

Machine Learning Research Trends in Traditional Chinese Medicine: A Bibliometric Review

Jiekee Lim ¹, Jieyun Li¹, Mi Zhou¹, Xinang Xiao¹, Zhaoxia Xu^{1,2}

¹School of Traditional Chinese Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, People's Republic of China; ²Shanghai Key Laboratory of Health Identification and Assessment, Shanghai, People's Republic of China

Correspondence: Zhaoxia Xu, School of Traditional Chinese Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, 201203, People's Republic of China, Email zhaoxia7001@shutcm.edu.cn

Background: Integrating Traditional Chinese Medicine (TCM) knowledge with modern technology, especially machine learning (ML), has shown immense potential in enhancing TCM diagnostics and treatment. This study aims to systematically review and analyze the trends and developments in ML applications in TCM through a bibliometric analysis.

Methods: Data for this study were sourced from the Web of Science Core Collection. Data were analyzed and visualized using Microsoft Office Excel, Bibliometrix, and VOSviewer.

Results: 474 documents were identified. The analysis revealed a significant increase in research output from 2000 to 2023, with China leading in both the number of publications and research impact. Key research institutions include the Shanghai University of Traditional Chinese Medicine and the China Academy of Chinese Medical Sciences. Major research hotspots identified include ML applications in TCM diagnosis, network pharmacology, and tongue diagnosis. Additionally, chemometrics with ML are highlighted for their roles in quality control and authentication of TCM products.

Conclusion: This study provides a comprehensive overview of ML applications' development trends and research landscape in TCM. The integration of ML has led to significant advancements in TCM diagnostics, personalized medicine, and quality control, paving the way for the modernization and internationalization of TCM practices. Future research should focus on improving model interpretability, fostering international collaborations, and standardized reporting protocols.

Keywords: artificial intelligence, bibliometric, machine learning, review, traditional Chinese medicine

Introduction

Traditional Chinese Medicine (TCM), with its unique theoretical system and treatment methods, has a rich history spanning thousands of years. However, despite its profound heritage, TCM faces significant challenges in modern medical research due to its inherent complexity and the qualitative nature of its diagnostic and therapeutic systems. One major obstacle is the gap between TCM's holistic and often subjective framework and the data-driven, quantitative nature of modern scientific methodologies, including machine learning (ML). Therefore, integrating traditional TCM knowledge with modern technology is a crucial research area in contemporary TCM studies.

Machine learning (ML) is a field within artificial intelligence (AI) that develops algorithms that enable computers to learn and improve from experience and data without explicit programming.^{1,2} ML has shown immense potential in the medical field in recent years.³ ML algorithms can process and analyze vast amounts of data, uncovering hidden patterns and associations within the data. This capability is particularly valuable for TCM areas, such as disease diagnosis, syndrome differentiation, and Chinese herbal medicine analysis. For instance, Chen and He highlighted the importance of ML in TCM, showing how these algorithms can enhance research efficiency and accuracy by extracting valuable insights from large datasets.⁴

Applying ML to TCM research has led to significant breakthroughs. For example, a study utilized ML techniques to analyze TCM's four diagnostic methods, successfully predicting the progression of chronic diseases.⁵ The study trained algorithms capable of recognizing disease patterns by analyzing extensive patient diagnostic data, providing a new

diagnostic tool for TCM.⁵ Analyzing large-scale clinical data through ML enhances our understanding of TCM's mechanisms of action and elucidates the differential responses of various constitutions to treatments, thus advancing personalized medicine.

Bibliometric analysis, a scientific method involving the quantitative assessment of literature data,⁶ is particularly important for understanding ML research in TCM. This method helps researchers identify research hotspots and trends in the TCM-ML field. Visualization tools such as network maps and trend graphs provide a clear depiction of research structures and evolution paths. Through this bibliometric analysis, we aim to systematically comprehend the development trends and research landscape in this field, offering scientific data support and decision-making insights for the modernization and internationalization of TCM.

The findings demonstrated that the introduction of ML techniques provides new perspectives and methods for TCM research, making traditional TCM diagnosis and pharmacological studies more precise and efficient. Additionally, bibliometric analysis allows for a better understanding of the developmental trajectory and research framework of this field, providing valuable data support and strategic insights for the modernization and global integration of TCM.

Materials and Methods

Data Collection

Data for this study was sourced from the Web of Science Core Collection (WoSCC), focusing on document types limited to "articles" and "review articles". Our search strategy utilized "Topic" search terms related to ML and Traditional Chinese Medicine (TCM). The search terms including: "machine learning", "artificial intelligence", "deep learning", "artificial neural network*", "computational intelligence", "deep neural network*", "traditional chinese medicine*", "chinese medicine*", "chinese traditional medicine*", "chinese herbal medicine*". Detailed search strategies are documented in [Table S1](#). Two researchers independently conducted the literature search to ensure thoroughness and minimize selection bias. All electronic searches were completed on May 8th, 2024. To maintain data integrity, publications from the year 2024 were excluded from the analysis, as data for this period were not yet finalized.

Data Screening and Quality Control

The literature screening process was conducted by two researchers (J. Lim and M. Zhou) to ensure accuracy and reliability. Studies were included if they focused on ML applications in TCM and were published in peer-reviewed journals. Studies were excluded if they did not contain a clear ML component or were unrelated to TCM. Each researcher independently reviewed the collected literature and cross-checked their findings. In cases where discrepancies arose between the two researchers, discussions were held to resolve the differences. If consensus could not be reached through discussion, a third researcher (J. Li) was brought in to adjudicate and provide a final decision. This rigorous three-step process was implemented to enhance the precision of literature selection and to maintain high-quality standards in our research.

Data Analysis

Data analysis and visualization in this study were conducted using Microsoft Office Excel 365, Bibliometrix,⁷ and VOSviewer 1.6.20.⁸ Specifically, Microsoft Office Excel 365 was employed for data collation and preliminary analysis. VOSviewer 1.6.20 was utilized to investigate co-cited literature. Bibliometrix facilitated a comprehensive analysis of publication trends, identification of the most cited papers, historiographic mapping, keyword analysis, and thematic analysis.

Results

General Publication Trends

Publications by Year

Out of 502 papers identified from the WoSCC database, 28 were excluded, and 474 were included in our database. [Figure 1](#) illustrates the publication trend from 2000 to 2023, highlighting a gradual increase in the number of articles,

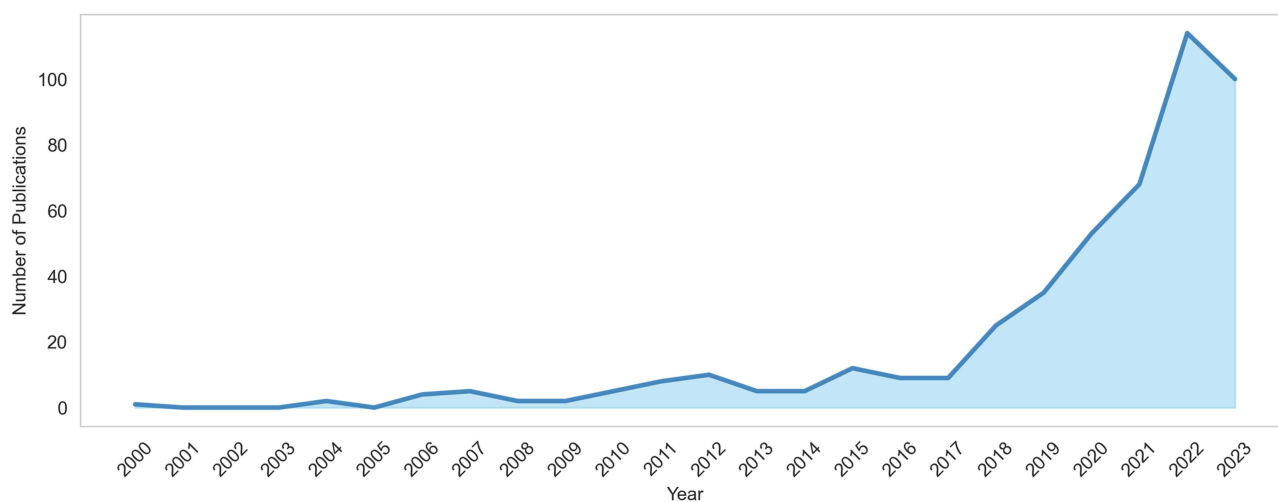


Figure 1 Annual number of publications.

indicating growing interest in the field. Initially, there was only one publication in 2000, with minimal output observed until 2006. Starting in 2018, research output significantly surged, with the number of articles doubling from 25 in 2018 to 53 in 2020, and peaking at 114 in 2022. Although there was a slight decline to 100 publications in 2023, the overall trend demonstrates an expanding research landscape in applying ML techniques within the TCM field. This growth pattern is characterized by initial slow growth followed by a recent rapid escalation.

Publications by Country

Publications involved authors from a total of 28 countries, and among those, the corresponding authors were from 16 countries. [Table 1](#) shows the number of publications and total citations for the corresponding authors' countries (with at least two publications). China leads by a significant margin with 441 papers and 5435 total citations, indicating a high level of research activity and impact. Other countries with contributions include India, Korea, the United Kingdom, Singapore, and the USA.

[Figure 2](#) highlights the research collaboration between countries, with China being the most collaborative, partnering with 24 countries. The highest collaboration frequency was between China and the USA, counting 23 instances. This is followed by China and Australia with 17 collaborations.

Publications by Affiliation

A total of 419 affiliations were identified across the publications. [Figure 3](#) lists the top 15 affiliations ranked by the number of articles published. The figure highlights the institutions with the highest output of publications in the field, with Shanghai University of Traditional Chinese Medicine leading with 66 articles, indicating its prominence as a research center in this domain. This is followed by the China Academy of Chinese Medical Sciences with 43 publications, and China Medical University Taiwan with 40 publications, suggesting their key roles in the research

Table 1 Number of Publications and Citations by Corresponding Author's Country

Country	Papers	Total Citations
CHINA	441	5435
INDIA	6	54
KOREA	5	86
UNITED KINGDOM	4	33
SINGAPORE	3	102
USA	3	52



Figure 2 Research collaboration between countries.

