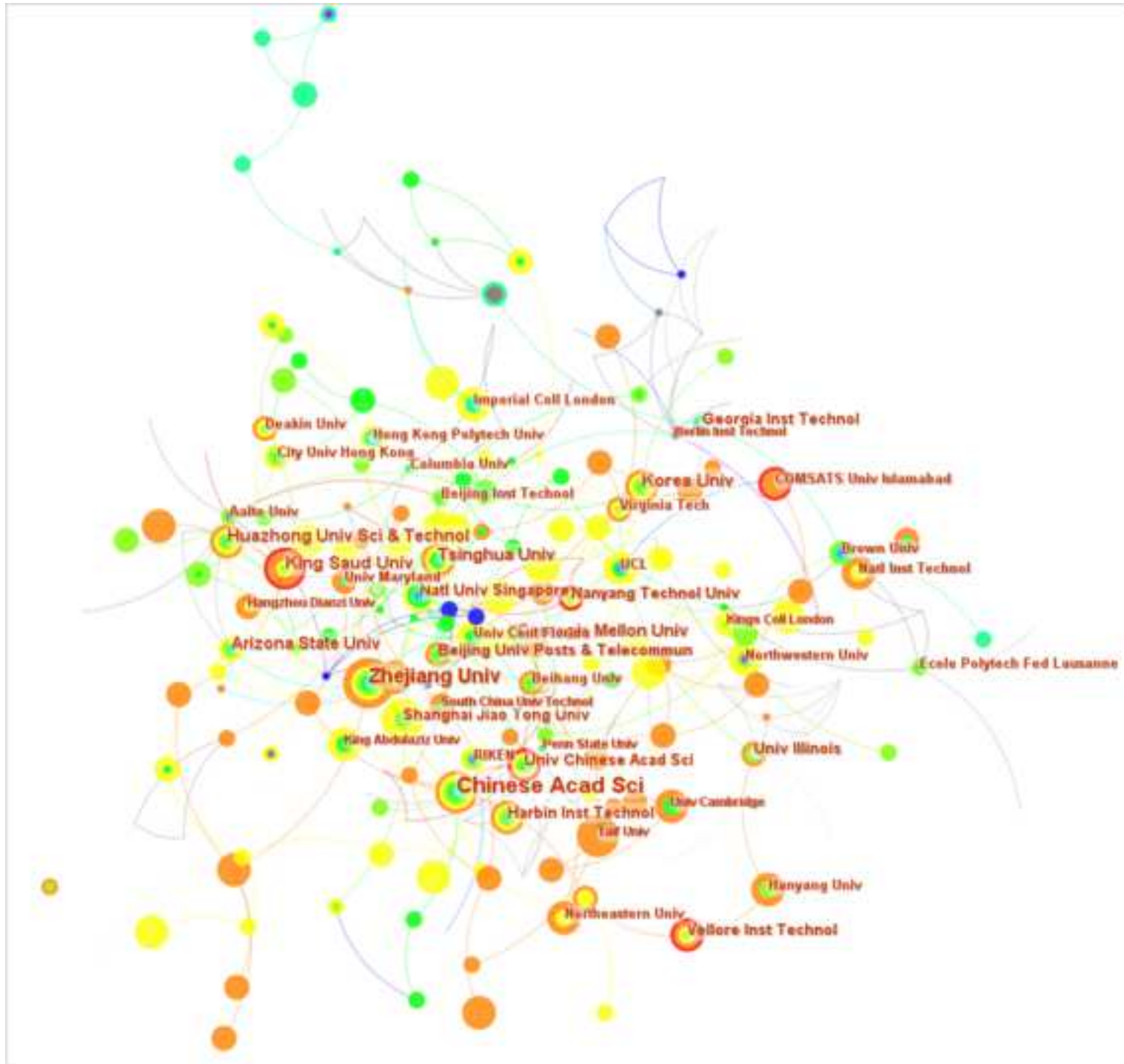




