American Journal of Health Research
2025, Vol. 13, No. 6, pp. 343–351
https://doi.org/10.11648/j.ajhr.20251306.17

SciencePG
Science Publishing Group

Research Article

A Review of the Current State of Multimodal Data Annotation in Traditional Chinese Medicine

Du Yutong, Yao Yuan, Zhou Hongwei*

Chinese Medicine Data Center, China Academy of Chinese Medical Sciences, Beijing, China

Abstract

Integrating artificial intelligence with traditional Chinese medicine (TCM) is a vital step towards the modernization and intelligent development of TCM. The aim of this article is to explore the current state of research and development in the annotation of multimodal data for TCM. First, the paper elucidates the background and core significance of multimodal data annotation. Then, it provides a systematic analysis of the various sources of TCM data, including textual materials such as classical texts and modern medical case records; imaging data such as facial, tongue, and ocular diagnoses; and signal data such as pulse and auscultation diagnoses. It also covers the technical characteristics of the annotation of these data types. Subsequently, the article explores detailed annotation methodologies tailored to various application scenarios. These include knowledge representation for classical Chinese medicine texts, in-depth annotation to capture the expertise of renowned traditional Chinese medicine practitioners, and standardized annotation for modern clinical information systems. Finally, this paper focuses on exploring the cross-disciplinary integration of global workspace theory (GWT) in cognitive neuroscience with multi-granularity computing and multimodal annotation in computer science. GWT offers an approach to understanding how diverse diagnostic information is synthesized into a coherent awareness, mirroring the process of TCM pattern differentiation. Multi-granularity computation is a methodological approach to processing this information through attention mechanisms at different levels. By integrating these advanced concepts with a practical framework for multimodal annotation, this paper aims to provide theoretical support and methodological insights for constructing intelligent diagnostic and therapeutic systems tailored to the cognitive characteristics of TCM.

Keywords

Traditional Chinese Medicine, Multimodal Fusion, Data Annotation, Review

1. Introduction

Traditional Chinese Medicine, a treasure of Chinese civilization, urgently requires deep integration with artificial intelligence technology for its modern development. The Healthy China 2030 Plan explicitly calls for advancing the application of health and medical big data, while the Several Opinions on Promoting the Development of Digital Traditional Chinese Medicine also emphasizes strengthening the construction of TCM data resource systems. The value of TCM data is increasingly evident, with data empowerment emerging as a key engine driving the inheritance and innovation of TCM. In this process, multimodal TCM data annotation—as the critical step in converting clinical information

*Corresponding author: zhwcatcm@163.com (Zhou Hongwei)

Received: 28 October 2025; **Accepted:** 12 November 2025; **Published:** 11 December 2025

into machine-understandable formats—has become the cornerstone for advancing intelligent TCM research [1]. Multimodal TCM data annotation refers to the process of standardizing and structuring annotated data, as well as semi-structured and unstructured data obtained through multiple channels [2]. Annotation is not merely a labeling task but involves constructing cognitive frameworks that enable machines to comprehend the TCM world. The quality and depth of multimodal data annotation directly impact the effectiveness of artificial intelligence applications [3].

This study systematically reviews the diverse sources of TCM multimodal data and their corresponding specific scenarios, aiming to provide theoretical support and methodological insights for building intelligent TCM systems.

2. Analysis of Multimodal Data Sources in TCM

2.1. Text Data

TCM textual data encompasses diverse sources, including ancient classical texts, modern clinical case records, herbal formulae materials, journal articles, and ethnic medicine compendiums [4]. At the forefront of intelligent processing for TCM textual data, large language models (LLMs) have emerged as critical tools due to their robust semantic understanding and generative capabilities. To adapt them for TCM applications, researchers have developed a series of key technologies: Retrieval-Enhanced Generation (REG) enhances answer accuracy by incorporating external knowledge bases (e.g., TCM knowledge graphs), effectively mitigating hallucination issues; Prompt Engineering and Chain of Thought guide models through multi-step reasoning to simulate TCM pattern differentiation; Human Feedback Reinforcement Learning integrates renowned TCM practitioners' expertise into model optimization, ensuring outputs align with clinical practice. Additionally, techniques like efficient parameter fine-tuning enable general-purpose large models to precisely adapt to specialized TCM tasks such as question-answering, medical case generation, and prescription recommendation [5]. The integrated application of these technologies propels TCM text processing from simple information extraction toward a new phase of deep understanding and decision support.