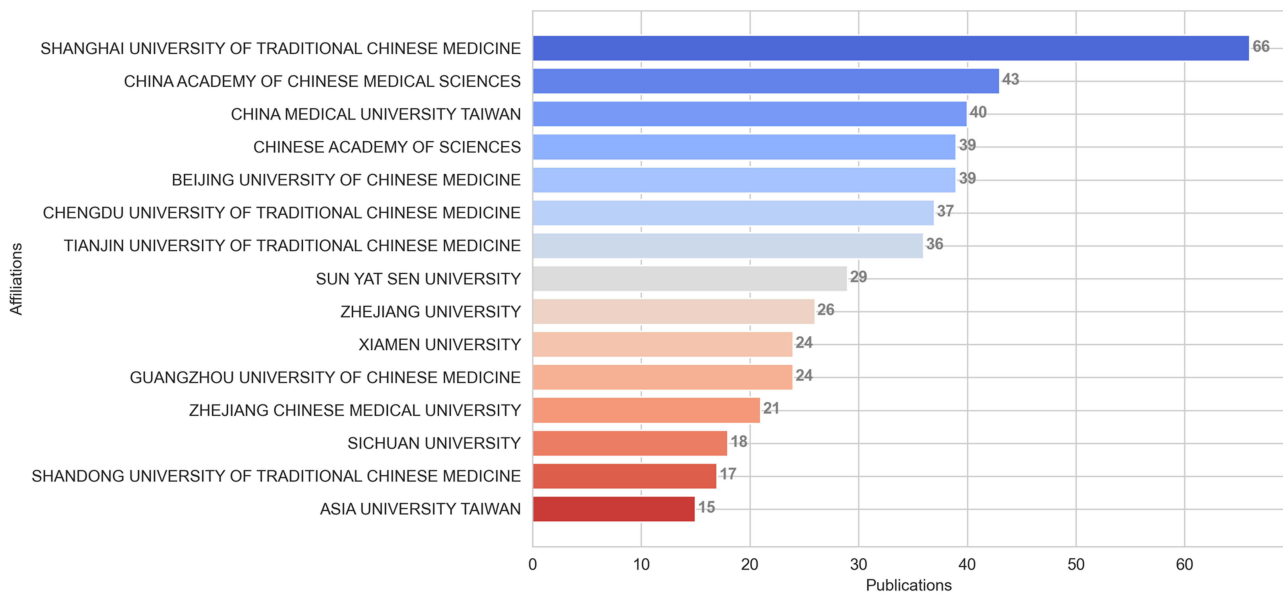


Figure 3 The top 15 affiliations by number of publications.

community. The results reveal a significant concentration of scholarly output within Chinese universities and institutes, particularly in the field of Traditional Chinese Medicine.

Publications by Journal

Research publications spread across 226 journals, with 17 journals having 6 or more publications (Table 2). “Evidence-Based Complementary and Alternative Medicine” tops the list with 27 publications and with the highest H-index (9), followed by “IEEE Access” (n = 20, H-index = 7) and “Frontiers in Pharmacology” (n = 15, H-index = 8). The H-index (Hirsch index) measures a journal’s productivity and citation impact by indicating the number of published articles that have been cited at least H times.

Table 2 Top Journals by Number of Publications (n ≥ 6)

Rank	Sources	Articles	H_index	IF 2022	CiteScore 2022
1	EVIDENCE-BASED COMPLEMENTARY AND ALTERNATIVE MEDICINE	27	9	0	3.5
2	IEEE ACCESS	20	7	3.9	9
3	FRONTIERS IN PHARMACOLOGY	15	8	5.6	6.3
4	SPECTROSCOPY AND SPECTRAL ANALYSIS	12	4	0.7	1.4
5	BIOMED RESEARCH INTERNATIONAL	9	5	0	5.3
6	CHINESE JOURNAL OF INTEGRATIVE MEDICINE	9	5	2.9	5.3
7	CHINESE MEDICINE	9	5	4.9	6.9
8	SPECTROCHIMICA ACTA PART A-MOLECULAR AND BIOMOLECULAR SPECTROSCOPY	9	8	4.4	7.9
9	JOURNAL OF ETHNOPHARMACOLOGY	8	6	5.4	8.6
10	JMIR MEDICAL INFORMATICS	7	5	3.2	5.6
11	MOLECULES	7	3	4.6	6.7
12	PHARMACOLOGICAL RESEARCH	7	6	9.3	15.8
13	ARTIFICIAL INTELLIGENCE IN MEDICINE	6	6	7.5	14
14	COMPUTATIONAL AND MATHEMATICAL METHODS IN MEDICINE	6	3	0	2.3
15	CONCURRENCY AND COMPUTATION-PRACTICE & EXPERIENCE	6	6	2	4.2
16	JOURNAL OF PHARMACEUTICAL AND BIOMEDICAL ANALYSIS	6	5	3.4	6.6
17	PLOS ONE	6	4	3.7	6

Notes: “Evidence-based Complementary and Alternative Medicine”, “Biomed Research International”, and “Computational and Mathematical Methods in Medicine” were removed from the SCIE catalog on March 20, 2023.

Abbreviation: IF, impact factor.

Among the listed journals, it is notable that “Pharmacological Research” holds the highest impact factor (IF) of 9.3 and the highest CiteScore of 15.8. “Artificial Intelligence in Medicine” (IF = 7.5, CiteScore = 14) is also a notable journal.

The diversity of journals reflects a multidisciplinary approach within the field, [Figure 4](#) shows a comparison of journal metrics, including the IF, CiteScore for 2022, and the H-index.

Citation Analysis

Most Cited Papers

The top cited works mainly reflect the growing recognition of ML in drug discovery from natural products of TCM. The top 10 most cited papers are listed in [Table 3](#). Leading the list, the review paper “Natural Products for Drug Discovery in the 21st Century: Innovations for Novel Drug Discovery” by Thomford NE,⁹ published in the International Journal of Molecular Sciences in 2018, tops with 627 citations. The article discussed the importance of natural products in drug discovery and the integration of innovative technologies, for instance, the discovery of the antimalarial drug Artemisinin (*Artemisia annua*) in traditional Chinese herbal medicine and the utilization of AI in modern drug discovery.

Overall, the analysis reveals a strong trend toward the discovery of drugs from natural products, in silico predictions, and analytical evaluations of herbal medicines within TCM research intersecting with ML. This interdisciplinary approach is increasingly acknowledged and cited within the scientific community, underscoring the potential of integrating traditional knowledge with contemporary pharmaceutical research.

Historiographic Mapping

[Figure 5](#) presents a historiographic map of the top 10 co-cited papers, with the number of citations detailed in [Table 4](#). This map illustrates the relationships between papers, showing how more recent publications may have cited earlier ones. This suggests a level of interconnectivity and highlights the influence of prior research on subsequent studies. A historical pathway, revealing a research topic and its key documents, was identified, highlighting one significant development path within the field.

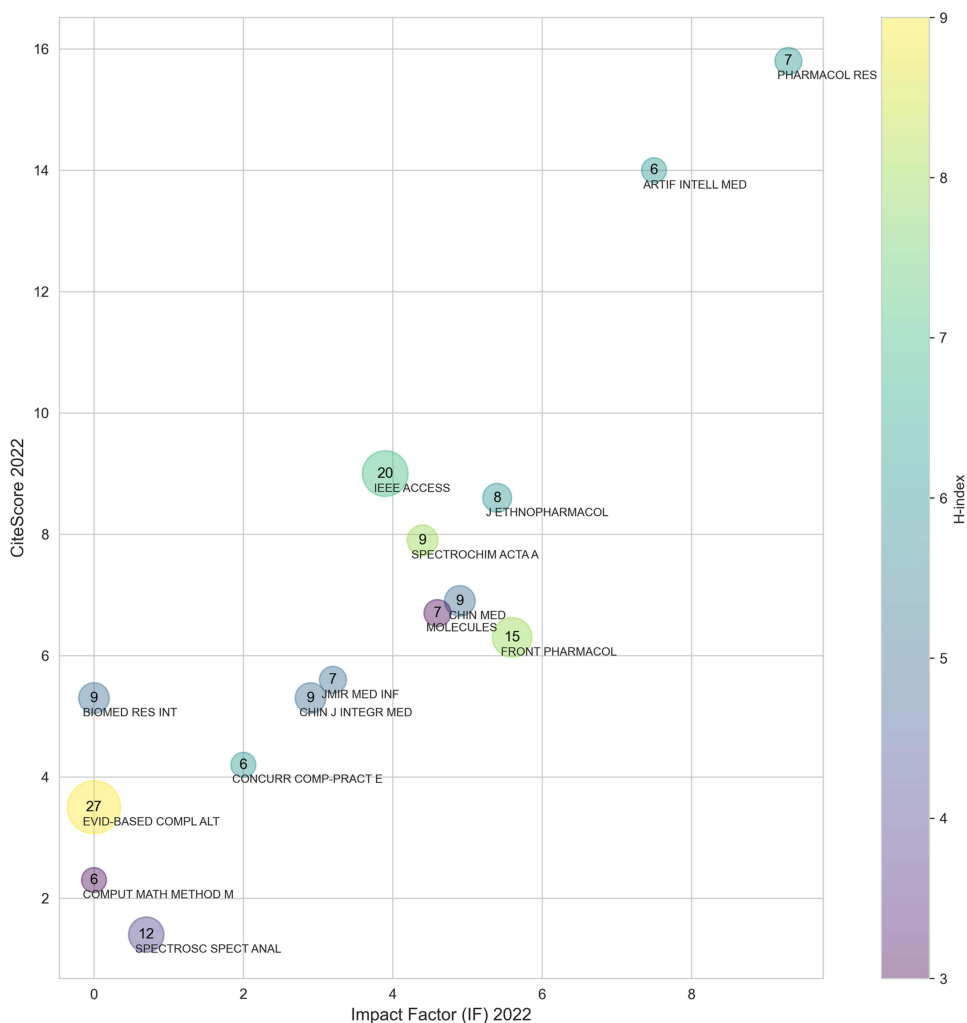


Figure 4 Top journals metrics comparison. The sizes of the bubbles represent the number of publications, with larger bubbles indicating more publications. The colors of the bubbles represent the H-index, with the color gradient ranging from purple (lower H-index) to yellow-green (higher H-index).