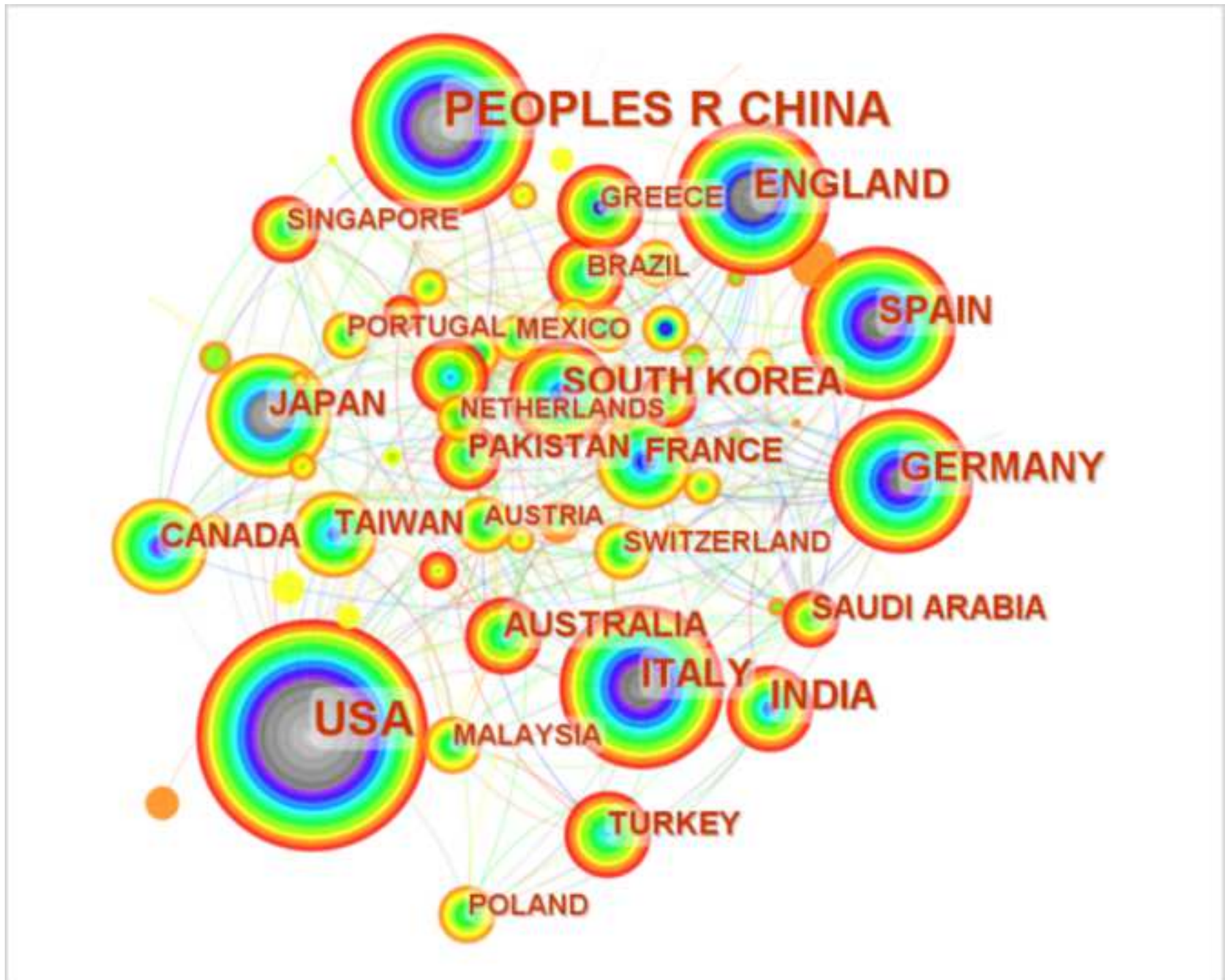


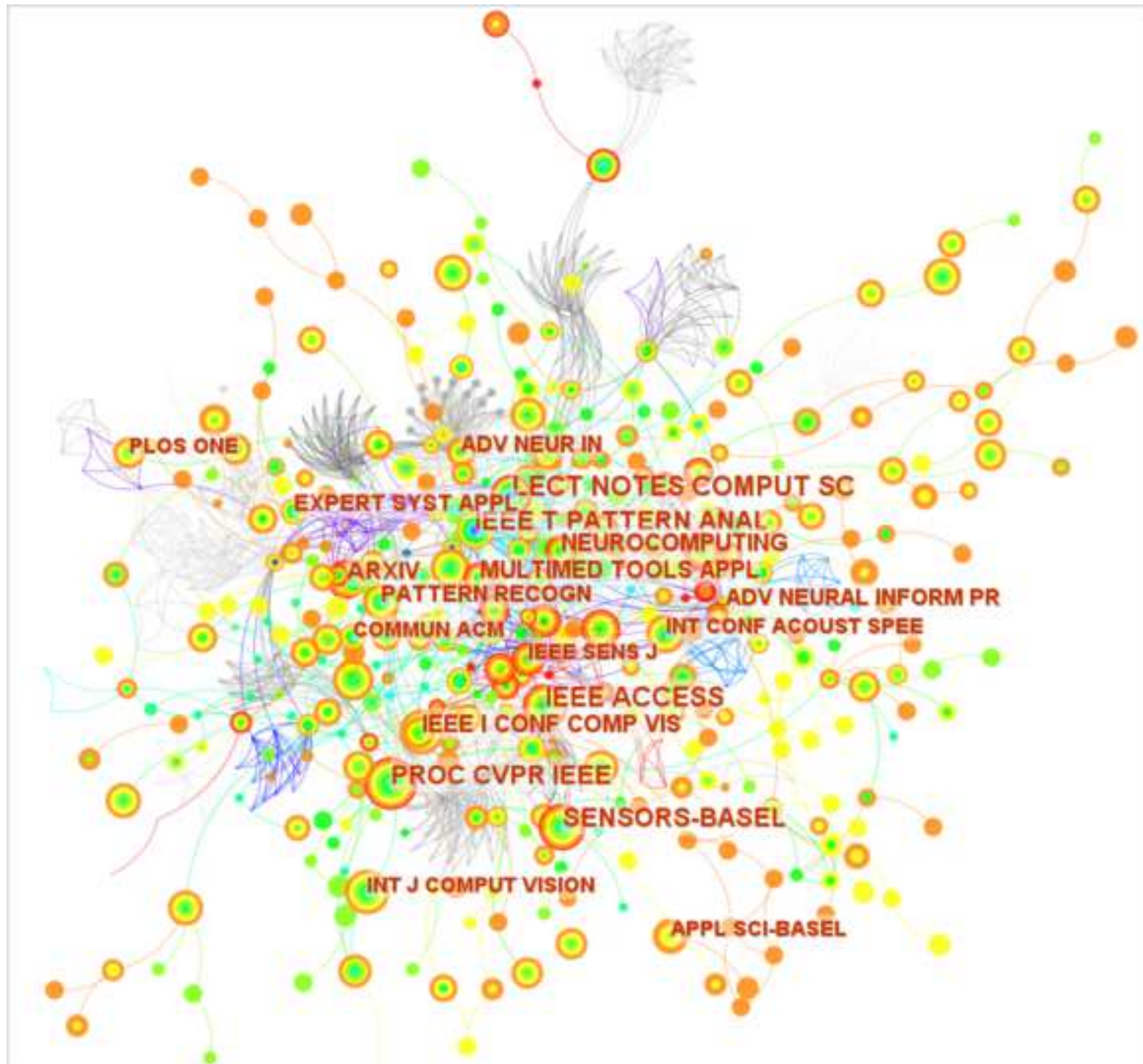