2.2. Image Data

Imaging data in TCM primarily refers to visual representations such as facial, tongue, and ocular diagnosis recorded in image format.

Facial diagnosis imaging data mainly involves steps including image acquisition, face detection and segmentation, facial region localization, and feature extraction [6]. Researchers have comprehensively employed multiple pattern recognition methods—such as ensemble learning, clustering, and regression analysis—to quantitatively extract facial color and luster features [7].

Research on tongue diagnosis objectification focuses on digital acquisition and feature analysis of tongue patterns. Acquisition methods are primarily categorized into standardized collection under natural lighting and closed-system collection using specialized tongue imaging devices. Post-acquisition feature analysis employs 3D imaging, hospital information system models, and RGB numerical analysis. Hierarchical analysis and edge correction algorithms enable automatic classification of tongue shapes, while tongue segmentation technology achieves precise identification of morphological features such as tooth marks and fissures [8]. Automated annotation of tongue images represents a frontier in tongue diagnosis image processing. Song Xin et al. [9] explicitly proposed an annotation algorithm based on adversarial autoencoders, aiming to translate visual morphology into structured semantic descriptions to provide objective evidence for clinical diagnosis.

Digital research in ocular and labial diagnosis has also advanced. For ocular diagnosis, disease classification and feature quantification are achieved primarily through iris and fundus image analysis combined with deep learning models. Regarding labial diagnosis, image segmentation and color space analysis are employed for lip diagnosis localization and lip color classification [6].

2.3. Signal Data

Pulse diagnosis, as the most distinctive diagnostic method in TCM, primarily relies on the eight elements of pulse characteristics (pulse position, pulse force, pulse rate, pulse rhythm, pulse length, pulse width, fluidity, and tension) to construct a quantitative annotation system. Data sources are concentrated on two pathways: pulse waveform analysis and radial artery ultrasound. Pulse waveform analysis extracts feature through methods like time-domain, frequency-domain, and combined time-frequency domain approaches, while radial artery ultrasound provides hemodynamic parameters such as lumen volume and blood flow velocity. Yi et al. [10] noted that three-dimensional pulse waveforms incorporating positional information combine the advantages of both methods, offering a more comprehensive reflection of the three-dimensional motion of the pulse vessel.

Auscultation encompasses both auditory and olfactory aspects. Although the final data for olfactory diagnosis relies on chemical composition analysis, the front-end process of capturing odor information via sensors like electronic noses fundamentally converts intangible odor biomarkers into analyzable electrical signals or digital spectra. This aligns with the technical workflow of signal data from acquisition to analysis. Therefore, this study places both auditory and olfactory diagnosis within the signal data framework for exploration. Specifically, auditory diagnosis employs high-precision sensors to capture patient speech, constructs a mathematical model based on the Fast Fourier Transform (FFT) to digitize sound signals, and then utilizes deep learning algorithms such as convolutional neural networks for feature extraction and Five-Element Constitution classification [11]. Olfactory diagnosis, grounded in traditional Chinese medical theories like "five odors corresponding to five organs," relies on analytical instruments such as electronic noses and gas chromatography-mass spectrometry systems. By capturing and analyzing human volatile organic compounds (VOCs), it constructs odor fingerprint spectra to enable early disease screening and auxiliary identification of TCM syndromes. Artificial intelligence and machine learning algorithms establish mapping models between VOCs data and diagnostic conclusions, driving the intelligent evolution of olfactory diagnosis [12].

Compared to single-modality approaches, multimodal TCM data offers stronger support for clinical decision-making. However, researchers [3] explicitly note that applying this to real-world TCM practice requires solving the core challenge of accurately mapping non-textual modalities (tongue patterns, pulse patterns) to knowledge graph nodes. The semantic disconnect across modalities remains a critical bottleneck hindering deep cognitive integration in TCM AI.

3. Annotation Systems and Methods for Different Scenarios

3.1. Knowledge of Ancient Chinese Medical Texts

The digitization of classical Chinese medical texts is a process of information reorganization that elevates it to knowledge. Its primary task is to bridge the linguistic gap between these texts and modern Chinese. Sun Haishu et al. [13] were the first in the field of traditional Chinese medicine to explicitly categorize annotation into semantic annotation and guiding annotation, revealing its dual role in actively constructing cognitive frameworks—both restoring the original meaning of knowledge and guiding machine reasoning pathways.