The network reveals a progression from early influential work, such as Zhang NL (2008)²⁰ introducing latent tree models for TCM diagnosis and Zhou XZ (2010)¹⁵ conducting a comprehensive survey on text mining for TCM knowledge discovery, to significant advancements in ML applications by Liu GP (2014)²¹ and Zhao CB (2015)¹⁷ for

Table 3 The Top 10 Cited Papers

Rank	Author	Year	Journal	Title	TC	TC/Year
1	Thomford NE ⁴	2018	International Journal of Molecular Sciences	Natural Products for Drug Discovery in the 21st Century: Innovations for Novel Drug Discovery	627	89.57
2	Tian S ⁵	2015	Advanced Drug Delivery Reviews	The Application of In Silico Drug-Likeness Predictions in Pharmaceutical Research	272	27.20
3	Tian RT ⁶	2009	Journal of Chromatography A	Evaluation of Traditional Chinese Herbal medicine: Chaihu (Bupleuri Radix) by Both High-Performance Liquid Chromatographic and High-Performance Thin-Layer Chromatographic Fingerprint and Chemometric Analysis	144	9.00
4	Xu Q ⁷	2020	IEEE Journal of Biomedical and Health Informatics	Multi-Task Joint Learning Model for Segmenting and Classifying Tongue Images Using a Deep Neural Network	121	24.20

(Continued)

Table 3 (Continued).

Rank	Author	Year	Journal	Title	TC	TC/Year
5	Zhu YY ⁸	2022	Acta Pharmaceutica Sinica B	New Opportunities and Challenges of Natural Products Research: When Target Identification Meets Single-Cell Multiomics	116	38.67
6	Tian S ⁹	2012	Molecular Pharmaceutics	Drug-likeness Analysis of Traditional Chinese Medicines: Prediction of Drug-likeness Using Machine Learning Approaches	89	6.85
7	Zhou XZ ¹⁰	2010	Journal of Biomedical Informatics	Text Mining for Traditional Chinese Medical Knowledge Discovery: A Survey	77	5.13
8	Yang M ¹¹	2013	Evidence-Based Complementary and Alternative Medicine	Navigating Traditional Chinese Medicine Network Pharmacology and Computational Tools	74	6.17
9	Ung CY ¹²	2007	Journal of Ethnopharmacology	Are Herb-Pairs of Traditional Chinese Medicine Distinguishable From Others? Pattern Analysis and Artificial Intelligence Classification Study of Traditionally Defined Herbal Properties	73	4.06
10	CHANG SS ¹³	2011	PLOS Computational Biology	Two Birds with One Stone? Possible Dual-Targeting H1N1 Inhibitors from Traditional Chinese Medicine	70	5.00

Abbreviation: TC, Total Citations.

syndrome diagnosis and patient classification. Notable contributions from Xu Q (2019, 2020)^{12,22} include studies on syndrome differentiation and tongue image analysis using neural networks. Recent developments by Wang X (2020)²³ and Zhang H (2020)²⁴ highlight the use of deep convolutional neural networks (CNN) for tongue diagnosis and AI-based

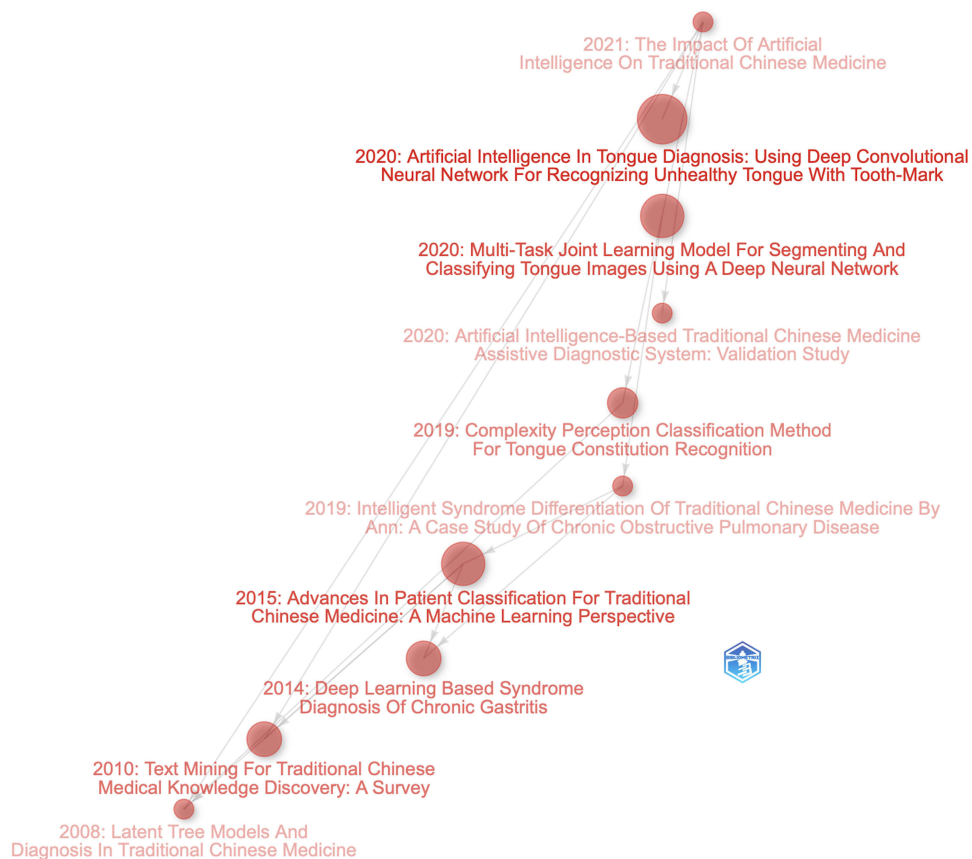


Figure 5 Historiographic map of the top 10 co-cited papers. The titles of the research papers are ordered by the year of publication. The lines connecting the nodes indicate that later papers have cited earlier ones, demonstrated interconnectivity, and highlighted the influence of preceding research on subsequent studies. The title darkness corresponds to the number of local citations, with darker titles indicating higher local citation counts as detailed in Table 4.

Table 4 Paper Citations of Historiographic Mapping

Author	Year	Journal	Title	Local citations	Global citations
Zhang NL ¹⁴	2008	Artificial Intelligence in Medicine	Latent Tree Models and Diagnosis in Traditional Chinese Medicine	10	66
Zhou XZ ¹⁰	2010	Journal of Biomedical Informatics	Text Mining for Traditional Chinese Medical Knowledge Discovery: A Survey	12	77
Liu GP ¹⁵	2014	Computational and Mathematical Methods in Medicine	Deep Learning Based Syndrome Diagnosis of Chronic Gastritis	12	26
Zhao CB ¹⁶	2015	Evidence-Based Complementary and Alternative Medicine	Advances in Patient Classification for Traditional Chinese Medicine: A Machine Learning Perspective	15	40
Ma JJ ¹⁷	2019	Artificial Intelligence in Medicine	Complexity Perception Classification Method for Tongue Constitution Recognition	11	29
Xu Q ¹⁸	2019	IEEE ACCESS	Intelligent Syndrome Differentiation of Traditional Chinese Medicine by ANN: A Case Study of Chronic Obstructive Pulmonary Disease	10	23
Xu Q ⁷	2020	IEEE Journal of Biomedical and Health Informatics	Multi-task Joint Learning Model for Segmenting and Classifying Tongue Images Using a Deep Neural Network	15	121
Wang X ¹⁹	2020	Computational and Structural Biotechnology Journal	Artificial Intelligence in Tongue Diagnosis: Using Deep Convolutional Neural Network for Recognizing Unhealthy Tongue with Tooth-Mark	18	40
Zhang H ²⁰	2020	JMIR Medical Informatics	Artificial Intelligence-Based Traditional Chinese Medicine Assistive Diagnostic System: Validation Study	10	30
Wang YL ²¹	2021	The American Journal of Chinese Medicine	The Impact of Artificial Intelligence on Traditional Chinese Medicine	10	18

Notes: Local citations refer to the number of times an article has been cited by other articles within the analyzed dataset; Global citations refer to the total number of citations an article has received according to the WoS database.

assistive diagnostic systems, while Wang YL (2021)²⁵ discusses the broader impact and challenges of AI integration in TCM.

The overall progression of the network path reveals the evolution and interconnectedness of deep learning and Artificial Neural Networks (ANN) in TCM diagnosis, particularly in syndrome or disease diagnosis and tongue image analysis.

Co-Cited References

In the analysis of co-cited references, the inclusion criterion was set at a minimum of 15 citations per reference. Out of a total of 20,806 cited references, only 15 met this threshold. The analysis identified three main clusters within the co-citation network, color-coded red, blue, and green, as depicted in Figure 6. The red cluster emphasized machine learning and TCM pharmacological databases, the green cluster focused on deep learning and image recognition, and the blue cluster addressed syndrome differentiation and patient classification in TCM. The specific references for each cluster can be found in Table S2, and its reference list.

The article receiving the highest number of citations (36) within this dataset is “Deep Residual Learning for Image Recognition” by He KM,²⁶ published in 2016. Following closely, the second most-cited paper, with 23 citations, is “Very Deep Convolutional Networks for Large-Scale Image Recognition” by Simonyan K,²⁷ from 2015.