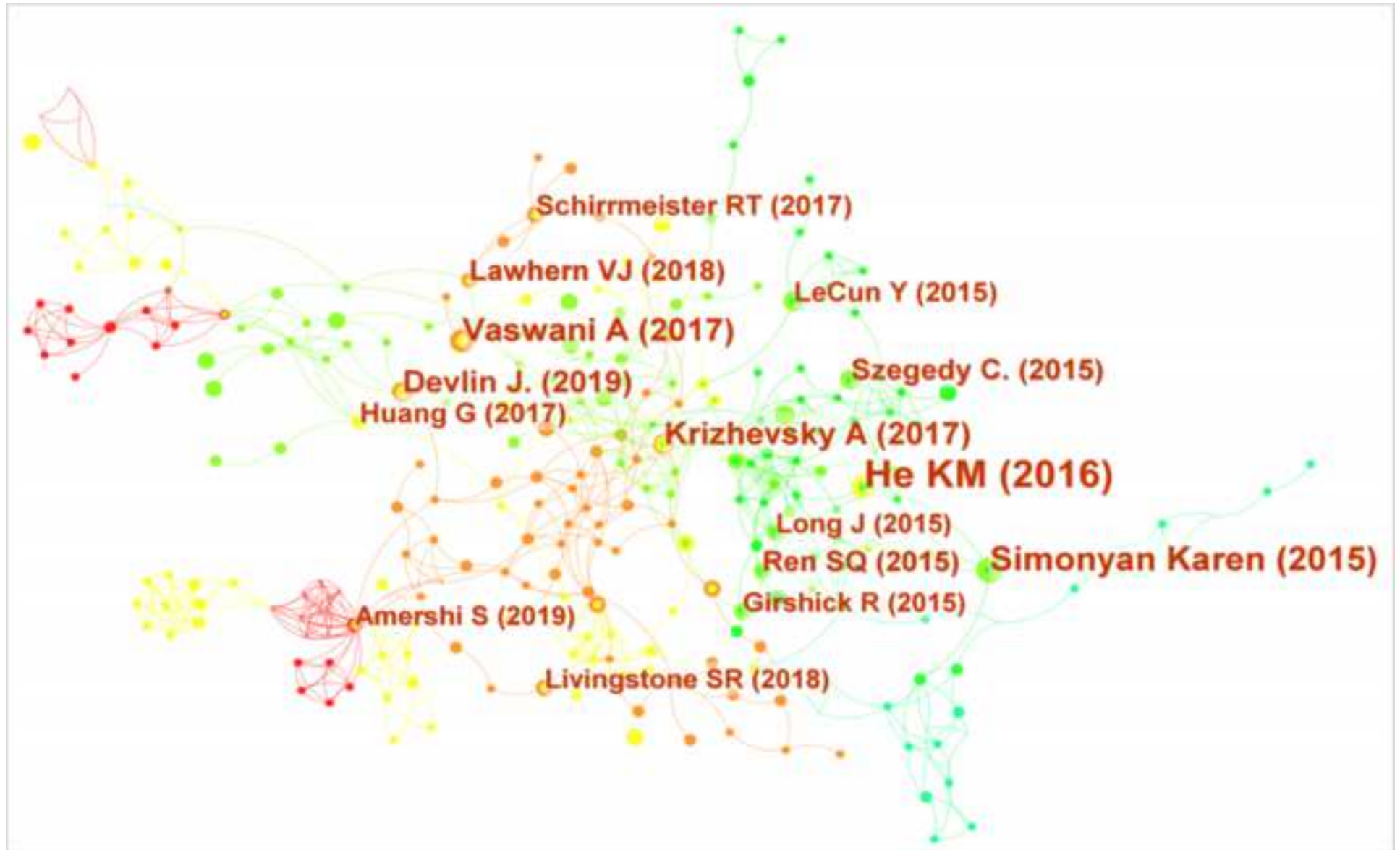