In our initial exploration of digitizing classical Chinese medical texts, we first encountered numerous succinct symptom descriptions such as "dry throat and mouth" or "scanty and dry tongue coating." Directly applying generic natural language processing models or complex knowledge representation frameworks proved cumbersome and inefficient. To address this challenge, Wang et al. [14] proposed a symptom information extraction method using Hidden Markov Models (HMM), specifically a diagnostic marker set based on healthy-value pairs. This approach defines symptom location as the healthy state (K) and symptom manifestation as the value state (V), enabling rapid and accurate extraction of core diagnostic information from classical texts while effectively filtering irrelevant vocabulary (UL). While highly efficient for symptom description extraction, this model inherently lacks expressive power to capture more complex semantic relationships. It struggles with phenomena like concise classical language and polysemy (e.g., the differing meanings of "hòu" in '午后颧红' [red cheeks in the afternoon] and "里急后重" [urgent tenesmus]). Liu Bo et al. [15] introduced a second-order hidden Markov model. When processing "午后颧红" and "里急后重," the model accurately classifies "后" as a temporal adverb or body part based on the semantic context of the following words '红' or "重." This research represents a technical approach to lexical-level parsing.

The methodological evolution from lexical to relational analysis is fully realized in Lin Ruifan et al.'s [16] research on constructing a knowledge graph for the Tang Dynasty edition of Shang han Lun. This work employs ontology engineering as its overarching design, systematically defining 25 core concepts and 22 semantic relations. By constructing 3,651 triples from the original text, it transforms isolated symptom descriptions into a knowledge network featuring a complete logical chain encompassing cause, mechanism, syndrome, treatment method, and medication. Building upon this precision knowledge network construction, Lu Yongmei et al. [17] further introduced the Binary-BERT model for zero-shot learning. This approach avoids memorizing specific disease patterns by learning abstract semantic associations between classical clinical experience and disease labels, enabling automatic classification of diseases unseen during training.

3.2. In-Depth Annotation of the Experience Inheritance from Renowned Senior TCM Practitioners

The goal of inheriting the experience of renowned senior TCM practitioners is to construct a cognitive framework capable of reproducing unique clinical reasoning. At the TCM education level, this approach helps these practitioners better transmit their medical thinking while activating the self-learning abilities of successors [18]. Taking the research on inheriting the "Universal Treatment Formula" concept of Master Yu Yingao as an example [5], the annotation process for inheriting renowned senior practitioners' experience first involves structuring diagnostic and therapeutic elements—standardizing information such as symptoms, syndrome patterns, treatment methods, and prescriptions. Next, it entails annotating the relational logic of diagnosis and treatment, using association rules and complex networks to reveal intrinsic connections among symptoms, treatment methods, and medications. and ultimately achieving model-based annotation of clinical reasoning to construct a clinical thinking model for renowned senior practitioners. This paradigm is equally applicable to interpreting and transmitting the clinical wisdom of other distinguished practitioners [19], providing a replicable, universal pathway for academic inheritance in traditional Chinese medicine.

3.3. Standardized Annotation for Modern Clinical Information Systems

Building high-quality, standardized clinical data is the core of modern clinical information systems, a need that has received significant attention and implementation in national-level research projects. Taking the Big Data Knowledge Engineering Project for Liver Diseases in Traditional Chinese Medicine as an example [20], the primary task for researchers is to establish a standardized system for liver diseases. This system first annotates clinical standards, encompassing diagnostic criteria for syndromes and clinical guidelines. It then annotates data standards, covering multiple aspects such as classification and coding of tongue pattern information, classification and coding of pulse pattern information, and data element standards. This work demonstrates that standardized coding and annotation of clinical data are prerequisites for transforming raw case data into research-ready data.