Keywords Analysis

Most Frequent Keywords and the Co-Occurrence Network

Figure 7 illustrates the ranking of the top 15 most frequent keywords, which includes terms such as “traditional chinese medicine”, “machine learning”, “deep learning”, “artificial intelligence”, “artificial neural network”, “near-infrared spectroscopy”, “convolutional neural network”, “syndrome differentiation”, “network pharmacology”, “tongue diagnosis”, “feature extraction”, “tongue image”, “support vector machine”, “back propagation neural network”, and “chemometrics”. The

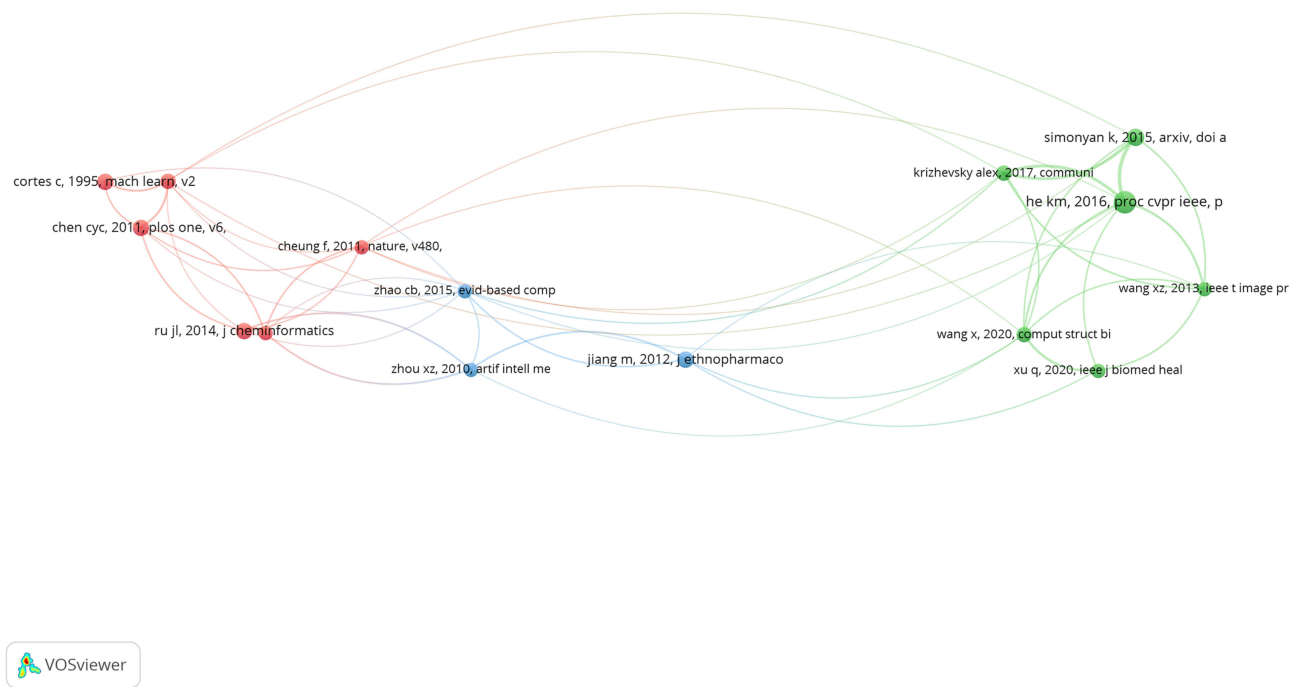


Figure 6 The co-citation network analysis of references. The red cluster represents research on machine learning and TCM pharmacological databases, the green cluster focuses on deep learning and image recognition, and the blue cluster addresses syndrome differentiation and patient classification in TCM.

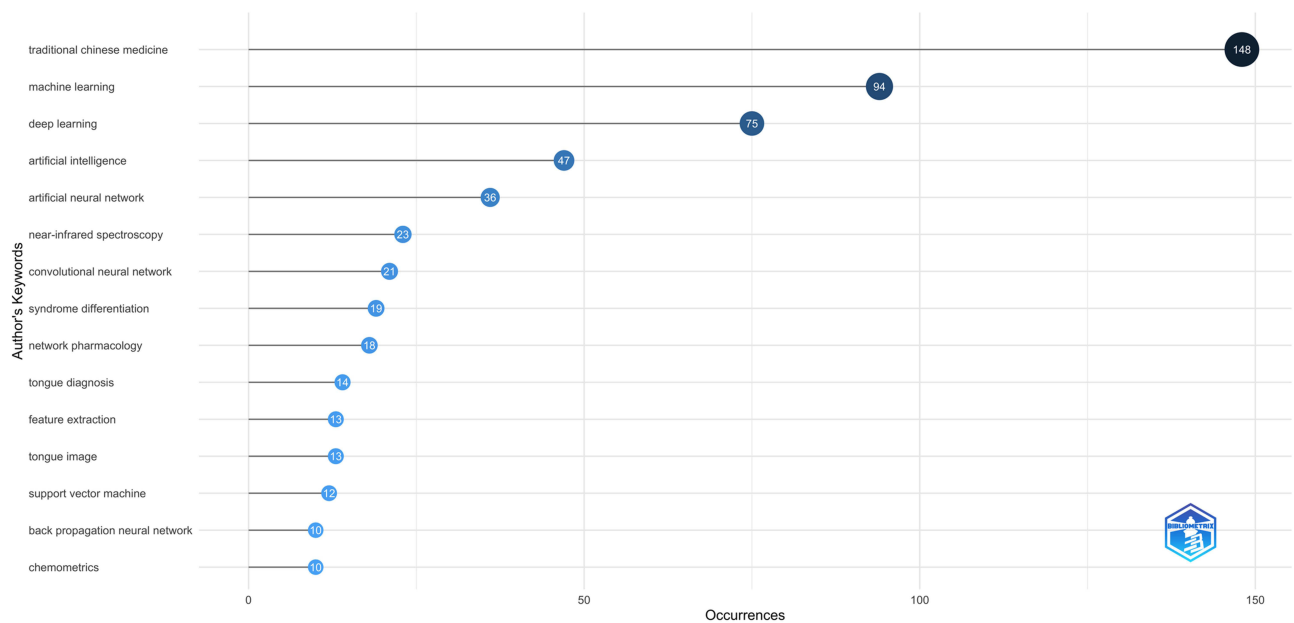


Figure 7 Frequency ranking of top 15 keywords.

prominence of these keywords indicates a significant focus within the analyzed dataset on traditional Chinese medicine, tongue diagnosis, and a range of ML and computational methodologies.

The co-occurrence network illustrated in Figure 8 visualizes the interconnections between keywords. This network reveals the multidisciplinary integration of ML with TCM, highlighting various applications such as identifying TCM syndromes, tongue and pulse diagnosis, feature extraction in tongue image and pulse wave, data mining, drug discovery, medical simulations, herbal analysis, disease diagnosis, and network pharmacology. The network nods to major health

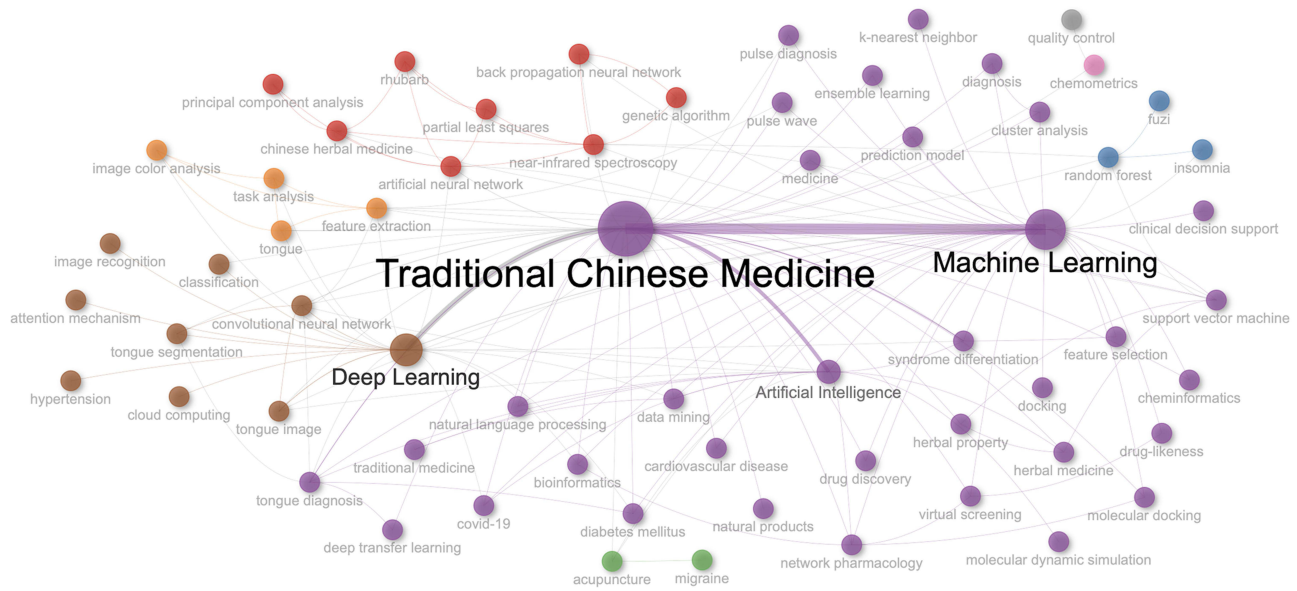


Figure 8 Co-occurrence network of keywords. The prominent nodes represent major research themes, while the connections illustrate how these themes are interrelated.

issues such as migraines, acute infectious diseases like COVID-19, cardiovascular disease, diabetes, and hypertension. It also showcases the objective technical methods of TCM diagnostics, including advanced tongue and pulse analysis with deep learning. ML techniques featured in the network include Support Vector Machines (SVM), neural networks, clustering algorithms, genetic algorithms, natural language processing, random forests, and k-nearest neighbors.