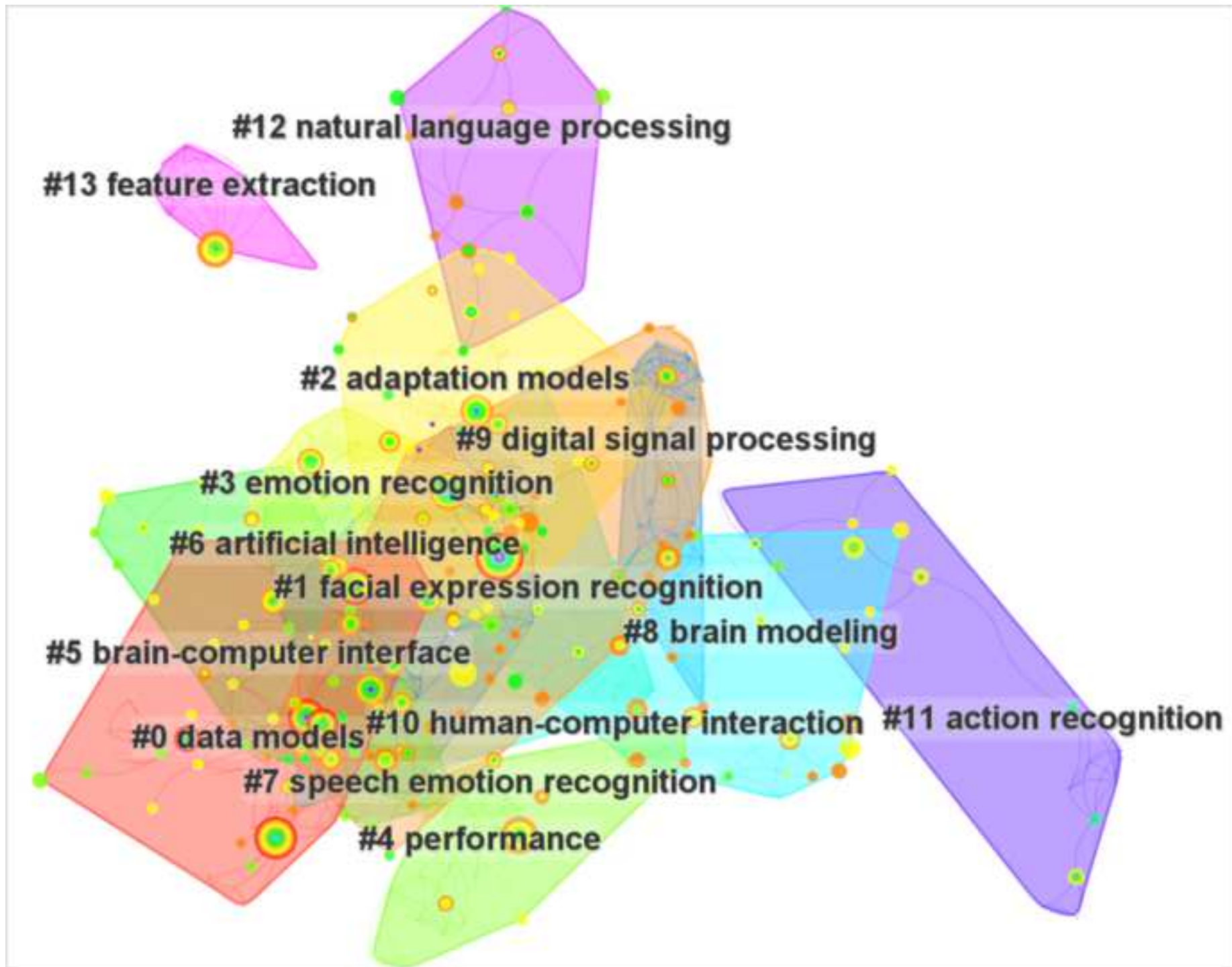