This approach of structurally deconstructing clinical information demonstrates remarkable vitality across different granularities of TCM knowledge representation. At the symptom name recognition level, Wei Zunqiang et al. [21] employed HMM to construct a symptom name recognition model. In this model, the observation sequence consists of characters in clinical texts, while the hidden state sequence represents the BIO tags to be identified. By modeling state transition probabilities and observation emission probabilities, HMM describes the generation process of clinical texts. Through a decoding algorithm, the system infers the most probable hidden label sequence from observed characters, thereby precisely delineating symptom name boundaries. At the syndrome decision level, You Zhengyang et al. [22] addressed the issue of synonymous but distinct syndrome names, such as “heart qi deficiency syndrome” and “heart qi deficiency syndrome.” They defined a set of cleansing rules for disorganized raw syndrome definitions: content within parentheses () is marked as ‘ignorable’ metadata, while content within square brackets [] is marked as “replaceable” metadata. Based on these rules, machines can automatically decompose “Heart Qi (Deficiency) Syndrome” into “Heart Qi Deficiency Syndrome” and “Chong-Ren Dysregulation Syndrome” into “Chong-Ren Dysregulation Syndrome” and “Chong-Ren Imbalance Syndrome.” This work automates the terminology normalization process by introducing a simple structured annotation system that formalizes experts' semantic judgments on syndromes.

This section systematically presents annotation methods tailored for three primary scenarios: classical texts, renowned physicians' experience, and clinical information systems. While these methodologies differ in focus, their shared objective is to achieve the structured transformation of knowledge within the TCM domain, thereby establishing foundational support for subsequent deep mining and intelligent empowerment.

4. An Approach to Multimodal Data Annotation in TCM

4.1. Theoretical Foundation: Global Workspace Theory and Multi-Granularity Computing

4.1.1. Global Workspace Theory

GWT proposed by Baars [23, 24] serves as a foundational theory in modern cognitive neuroscience [25]. This theory describes the brain as a cognitive architecture composed of numerous specialized modules and a central workspace. Through selective attention mechanisms [26], information from each module can be selected into the global workspace (GW) for global broadcasting and shared among modules.

In recent years, neuroscientists have identified similarities between the brain and deep artificial neural networks, noting that deep learning simulates the brain's perceptual and cognitive functions through multi-layer neural network computations [27, 28]. Consequently, the application scope of GWT has expanded beyond explaining conscious phenomena in the brain to encompass the virtual realm of artificial intelligence. Jaegle et al. [29] demonstrated that the Perceiver architecture satisfies GWT's functional criteria. By processing heterogeneous inputs through a core latent space with finite capacity, it implements GWT's core principles—cross-module information broadcasting, selective attention, and working memory—thus constructing a functional GW within artificial systems. Dossa et al. [26] implemented and evaluated an audiovisual-input-based 3D navigation agent in a real multimodal environment. Its architecture satisfied all four GWT metric attributes proposed by Butlin et al. [30]. Analysis of attention weights revealed the agent primarily integrates audiovisual information through cross-modal attention—a process highly consistent with Traditional Chinese Medicine's emphasis on cross-verification and interrelatedness among the Four Diagnostic Methods.

Based on global workspace theory, the integrated application of the Four Diagnostic Methods in TCM can be conceptualized as follows: multimodal information obtained through observation, auscultation, inquiry, and palpation enters the brain's perceptual buffer via distinct sensory channels, competitively activating corresponding cognitive processing modules within the GW [3]. During this process, the brain's central executive system selectively attends to perceptual information based on the current diagnostic context, forming core conscious events temporarily stored in working memory. Subsequently, this conscious event dynamically interacts with and refines diagnostic knowledge stored in long-term memory, including syndrome concept networks and clinical experience. Ultimately, diagnostic information distributed across modules is integrated into a unified diagnostic decision through the GW as an information hub. The TCM diagnostic and therapeutic process embodies the GWT description that “consciousness arises from the integration and broadcasting of information within the global workspace.”

4.1.2. Multi-Granularity Computation

Multi-granularity computation refers to processing uncertain complex problems at different scales or levels to obtain a more comprehensive and robust information processing paradigm [31]. It addresses the inherent hierarchical nature of information processing, providing concrete methods for achieving refined information integration within GW. This technique has been successfully applied across various AI tasks: In Chinese medicine text relation extraction, Wang et al. [32] constructed a model integrating character-level and word-level granularity to adaptively capture semantic units ranging from characters to words. In medical visual question answering, Wang et al. [33] developed a multi-granularity medical VOA model that represents question features at word, phrase, and sentence levels. Attention modules fuse these multi-granularity question features with image features to uncover correlations across granularity dimensions; For facial expression recognition, Xia et al. [34] a multi-granularity multi-scale feature fusion network. By progressively inputting facial puzzles from fine to coarse granularity, it extracts local fine-grained information, coarse-grained information, and global information; In medical image analysis, Liu et al. [35] constructed multi-granularity inputs through unsupervised methods and employed attention mechanisms for adaptive fusion of features across different granularities, significantly enhancing performance in hand bone age assessment.