Keywords Trend Topics

Figure 9 displays the analysis of trending topics over time, based on bigrams extracted from research titles. The graph identifies the top three keywords for each year that have surpassed a minimum occurrence frequency of six. The analysis shows a dominance of “Chinese medicine” with 250 occurrences, gaining prominence from 2019 to 2022. “Machine

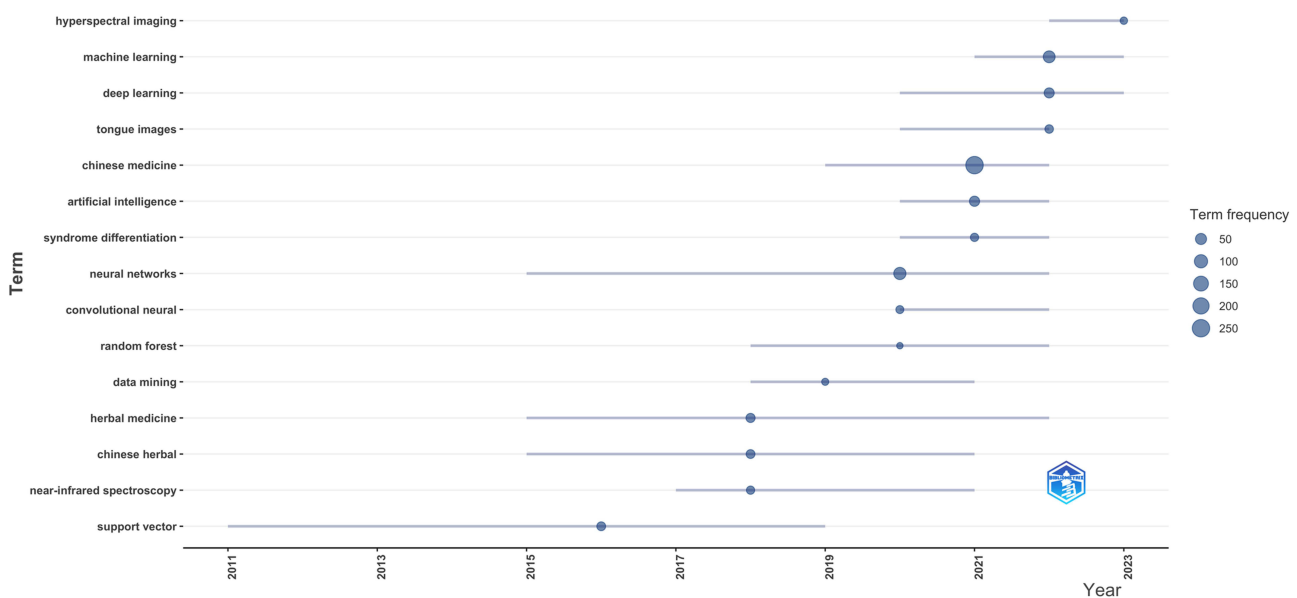


Figure 9 Trend topics of keyword frequencies in research titles. Each bubble on the graph represents a specific keyword topic, with the bubble size being proportional to the number of occurrences of that keyword. The gray bar shows the time span from the first quartile to the third quartile within which the keyword was actively mentioned in research titles.

learning” with 66 occurrences peaks in 2022, reflecting growing research interest in recent years. The second highest occurrence “neural networks” peaks in 2020 with 74 occurrences and is active from 2015 to 2020, paving the way for the continuous rise of “deep learning” from 2020 to 2023, as researchers build on neural network foundations to develop more sophisticated deep learning models. Notably, “support vector”, “herbal medicine”, and “chinese herbal” display longevity in research discussions over extended periods. In recent years, the increasing presence of “artificial intelligence”, “syndrome differentiation”, “tongue images”, and “hyperspectral imaging” indicates a shift toward advanced computational methods in image analysis and diagnosis in TCM research.

Thematic Analysis

The analysis of research themes has identified six clusters categorized into four different quadrants, as represented in Figure 10. The quadrants represent motor, niche, basic, and emerging or declining themes, each with distinct characteristics and implications for the field. The clusters have been grouped into quadrants based on two key dimensions: development degree (density) on the y-axis and relevance degree (centrality) on the x-axis.

Motor themes are the most influential themes with a high degree of development and centrality. This quadrant is composed of two clusters, Cluster 1: the “traditional Chinese medicine” cluster is the most developed and important cluster, highlighting the application of ML in network pharmacology and syndrome differentiation; Cluster 2: the “deep learning” cluster focuses on tongue diagnosis and tongue image extraction using deep learning.

Niche themes have high density but lower centrality, indicating specialized topics that might be important within sub-fields or particular domains. This quadrant includes the Cluster 3, the “chemometrics” cluster, which study the advanced data analysis techniques to enhance the quality control and authentication of TCM pharmaceuticals.

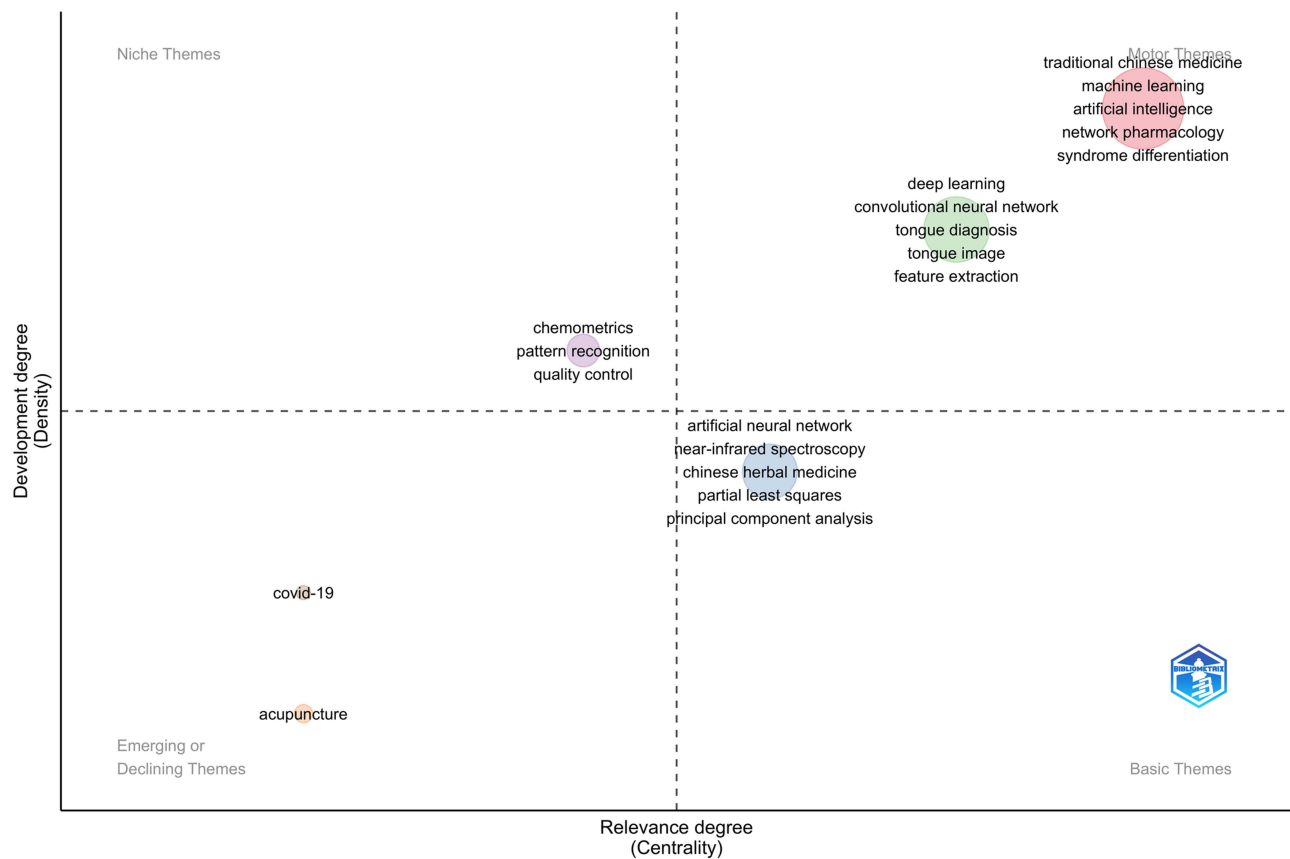


Figure 10 Thematic map of research trends in ML applications for TCM analysis. Six clusters identified: Cluster 1 “traditional Chinese medicine”, Cluster 2 “deep learning”, Cluster 3 “chemometrics”, Cluster 4 “artificial neural network”, Cluster 5 “covid-19” and Cluster 6 “acupuncture”.

Basic themes represent foundational topics essential to the field but may not currently be the focus. The themes are important but are not well-developed, with high centrality but low density. Cluster 4, the “artificial neural network” cluster in this quadrant focuses on using neural network technologies to enhance analytical and predictive methodologies within the scope of Chinese herbal medicine.

Emerging or declining themes are neither central nor well-developed, they have either not yet achieved widespread adoption or are being phased out in favor of more contemporary methods. The Cluster 5 and 6, “covid-19” and “acupuncture”, fall into this category.

Discussion

Publication Trends and Collaboration

The first publication on applying ML techniques to TCM appeared in 2000.²⁸ For many years, the number of publications remained minimal, reflecting an early stage of development. However, beginning in 2018, there was a significant surge in research output, marking a rapid growth and maturation period in this area. China leads in research output, highlighting its significant role in research development. The research collaboration network further reveals that China is the most collaborative nation, partnering with 24 countries. This international collaboration is crucial for the exchange of ideas and technologies, driving innovation. Leading institutions such as the Shanghai University of Traditional Chinese Medicine and the China Academy of Chinese Medical Sciences are at the forefront of this research. The work is distributed across 226 journals, including high-impact ones like “Pharmacological Research” and “Artificial Intelligence in Medicine”, indicating its relevance to various scientific domains like pharmacology, bioinformatics, and medical informatics.

Motor Themes: Machine Learning Applications in TCM Diagnosis and Network Pharmacology

Motor themes, including the application of ML in TCM diagnosis, network pharmacology, and the application of deep learning in tongue diagnosis, are highly relevant, well-developed, and central to the field. Unlike basic or niche themes, which may focus on foundational knowledge or emerging areas, motor themes represent key research hotspots that drive significant advancements and propel the field forward. These themes not only demonstrate maturity in research but also have a strong impact on the development of TCM and ML applications. The main motor themes include machine learning applications in TCM diagnosis and network pharmacology.