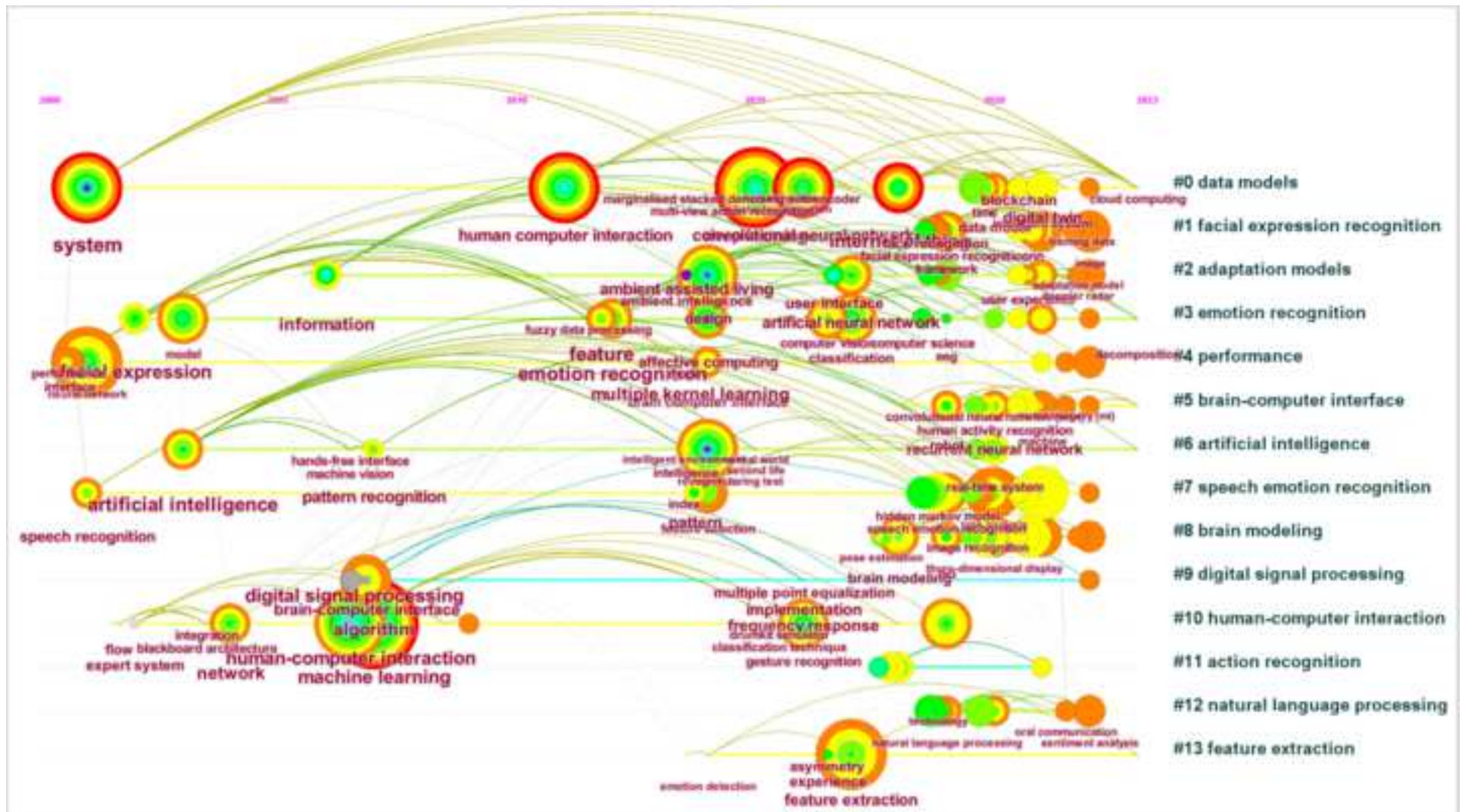


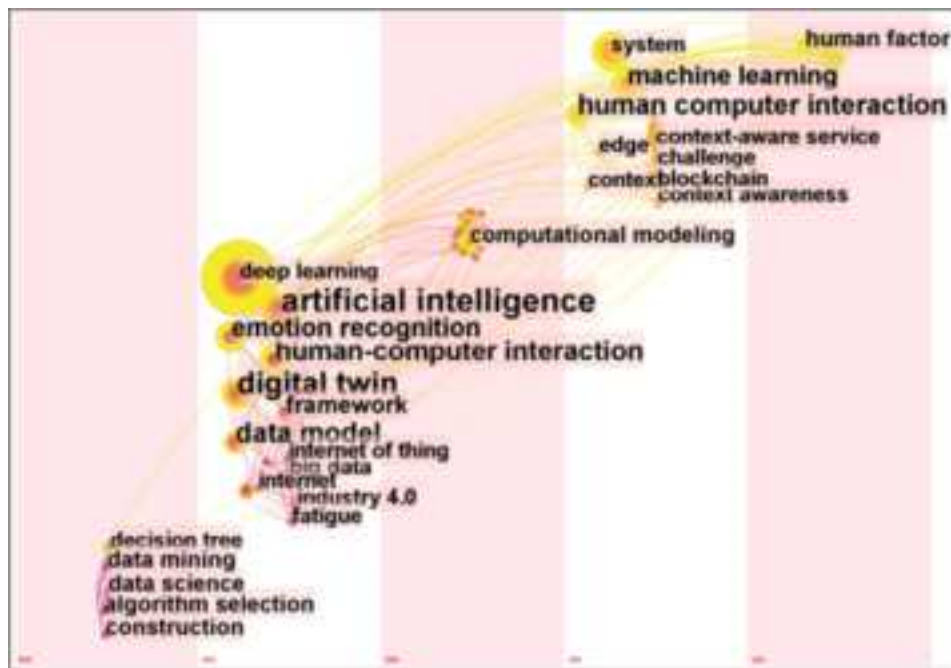