Within the framework of a global workspace, multimodal information in traditional Chinese medicine inherently possesses multi-level characteristics. Zhang Pengfei et al. [36] developed a multi-granularity patio-temporal data fusion model by combining recurrent neural networks with sparse coding networks and Gaussian cloud transformation mechanisms. This model simulates the cognitive process of traditional Chinese physicians—“from fine to coarse, from local to holistic”—enabling switching between different cognitive levels: for instance, identifying subtle differences in single symptoms at the spatial granularity level (e.g., pale and swollen tongue body, weak and slow pulse); at the temporal granularity level, grasping the evolutionary trends of syndromes (e.g., tongue coating changing from white to yellow, or the improvement process of poor appetite symptoms); and at the patio-temporal granularity level, integrating macro variables such as solar terms and geography. This multi-level information provides structured inputs for GW to perform cross-modal and cross-level fusion and decision-making.

4.2. Core Paradigm: Multimodal Fusion Annotation

The essence of pattern identification and treatment in TCM is a multimodal fusion cognitive process, which aims to integrate information from multiple modalities at different levels to achieve more accurate clinical decision-making [37]. Based on this understanding, scholars [36] have proposed a multi-granularity computing and multimodal deep learning algorithm model for multi-source information fusion. This model has been applied in fields such as disease auxiliary diagnosis [37], rumor detection [38], and natural language processing [39].

In modern medicine, this approach demonstrates equally powerful potential. Huang et al. [40] proposed multimodal multi-granularity fusion neural networks that simultaneously integrate local granularity (nodes and edges) and global granularity (whole-brain small-world topological properties) across structural and functional brain networks, elevating epilepsy recognition rates to 83.94%. Yu et al. [41] integrated MRI and CT images of cervical spondylosis, employing pixel-level and feature-level fusion strategies with DenseNet-121 as the baseline model to explore multimodal data's efficacy in predicting postoperative neck pain. Park et al. [42] successfully classified Parkinson's disease patients into distinct severity subtypes by integrating clinical scales, physical fitness tests, and wearable sensor data using unsupervised clustering algorithms. Simultaneously, they identified core digital biomarkers from multimodal data through mutual information and recursive feature elimination algorithms, revealing that gait parameters from wearable sensors correlate most strongly with PD severity.

Although systematic research on multimodal, multi-granularity fusion annotation remains unexplored in the field of objective analysis of TCM data, preliminary studies align with this paradigm. For instance, Kang et al. [43] constructed a depression recognition model using distributed representations based on multimodal correlation annotation, systematically demonstrating the integration of TCM depression diagnosis across three granularity levels: At the fine-grained level, local features such as fixation points and pupil diameter in eye-tracking data, and mouth corner movements and brow state in facial expression data were annotated. At the medium-grained level, multimodal associations were established between eye-tracking features (e.g., vacant gaze), facial features (e.g., drooping mouth corners), and speech features (e.g., monotonous intonation), collectively pointing to the pathogenesis of liver qi stagnation. At the coarse-grained level, all modality information is integrated to output a standardized diagnostic conclusion for depression with liver qi stagnation syndrome. This study confirms that multimodal fusion techniques at either the feature layer or decision layer significantly outperform unimodal models in both recognition accuracy and F1 score. Li Rongyao et al. [44] constructed a multimodal knowledge graph for

Compendium of Materia Medica through graph representation, linking textual and image entities to form a semantic network comprising 10,799 entities and 14,686 relationships. This graph includes multimodal entities such as Chinese herbal medicines, herbal products, and diseases, providing a pioneering model for constructing multimodal knowledge graphs in TCM.

The global workspace theory and multi-granularity computation provide theoretical foundations for constructing multimodal fusion models aligned with TCM cognition. Multimodal co-annotation, serving as the data foundation for achieving deep model-internal fusion, has demonstrated significant advantages in some studies. However, current research predominantly relies on traditional neural networks, and studies on unified representation and deep fusion based on the Transformer architecture remain unexplored in the TCM domain [3]. This represents both a current research bottleneck and a key direction for future breakthroughs.