TCM Diagnosis

TCM diagnosis, which involves the objectification of the four diagnostic methods of TCM, including syndrome differentiation and treatment, as well as disease diagnosis, is part of the most developed and important research in Cluster 1.

Syndrome differentiation is critical for categorizing patients based on their symptoms and signs to determine the appropriate treatment. Significant milestones in syndrome differentiation include Liu GP’s 2014 study demonstrating that deep belief networks with multilabel learning enhance syndrome recognition accuracy for chronic gastritis.²¹ ANN has been effectively utilized for differentiating syndromes in diseases like Chronic Obstructive Pulmonary Disease (COPD) and Rheumatoid Arthritis, improving diagnostic precision and treatment personalization.^{22,29} Deep learning models have also shown remarkable success, as evidenced by a fusion model for syndrome differentiation in lung cancer diagnosis achieving an F1 score of 0.8884.³⁰ Additionally, extreme learning machine networks combined with particle swarm optimization have created robust syndrome classification models for primary liver cancer, incorporating diverse diagnostic inputs such as TCM symptoms, signs, tongue diagnosis, and pulse information.³¹ However, there was study showed that SVM proved effective in approximating TCM syndromes, outperforming Back Propagation Neural Network (BPNN) and extreme learning machines.³² This finding suggests the need for more targeted research to identify optimal ML algorithms for specific diagnostic tasks within TCM.

From the analysis of keywords trend topics, SVM and random forests are part of the most commonly used ML algorithms for TCM prediction models. For instance, Tang et al demonstrated that random forests can effectively predict TCM prescriptions for insomnia with a high accuracy of 0.85 by analyzing integrated diagnostic and syndrome

differentiation data;³³ Lim et al found that SVM performed better than other models with pulse wave parameters and TCM clinical indices in diagnosing polycystic ovary syndrome.³⁴ Other than that, Zhang et al developed an AI-based diagnosis system from health records, which applied a CNN for multiple disease prediction, achieving over 80% accuracy, and an integrated learning model (BPNN, Random Forest, XGboost, and SVM) for syndrome prediction.²⁴ These findings emphasize the importance of algorithm selection and model combinations, as well as careful consideration of data variety and completeness, to enhance predictive accuracy.

The high development and centrality of ML in TCM diagnosis underscore its widespread adoption and effectiveness in analyzing complex TCM medical data. This progress has significantly improved diagnostic accuracy and predictive capabilities, enabling highly individualized treatment plans. However, the lack of standardized ML frameworks in TCM, coupled with variable data quality and limited interpretability, raises concerns about its clinical applicability.

Integrating Network Pharmacology with Machine Learning

Integrating network pharmacology with ML is crucial for understanding TCM treatment mechanisms. The research is also included in Cluster 1. This interdisciplinary approach provides deep insights into the complex interactions of herbal compounds with biological systems, aiding in the development of new therapeutic strategies. For instance, Yang et al utilized network pharmacology and clustering algorithms to explore the mechanisms of TCM formulas in treating various types of coronary heart disease,³⁵ their study identified key bioactive compounds and potential therapeutic targets, providing valuable insights into the efficacy of these traditional remedies. Similarly, Shen et al used network pharmacology and ML algorithms (naïve Bayesian and recursive partitioning) to investigate Xiao-Xu-Ming Decoction for Alzheimer's disease,³⁶ the research uncovered multiple bioactive compounds with promising neuroprotective properties, highlighting the potential of this decoction in managing neurodegenerative conditions.

While these studies demonstrated the potential of combining network pharmacology and ML to enrich our understanding of the therapeutic effects of herbal medicine, their focus was limited to specific conditions and formulas. This specificity might restrict the applicability of their findings to broader TCM practices. Investigating a single TCM formula across various conditions could reveal common therapeutic pathways, thus broadening the relevance of the results. Furthermore, integrating these computational insights with traditional TCM theories and practices could provide a more comprehensive understanding, aiding in the broader application of these findings across different treatments.

Deep Learning in Tongue Diagnosis

Tongue diagnosis is a fundamental aspect of TCM diagnosis. Deep learning techniques, particularly CNN, are being employed to analyze tongue images for diagnostic purposes. This research hotspot is identified in Cluster 2, which is well-developed with numerous studies demonstrating its potential to automate and enhance the diagnostic process. The advancement of tongue diagnosis is notable by historiographic mapping.

For feature extraction research in tongue image analysis, Xu et al employed a multi-task joint learning model (UNET and discriminative filter learning) to segment and classify tongue images, utilizing selected features such as color and texture, to train the deep neural network.¹² This approach significantly improves the model's ability to differentiate between various conditions. Similarly, Yan et al proposed the SegTongue network, which combines CNN-based feature extraction with the center loss function, demonstrating superior performance for automated tongue image segmentation and color classification.³⁷

In tongue structure diagnosis, Ma et al developed a complexity perception classification method using CNN for tongue constitution recognition amid environmental variations, significantly improving recognition performance.³⁸ Wang et al focused on using CNN ResNet34 to recognize tooth-marked tongues in TCM, achieving over 90% accuracy with generalizability to different image sources and illuminations.²³

In disease diagnosis utilizing tongue images, the techniques were applied to conditions such as diabetes, gastric cancer, and lung cancer.^{39–41} Yuan et al demonstrated that tongue images and tongue coating microbiome, analyzed via deep learning models, can effectively diagnose gastric cancer, outperforming conventional blood biomarkers, with high accuracy, suggesting a noninvasive, cost-efficient screening method applicable in various settings.⁴⁰

The highly cited papers in the field have laid the foundation for advanced image recognition techniques. The most-cited paper, “Deep Residual Learning for Image Recognition” by He KM (2016), highlights the significance of deep learning in image analysis.²⁶ The second most-cited paper, “Very Deep Convolutional Networks for Large-Scale Image Recognition” by Simonyan K (2015), discusses the impact of deep convolutional network depth on image recognition accuracy.²⁷ The focus on deep learning in image recognition further signifies the advancements in diagnostic tools, especially in analyzing complex medical images such as tongue patterns.

Integrating advanced deep learning techniques enhances automated tongue diagnosis in TCM, leading to precise identification of tongue features that correlate with various health conditions, thus providing modernization of the ancient diagnostic method, and promoting mobile and personalized diagnostics innovations.

Niche Themes: Chemometrics in TCM Quality Control and Authentication

The Cluster 3 in the niche themes quadrant is well-developed but has limited influence on the broader field. It revolves around integrating ML techniques such as ANN and chemometric analyses, to evaluate, authenticate, and ensure the quality of TCM products.

Chemometrics involves applying mathematical and statistical methods to chemical data to extract relevant information. Studies within this cluster demonstrate the integration of ML algorithms with advanced chromatographic and spectroscopic methods. This combination ensures the precision, reliability, and efficiency of TCM evaluations. By incorporating techniques such as high-performance liquid chromatography (HPLC),⁴² ultra-performance liquid chromatography coupled with quadrupole time-of-flight mass spectrometry,⁴³ Fourier-transform infrared spectroscopy,⁴⁴ and near-infrared spectroscopy (NIRS),⁴⁵ the integration of ML and chemometrics significantly enhances the accuracy and consistency of TCM product evaluation. For instance, Tian et al applied both HPLC and high-performance thin-layer chromatographic fingerprinting coupled with chemometric analysis including ANN and K-Nearest Neighbors to assess Chaihu (*Bupleuri Radix*) for pattern recognition, revealing distinctions between various species and commercial samples.¹¹ Similarly, Wang et al’s work on Epimedium species authentication using HPLC fingerprint analysis for data collection; principal component analysis (PCA) and hierarchical cluster analysis for data analysis; and BPNN for pattern recognition to classify the species.⁴⁶ In the study on *Radix Scutellariae*, PCA streamlined data from multiple samples, identifying key markers differentiating crude and processed *Radix Scutellariae*, followed by SVM prediction.⁴⁷

This synergy supports the standardization and modernization of TCM practices, ultimately improving the overall quality control and authentication processes.

Basic Themes: Artificial Neural Networks in Chinese Herbal Medicines Analysis

Cluster 4 involving ANN in Chinese herbal medicines analysis is more central to the field than Cluster 3 but is less developed. This cluster underscores the advances of combining ANN with NIRS or analytical techniques, such as PCA and partial least squares (PLS). For example, Xiong et al combined deep Boltzmann machine for feature extraction and PLS analysis to analyze the dose-effect relationship of TCM, using experimental data from Ma Xing Shi Gan Decoction and datasets from the UCI Machine Learning Repository.⁴⁸ Chen et al proposed a 1D-CNN model based on NIRS to differentiate aristolochic acids and analogs, demonstrating strong feature learning capabilities.⁴⁹ By utilizing these techniques, researchers can achieve rapid, accurate, and comprehensive analysis of herbal medicines.

As observed from the top cited papers, ML aids in TCM drug discovery by providing powerful tools for data analysis and predictive modeling, capable of handling complex datasets to identify bioactive compounds and predict their pharmacological properties. Deep learning facilitates the virtual screening of thousands of compounds, the prediction of drug-target interactions,⁵⁰ and the assessment of ADMET (absorption, distribution, metabolism, excretion, and toxicity) properties.⁵¹ ML improves efficiency in identifying promising drug candidates, accelerating drug development from natural products particularly Chinese herbal.⁹ Recent advances in technologies, such as big data integration and predictive modeling, form multidimensional association networks and improve drug target reasoning algorithms.¹³ Notably, Zhong et al proposed the Siamese spectral-based graph convolutional network model for predicting drug targets and mechanisms, successfully identifying potential host targets of nelfinavir-cyclophilin A and interaction mechanisms

for SARS-CoV-2.⁵² These advancements demonstrate ML's transformative impact on modernizing TCM and developing Chinese herbal product-based therapies.