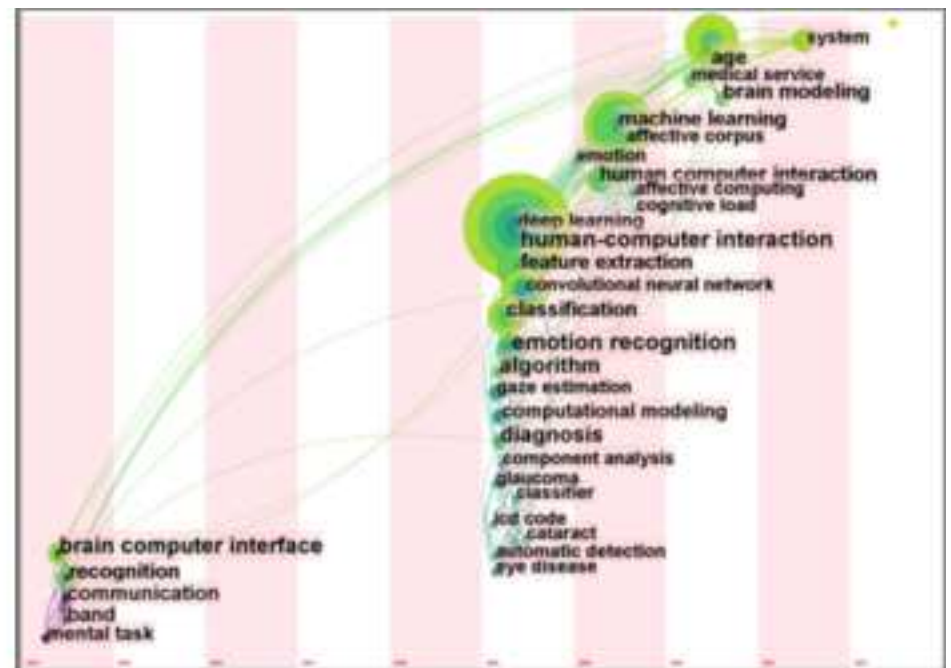