5. Conclusion and Outlook

Significant breakthroughs have been achieved in the technical application and standard exploration of multimodal data annotation for TCM, with outcomes empowering multiple TCM AI scenarios such as chronic disease management and Chinese herbal medicine quality inspection [45]. Regarding standard development, the industry has preliminarily established a comprehensive operational paradigm for TCM data annotation and semantic management, laying the foundation for standardized processing of multi-source data [2, 46, 47]. However, its further development and clinical implementation still face a series of challenges [36, 48, 49]: The scarcity of high-quality, large-scale annotated data severely limits the depth and breadth of model training; inconsistent standards among different data collection devices lead to prominent data heterogeneity issues, creating technical barriers for multi-source information fusion; existing fusion models have high computational complexity, posing practical difficulties for deployment in primary healthcare settings or edge devices; The black-box decision-making mechanism of models results in poor interpretability, undermining clinicians' trust and adoption of these models; current research risks unclear boundaries between technology and TCM theory, potentially allowing modern technical standards (such as Western quantitative indicators) to erode TCM ontology, failing to genuinely serve TCM pattern differentiation thinking; medical data involves substantial personal privacy, and the security and privacy protection systems during its sharing and use remain underdeveloped, becoming a bottleneck for cross-institutional and cross-platform application of models.

Future research should focus on the following directions: First, establish industry-standardized protocols for collecting data from the Four Examinations, construct high-quality multimodal TCM datasets, and leverage technologies like federated learning to enable cross-institutional collaboration while ensuring data security. Second, develop lightweight, explainable fusion models that reduce computational overhead, enhance decision transparency, and improve clinical utility and credibility. Third, adhere to TCM theoretical guidance to ensure intelligent technologies serve TCM pattern differentiation thinking, preventing the dilution of TCM characteristics. Fourth, explore human-machine collaborative intelligent diagnosis paradigms, positioning AI as an auxiliary tool while retaining final decision-making authority with clinicians, thereby achieving the organic integration of data intelligence and clinical wisdom.

Abbreviations

TCM	Traditional Chinese Medicine
LLMs	Large Language Models
REG	Retrieval-Enhanced Generation
FFT	Fast Fourier Transform
VOCs	Volatile Organic Compounds
HHM	Hidden Markov Models
GWT	Global Workspace Theory
GW	Global Workspace

Author Contributions

Du Yutong: Writing - original draft
Yao Yuan: Writing - review & editing
Zhou Hongwei: Supervision, Funding acquisition, Writing - review & editing

Funding

This work is supported by the 'Research on Standardization of Medical Case Records from Renowned Physicians' project, which is funded by the Basic Research Fund of the China Academy of Chinese Medical Sciences. (Grant No.ZZ1718-XRZ-102-SJ).

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Yang, T., Ren, H. Y., Zhou, Z. J., et al. Issues and Challenges in the AI-Empowered High-Quality Development of Traditional Chinese Medicine. Journal of Nanjing University of Traditional Chinese Medicine. 2024, 40(12), 1285-1290. [In Chinese] https://doi.org/10.14148/j.issn.1672-0482.2024.1285

[2] China Information Association of Chinese Medicine. Specification for Annotation of Multimedia Resources in Chinese Medicine (T/CIATCM 114—2024) [Z].2024. [In Chinese].

[3] Lin, S. Y., Huang, H. W., Liu, C., et al. Cognitive mechanisms and multimodal research methodologies of traditional Chinese medicine diagnostics under the large language model perspective. Chinese Journal of Traditional Chinese Medicine (formerly China Journal of Traditional Chinese Medicine and Pharmacy). 2025, 40(1), 98-102. [In Chinese]

[4] Wang, P., Research on Text Classification for Chinese Medicine Knowledge. Master's Thesis, North China University of Science and Technology. Tangshan, 2021. [In Chinese] https://doi.org/10.27108/d.cnki.ghelu.2021.000675

[5] Wang, X. P., Study on clinical thinking of "causes tracing, diseases differentiation, syndromes differentiation" and the decision model of spleen and stomach diseases under the guidance of Traditional Chinese Medicine master Yu Ying'ao's Tongzhi prescriptions. Ph. D. Thesis, China Academy of Chinese Medical Sciences. Beijing, 2025. [In Chinese]

[6] Wang, S. Y., Yang, J., Tian, Y. Research Progress of Key Technologies of TCM Clinical Information Collection Based on TCM Data Science. Zhonghua Zhongyiyao Xuekan. 2024, 42(3), 10-18. [In Chinese] https://doi.org/10.13193/j.issn.1673-7717.2024.03.003