Declining Theme: COVID-19

Cluster 5 is about utilizing ML techniques for COVID-19 research. For example, Wang et al evaluated TCM prescriptions recommended for COVID-19 using ontology-based side-effect prediction and deep learning, suggesting safe treatments and providing a novel evaluation method.⁵³ The cluster is less developed and currently has limited relevance. While COVID-19 research was initially a hotspot due to the pandemic, its relevance has decreased as the immediate crisis has passed.

Emerging Theme: Machine Learning in Acupuncture

Cluster 6 is about the emerging role of ML in acupuncture, the integration marks a significant advancement in TCM in both the research and clinical practice. Sun et al illustrated how auriculotherapy, which involves stimulating specific points on the ear, can greatly benefit from deep learning that accurately locates these points, ensuring consistent and effective treatments.⁵⁴ Additionally, the study by Liu et al used the Random Forest model to analyze exosomal miRNA biomarkers for diagnosing migraine and predicting acupuncture treatment outcomes, thus providing reliable diagnostic and prognostic insights.⁵⁵ By applying ML, researchers and clinicians can gain deeper insights into the biological underpinnings of acupuncture, thus enhancing its credibility and integration into contemporary medical practices.

Limitations

Several limitations exist in our study and the field of research. Firstly, our study only searched the WoSCC database and did not include other medical databases, which inevitably resulted in sample bias. Secondly, our inferences and demonstrations were based on “Topic” search terms, which may not fully capture the main content of publications. This approach could have introduced some bias in topic classification. Additionally, most studies have been conducted in China, leading to potential biases and a lack of generalizability to other populations. Another limitation is the interpretability of machine learning models, particularly deep learning models, which often function as “black boxes”, making it difficult to understand the rationale behind their predictions. Variability in data quality and consistency across different studies may affect the reliability of the findings.

Future Directions

Increasing international collaboration is essential to enhance the generalizability of findings and foster the exchange of diverse perspectives and expertise. Collaborative efforts to create large, diverse, and high-quality datasets—incorporating data from varied populations and clinical settings—will be crucial for developing robust and generalizable ML models. Establishing standardized data collection and reporting protocols can further improve data quality and consistency, making it easier to compare and synthesize findings across studies. To support these goals, we suggest creating international data repositories with standardized protocols of ML frameworks in TCM, securing joint funding and exchange programs, and validating models across diverse clinical settings to enhance the global applicability of TCM research.

Additionally, enhancing the interpretability of machine learning models is essential for their acceptance and integration into clinical practice. Developing explainable AI techniques such as SHAP (Shapley Additive Explanations)⁵⁶ and LIME (Local Interpretable Model-agnostic Explanations),⁵⁷ help analyze feature importance by showing how each input variable influences the model's output. Additionally, attention mechanisms can help visualize the focal areas of decision-making within neural networks.⁵⁸ These methods offer clearer insights into model decisions, helping bridge the gap between advanced technologies and clinical applications.

Huang et al discussed the future of TCM research, emphasizing that advancements in AI technologies will shape the field by 2035.⁵⁹ Key opportunities include AI-driven improvements in drug discovery, diagnostics, and deeper integration with modern biomedicine. The authors further proposed that advancing AI applications in TCM would benefit from a collaborative, open-source model that includes both open-source AI tools and shared TCM datasets. These AI-driven

advancements could thus facilitate the modernization and global integration of TCM, supporting its acceptance on an international scale.

Conclusions

ML applications in TCM research have significantly enhanced diagnostic precision, treatment personalization, and quality control. By improving syndrome differentiation, disease diagnosis, and tongue image analysis, ML enables more accurate diagnostics and individualized treatment plans. ML has also strengthened TCM product identification and analysis, ensuring greater consistency, efficacy, and safety, and making TCM more reliable for domestic and global use. These findings highlight the translational potential of ML to position TCM as an understandable and applicable complementary approach within global healthcare, supporting integrative medicine where traditional and modern practices coexist and mutually enhance patient care. As computational power and algorithmic innovations continue to advance, the future of ML in TCM is promising.

Abbreviations

AI, Artificial intelligence; ANN, Artificial Neural Networks; BPNN, Back Propagation Neural Network; CNN, Convolutional neural networks; COPD, Chronic Obstructive Pulmonary Disease; HPLC, High-performance liquid chromatography; IF, Impact factor; LIME, Local Interpretable Model-agnostic Explanations; ML, Machine learning; NIRS, Near-infrared spectroscopy; PCA, Principal component analysis; PLS, Partial least squares; SHAP, Shapley Additive Explanations; SVM, Support Vector Machines; TCM, Traditional Chinese Medicine; WoSCC, Web of Science Core Collection.

Funding

This work was supported by the National Natural Science Foundation of China [grant number 82374336, 82074333] and the Shanghai Key Laboratory of Health Identification and Assessment [grant number 21DZ2271000].

Disclosure

The authors report no conflicts of interest in this work.

References

1. Zhang B, Li C, Lin N. 1. Introduction of Machine Learning. In: *Machine Learning and Visual Perception*. De Gruyter; 2020. doi:10.1515/9783110595567-002
2. Bakshi K, Bakshi K. Considerations for artificial intelligence and machine learning: approaches and use cases. *IEEE Aerospace Conference Proceedings*. 2018. doi:10.1109/AERO.2018.8396488
3. Shehab M, Abualigah L, Shambour Q, et al. Machine learning in medical applications: a review of state-of-the-art methods. *Comput Biol Med*. 2022;145. doi: 10.1016/J.COMPBIOMED.2022.105458.
4. Chen H, He Y. Machine learning approaches in traditional Chinese medicine: a systematic review. *Am J Chin Med*. 2022;50(1):91–131. doi:10.1142/S0192415X22500045
5. Tian D, Chen W, Xu D, et al. A review of traditional Chinese medicine diagnosis using machine learning: inspection, auscultation-olfaction, inquiry, and palpation. *Comput Biol Med*. 2024;170:108074. doi:10.1016/J.COMPBIOMED.2024.108074
6. Ellegaard O, Wallin JA. The bibliometric analysis of scholarly production: how great is the impact? *Scientometrics*. 2015;105(3):1809–1831. doi:10.1007/S11192-015-1645-Z/TABLES/9
7. Aria M, Cuccurullo C. bibliometrix: an R-tool for comprehensive science mapping analysis. *J Informetr*. 2017;11(4):959–975. doi:10.1016/J.JOI.2017.08.007
8. van Eck NJ, Waltman L. Software survey: vOSviewer, a computer program for bibliometric mapping. *Scientometrics*. 2010;84(2):523–538. doi:10.1007/S11192-009-0146-3/FIGURES/7
9. Thomford NE, Senthebane DA, Rowe A, et al. Natural products for drug discovery in the 21st century: innovations for novel drug discovery. *Int J Mol Sci*. 2018;19(6):1578. doi:10.3390/IJMS19061578
10. Tian S, Wang J, Li Y, Li D, Xu L, Hou T. The application of in silico drug-likeness predictions in pharmaceutical research. *Adv Drug Deliv Rev*. 2015;86:2–10. doi:10.1016/J.ADDR.2015.01.009
11. Tao TR, Shan XP, Ping LH. Evaluation of traditional Chinese herbal medicine: chaihu (Bupleuri Radix) by both high-performance liquid chromatographic and high-performance thin-layer chromatographic fingerprint and chemometric analysis. *J Chromatogr A*. 2009;1216(11):2150–2155. doi:10.1016/J.CHROMA.2008.10.127
12. Xu Q, Zeng Y, Tang W, et al. Multi-task joint learning model for segmenting and classifying tongue images using a deep neural network. *IEEE J Biomed Health Inform*. 2020;24(9):2481–2489. doi:10.1109/JBHI.2020.2986376
13. Zhu Y, Ouyang Z, Du H, et al. New opportunities and challenges of natural products research: when target identification meets single-cell multiomics. *Acta Pharm Sin B*. 2022;12(11):4011–4039. doi:10.1016/J.APSB.2022.08.022