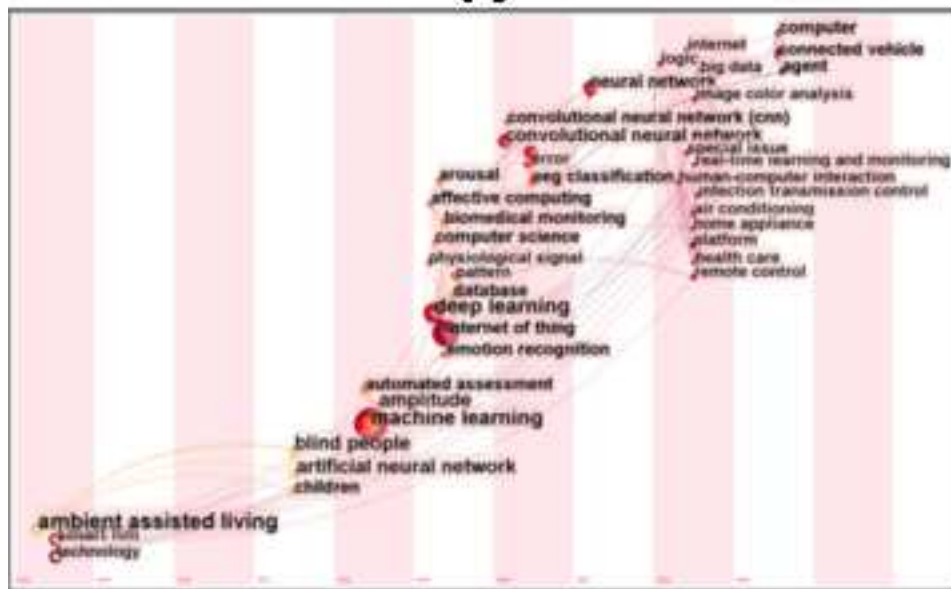