[7] Li, H. Y., Li, C., Lang, X. F., et al. Analysis of the Research Status and Hot Spot of Intelligent Four-Diagnosis in TCM. Journal of Nanjing University of Traditional Chinese Medicine. 2022, 38(2), 180-186. [In Chinese] https://doi.org/10.14148/j.issn.1672-0482.2022.0180

[8] He, J., Application Value of Traditional Chinese Medicine Syndrome Differentiation Artificial Intelligence System Based on Bayesian Network in Tongue Diagnosis and Pulse Diagnosis. Journal of Qiqihar Medical University. 2019, 40(22), 2872-2873. [In Chinese] https://doi.org/10.3969/j.issn.1002-1256.2019.22.040

[9] Song, X. Design and Implementation of a TCM Interrogation System Integrating Tongue Image Annotation. Master's Thesis, Dalian University of Technology, Dalian, 2020. [In Chinese] https://doi.org/10.26991/d.cnki.gdllu.2020.001537

[10] Yi, K., Zhang, M., Guo, S., et al. Research review on data annotation for intelligent pulse diagnosis devices in traditional Chinese medicine. Shanghai Journal of Traditional Chinese Medicine. 2024, 58(10), 5-10. [In Chinese] https://doi.org/10.16305/j.1007-1334. 2024.z20240620006

[11] Xu, R., Lai, Y., Niu, T., Niu, X. Simulation of Digital TCM Auscultation Based on Deep Learning. Computer Simulation. 2025, 42(7), 525-528. [In Chinese] https://doi.org/10.3969/j.issn.1006-9348.2025.07.089

[12] Xu, Z., Fu, Q., Jiang, Y., et al. Textual Research on the Origin and Evolution of Olfactory Diagnosis in Traditional Chinese Medicine. Modernization of Traditional Chinese Medicine and Materia Medica-World Science and Technology. 2025, 27(10), 2790-2799. [In Chinese] https://doi.org/10.11842/wst.20241116001

[13] Sun, H., Li, B., Wang, R., Fu, Y. Discussion on Several Issues in the Annotation of Traditional Chinese Medicine Ancient Book Bibliographic Databases. Chinese Journal of Information on Traditional Chinese Medicine. 2007, 14(10), 103-104. [In Chinese] https://doi.org/10.3969/j.issn.1005-5304.2007.10.058

[14] Wang, G., Du, J., Hao, Z., et al. POS Tagging and Feature Recombination for Ancient Prose of TCM Diagnosis. Computer Engineering and Design. 2015, 36(3), 835-841. [In Chinese] https://doi.org/10.16208/j. issn1000-7024.2015.03.054

[15] Liu, B., Du, J., Nie, B., et al. Part-of-speech Tagging of Traditional Chinese Medicine Diagnosis Ancient Prose Based on Second order HMM. Computer Engineering. 2017, 43(7), 211-216. [In Chinese] https://doi.org/10.3969/j.issn.1000-3428.2017.07.035

[16] Lin, R. Knowledge Graph of Tang-Dynasty-version Treatise on Febrile Diseases Constructed Based on Ontology Method. Master's Thesis, China Academy of Chinese Medical Sciences. Beijing, 2021. [In Chinese] https://doi.org/10.27658/d.cnki.gzzyy.2021.000210

[17] Lu, Y. M. Extracting and analyzing the fine-grained relationship of clinical experiences in ancient literature of traditional Chinese Medicine. Master's Thesis, Sichuan University. Chengdu, 2022. [In Chinese] https://doi.org/10.27342/d.cnki.gscdu.2022.004503

[18] Shi, W. Y. Clinical Multi-Modal Data for Famous Doctor Inheritance Integrated Acquisition and Analysis. M.S. Thesis, China Academy of Chinese Medical Sciences, Beijing, 2024. [In Chinese] https://doi.org/10.27658/d.cnki.gzzyy.2024.000385

[19] Jiao, X. Y., Wang, Y., Yao, L., et al. Analysis of the Disease-Symptom-Syndrome-Medicine Association Rules in Medical Cases of Primary Dysmenorrhea Treated by National Traditional Chinese Medicine Master GUO Chengjie Based on Complex Networks. Journal of Traditional Chinese Medicine. 2025, 31(7), 150-157. [In Chinese] https://doi.org/10.13862/j.cn43-1446/r.2025.07.026