14. Tian S, Wang J, Li Y, Xu X, Hou T. Drug-likeness analysis of traditional Chinese medicines: prediction of drug-likeness using machine learning approaches. *Mol Pharm.* 2012;9(10):2875–2886. doi:10.1021/MP300198D/SUPPL_FILE/MP300198D_SI_001.PDF
15. Zhou X, Peng Y, Liu B. Text mining for traditional Chinese medical knowledge discovery: a survey. *J Biomed Inform.* 2010;43(4):650–660. doi:10.1016/J.JBI.2010.01.002
16. Yang M, Chen JL, Xu LW, Ji G. Navigating traditional Chinese medicine network pharmacology and computational tools. *Evid Based Complement Alternat Med.* 2013;2013(1):731969. doi:10.1155/2013/731969
17. Zhao C, Li GZ, Wang C, Niu J. Advances in patient classification for traditional Chinese medicine: a machine learning perspective. *Evid Based Complement Alternat Med.* 2015;2015:1–18. doi:10.1155/2015/376716
18. Ung CY, Li H, Cao ZW, Li YX, Chen YZ. Are herb-pairs of traditional Chinese medicine distinguishable from others? Pattern analysis and artificial intelligence classification study of traditionally defined herbal properties. *J Ethnopharmacol.* 2007;111(2):371–377. doi:10.1016/J.JEP.2006.11.037
19. Sen CS, Huang HJ, Chen CYC. Two birds with one stone? Possible dual-targeting H1N1 inhibitors from traditional Chinese medicine. *PLoS Comput Biol.* 2011;7(12):e1002315. doi:10.1371/JOURNAL.PCBI.1002315
20. Zhang NL, Yuan S, Chen T, Wang Y. Latent tree models and diagnosis in traditional Chinese medicine. *Artif Intell Med.* 2008;42(3):229–245. doi:10.1016/J.ARTMED.2007.10.004
21. Liu GP, Yan JJ, Wang YQ, et al. Deep learning based syndrome diagnosis of chronic gastritis. *Comput Math Methods Med.* 2014;2014:1–8. doi:10.1155/2014/938350
22. Xu Q, Tang W, Teng F, et al. Intelligent syndrome differentiation of traditional Chinese medicine by ANN: a case study of chronic obstructive pulmonary disease. *IEEE Access.* 2019;7:76167–76175. doi:10.1109/ACCESS.2019.2921318
23. Wang X, Liu J, Wu C, et al. Artificial intelligence in tongue diagnosis: using deep convolutional neural network for recognizing unhealthy tongue with tooth-mark. *Comput Struct Biotechnol J.* 2020;18:973. doi:10.1016/J.CSBJ.2020.04.002
24. Zhang H, Ni W, Li J, Zhang J. Artificial intelligence-based traditional Chinese medicine assistive diagnostic system: validation study. *JMIR Med Inform.* 2020;8(6):e17608. doi:10.2196/17608
25. Wang Y, Shi X, Li L, Effert H, Shang D. The impact of artificial intelligence on traditional Chinese medicine. *Am J Chin Med.* 2021;49(6):1297–1314. doi:10.1142/S0192415X21500622
26. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2016; doi:10.1109/CVPR.2016.90.
27. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings. 2014. Available from: <https://arxiv.org/abs/1409.1556v6>. Accessed June 1, 2024.
28. Geng L, Luo AQ, Fu RN, Li J. Identification of Chinese herbal medicine using artificial neural network in pyrolysis-gas chromatography. *Chin. J. Anal. Chem.* 2000;28(5):549–553.
29. Xie J, Li Y, Wang N, Xin L, Fang Y, Liu J. Feature selection and syndrome classification for rheumatoid arthritis patients with Traditional Chinese Medicine treatment. *Eur J Integr Med.* 2020;34:101059. doi:10.1016/J.EUJIM.2020.101059
30. Liu Z, He H, Yan S, Wang Y, Yang T, Li GZ. End-to-end models to imitate traditional Chinese medicine syndrome differentiation in lung cancer diagnosis: model development and validation. *JMIR Med Inform.* 2020;8(6):e17821. doi:10.2196/17821
31. Ding L, You ZX, Yao WD, Ling LM. Application of an extreme learning machine network with particle swarm optimization in syndrome classification of primary liver cancer. *J Integr Med.* 2021;19(5):395–407. doi:10.1016/J.JOIM.2021.08.001
32. Yan E, Song J, Liu C, Luan J, Hong W. Comparison of support vector machine, back propagation neural network and extreme learning machine for syndrome element differentiation. *Artif Intell Rev.* 2020;53(4):2453–2481. doi:10.1007/S10462-019-09738-Z/METRICS
33. Tang Y, Li Z, Yang D, et al. Research of insomnia on traditional Chinese medicine diagnosis and treatment based on machine learning. *Chin Med.* 2021;16(1):1–21. doi:10.1186/S13020-020-00409-8
34. Lim J, Li J, Feng X, et al. Machine learning-based evaluation of application value of traditional Chinese medicine clinical index and pulse wave parameters in the diagnosis of polycystic ovary syndrome. *Eur J Integr Med.* 2023;64:102311. doi:10.1016/J.EUJIM.2023.102311
35. Yang J, Tian S, Zhao J, Zhang W. Exploring the mechanism of TCM formulae in the treatment of different types of coronary heart disease by network pharmacology and machine learning. *Pharmacol Res.* 2020;159:105034. doi:10.1016/J.PHRS.2020.105034
36. Shen Y, Zhang B, Pang X, et al. Network pharmacology-based analysis of xiao-xu-ming decoction on the treatment of Alzheimer's Disease. *Front Pharmacol.* 2020;11:1. doi:10.3389/FPHAR.2020.595254/FULL
37. Yan B, Zhang S, Yang Z, Su H, Zheng H. Tongue segmentation and color classification using deep convolutional neural networks. *Mathematics.* 2022;10(22):4286. doi:10.3390/MATH10224286
38. Ma J, Wen G, Wang C, Jiang L. Complexity perception classification method for tongue constitution recognition. *Artif Intell Med.* 2019;96:123–133. doi:10.1016/J.ARTMED.2019.03.008
39. Li J, Yuan P, Hu X, et al. A tongue features fusion approach to predicting prediabetes and diabetes with machine learning. *J Biomed Inform.* 2021;115:103693. doi:10.1016/J.JBI.2021.103693
40. Yuan L, Yang L, Zhang S, et al. Development of a tongue image-based machine learning tool for the diagnosis of gastric cancer: a prospective multicentre clinical cohort study. *EClinicalMedicine.* 2023;57:101834. doi:10.1016/j.eclinm.2023.101834
41. Shi Y, Guo D, Chun Y, et al. A lung cancer risk warning model based on tongue images. *Front Physiol.* 2023;14:1154294. doi:10.3389/FPHYS.2023.1154294/BIBTEX
42. Lai YH, Ni YN, Kokot S. Authentication of Cassia seeds on the basis of two-wavelength HPLC fingerprinting with the use of chemometrics. *Chin. Chem. Lett.* 2010;21(2):213–216. doi:10.1016/J.CCLET.2009.10.031
43. Han Y, Zhou M, Wang L, et al. Comparative evaluation of different cultivars of *Flos Chrysanthemi* by an anti-inflammatory-based NF- κ B reporter gene assay coupled to UPLC-Q/TOF MS with PCA and ANN. *J Ethnopharmacol.* 2015;174:387–395. doi:10.1016/J.JEP.2015.08.044
44. Lai Y, Ni Y, Kokot S. Classification of raw and roasted semen cassiae samples with the use of Fourier transform infrared fingerprints and least squares support vector machines. *Appl. Spectrosc.* 2010;64(6):649–656. doi:10.1366/000370210791414362
45. Dong WJ, Ni YN, Kokot S. Quantitative analysis of two adulterants in *Cynanchum stauntonii* by near-infrared spectroscopy combined with multi-variate calibrations. *Chem Papers.* 2012;66(12):1083–1091. doi:10.2478/S11696-012-0231-6/METRICS

46. Wang L, Wang X, Kong L. Automatic authentication and distinction of *Epimedium koreanum* and *Epimedium wushanense* with HPLC fingerprint analysis assisted by pattern recognition techniques. *Biochem Syst Ecol*. 2012;40:138–145. doi:10.1016/J.BSE.2011.10.014
47. Wang F, Wang B, Wang L, et al. Discovery of discriminatory quality control markers for Chinese herbal medicines and related processed products by combination of chromatographic analysis and chemometrics methods: radix *Scutellariae* as a case study. *J Pharm Biomed Anal*. 2017;138:70–79. doi:10.1016/J.JPBA.2017.02.004
48. Xiong W, Zhu Y, Zeng Q, et al. Dose-effect relationship analysis of TCM based on deep Boltzmann machine and partial least squares. *Math Biosci Eng*. 2023;20(8):14395–14413. doi:10.3934/MBE.2023644
49. Chen X, Chai Q, Lin N, Li X, Wang W. 1D convolutional neural network for the discrimination of aristolochic acids and their analogues based on near-infrared spectroscopy. *Anal. Methods*. 2019;11(40):5118–5125. doi:10.1039/C9AY01531K
50. Korotcov A, Tkachenko V, Russo DP, Ekins S. Comparison of deep learning with multiple machine learning methods and metrics using diverse drug discovery data sets. *Mol Pharm*. 2017;14(12):4462–4475. doi:10.1021/ACS.MOLPHARMACEUT.7B00578
51. Jing Y, Bian Y, Hu Z, Wang L, Xie XQS. Deep learning for drug design: an artificial intelligence paradigm for drug discovery in the big data era. *AAPS J*. 2018;20(3):58. doi:10.1208/S12248-018-0210-0
52. Zhong F, Wu X, Yang R, et al. Drug target inference by mining transcriptional data using a novel graph convolutional network framework. *Protein Cell*. 2022;13(4):281–301. doi:10.1007/S13238-021-00885-0
53. Wang Z, Li L, Song M, Yan J, Shi J, Yao Y. Evaluating the Traditional Chinese Medicine (TCM) officially recommended in China for COVID-19 Using Ontology-Based Side-Effect Prediction Framework (OSPF) and Deep Learning. *J Ethnopharmacol*. 2021;272:113957. doi:10.1016/J.JEP.2021.113957
54. Sun X, Dong J, Li Q, Lu D, Yuan Z. Deep learning-based auricular point localization for auriculotherapy. *IEEE Access*. 2022;10:112898–112908. doi:10.1109/ACCESS.2022.3215138
55. Liu L, Qi W, Wang Y, et al. Circulating exosomal microRNA profiles in migraine patients receiving acupuncture treatment: a placebo-controlled clinical trial. *Front Mol Neurosci*. 2023;15:1098766. doi:10.3389/FNMOL.2022.1098766/BIBTEX
56. Nohara Y, Matsumoto K, Soejima H, Nakashima N. Explanation of machine learning models using shapley additive explanation and application for real data in hospital. *Comput Methods Programs Biomed*. 2022;214:106584. doi:10.1016/J.CMPB.2021.106584
57. Barr Kumarakulasinghe N, Blomberg T, Liu J, Saraiva Leao A, Papapetrou P Evaluating local interpretable model-agnostic explanations on clinical machine learning classification models. *Proc IEEE Symp Comput Based Med Syst*. 2020; 7–12. doi:10.1109/CBMS49503.2020.00009.
58. Niu Z, Zhong G, Yu H. A review on the attention mechanism of deep learning. *Neurocomputing*. 2021;452:48–62. doi:10.1016/J.NEUCOM.2021.03.091
59. Huang N, Huang W, Wu J, Long S, Luo Y, Huang J. Possible opportunities and challenges for traditional Chinese medicine research in 2035. *Front Pharmacol*. 2024;15:1426300. doi:10.3389/FPHAR.2024.1426300/FULL

International Journal of General Medicine

Dovepress

Publish your work in this journal

The International Journal of General Medicine is an international, peer-reviewed open-access journal that focuses on general and internal medicine, pathogenesis, epidemiology, diagnosis, monitoring and treatment protocols. The journal is characterized by the rapid reporting of reviews, original research and clinical studies across all disease areas. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-general-medicine-journal>