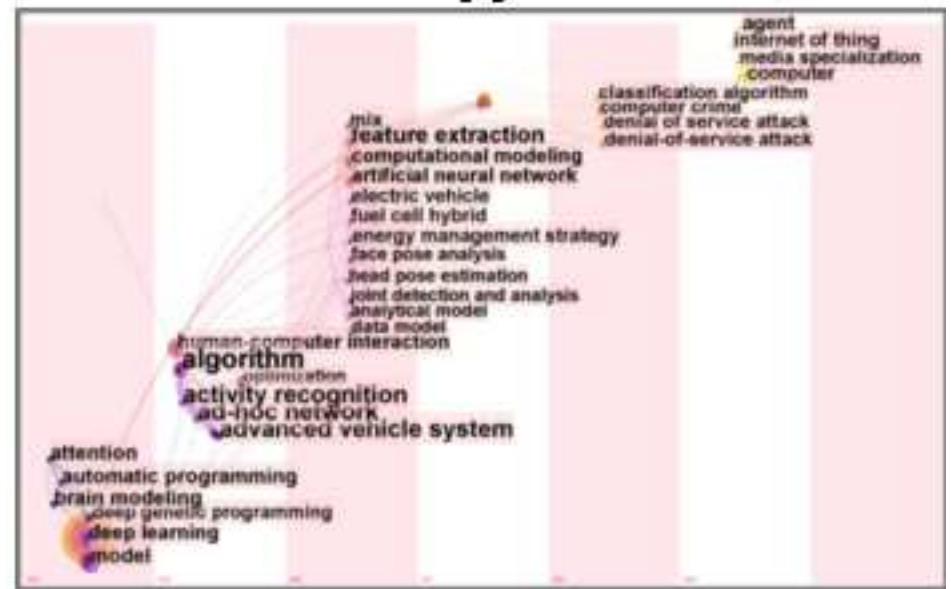
(a)



(b)



(c)



(d)