[20] Li, J., Li, X. D., Mao, S. S., et al. Research on big data knowledge engineering of TCM clinical liver diseases. Chinese Journal of Integrated Traditional and Western Medicine on Liver Diseases. 2023, 33(5). [In Chinese] https://doi.org/10.3969/j.issn.1005-0264.2023.005.001

[21] Wei, Z., Shu, H., Wang, Y. Research on Traditional Chinese Medicine Symptom Name Recognition Technology Based on Sequence Labeling. International Journal of Materials Science and Applications. 2022, 11(3), 66-75. [In Chinese] https://doi.org/0.11648/j.ijmsa.20221103.12

[22] You, Z., Wang, Y., Shu, H. Construction of a Part-of-Speech Tagged Corpus for Traditional Chinese Medicine Syndrome Names. International Journal of Materials Science and Applications. 2017, 6(21), 177-178. [In Chinese] https://doi.org/10.20109/j.cnki.etse.2017.21.139

[23] Baars B J. Global workspace theory of consciousness: toward a cognitive neuroscience of human experience [M]//LAUREYS S. Progress in Brain Research. Amsterdam: Elsevier, 2005, 150: 45-53. https://doi.org/10.1016/S0079-6123(05)50004-9

[24] Baars B. J., Geld N., & Kozma, R. (2021). Global Workspace Theory (GWT) and Prefrontal Cortex: Recent Developments. Frontiers in Psychology, 12, 749868. https://doi.org/10.3389/fpsyg.2021.749868

[25] VANRULLEN R, KANAI R. Deep learning and the Global Workspace Theory [J]. Trends in Neurosciences, 2021, 44(9): 692-702. https://doi.org/10.1016/j.tins.2021.04.005

[26] DOSSA R F J, ARULKUMARAN K, JULIANI A, et al. Design and evaluation of a global workspace agent embodied in a realistic multimodal environment [J]. Frontiers in Computational Neuroscience, 2024, 18: 1352685. https://doi.org/10.3389/fncom.2024.1352685

[27] Marblestone A H, Wayne G, Kording K P. Toward an Integration of Deep Learning and Neuroscience [J]. Frontiers in Computational Neuroscience, 2016, 10: 94. https://doi.org/10.3389/fncom.2016.00094

[28] VanRullen R. Perception Science in the Age of Deep Neural Networks [J]. Frontiers in Psychology, 2017, 8: 142. https://doi.org/10.3389/fpsyg.2017.00142

[29] Jaegle A., Gimeno F., Brock A., Vinyals O., Zisserman A., Carreira J., et al. (2021b). "Perceiver: general perception with iterative attention," in Proceedings of the 38th International Conference on Machine Learning (PMLR:), 4651-4664. https://doi.org/10.48550/arXiv.2103.03206

[30] Butlin P., Long R., Elmoznino E., Bengio Y., Birch J., Constant A., et al. (2023). Consciousness in artificial intelligence: insights from the science of consciousness. arXiv [preprint]. https://doi.org/10.48550/arXiv.2308.08708

[31] Zhang, Q. H., Wang, G., Hu, J. Multi-granularity Knowledge Acquisition and Uncertainty Measure. Beijing: Science Press; Beijing, 2013.

[32] Wang, Y. Research on TCM Text Relation Extraction Based on Multi-granularity Information. Master's Thesis, North China University of Science and Technology, Tangshan, China, 2021. [In Chinese] https://doi.org/10.27108/d.cnki.ghelu.2021.000119

[33] Wang H., Pan H., Zhang K., He S., Chen C. M2FNet: multi-granularity feature fusion network for medical visual question answering Pacific rim international conference on artificial intelligence, Springer (2022), pp. 141-154. https://doi.org/10.1007/978-3-031-20865-2_11

[34] Xia H., Lu L., Song S. Feature fusion of multi-granularity and multi-scale for facial expression recognition Vis Comput (2023), pp. 1-13. https://doi.org/10.1007/s00371-023-02900-3

[35] LIU B, HUANG Y, LI S, et al. Bone age assessment by multi-granularity and multi-attention feature encoding [J]. Quantitative Imaging in Medicine and Surgery, 2024, 14(8): 5902-5914. https://doi.org/10.21037/qims-23-806

[36] Zhang, P., Zeng, P., Wang, et al. Intelligent Integration Empowers Objectification of "Four Diagnostic Methods Integration" in Chinese Medicine. Chinese Journal of Integrated Traditional and Western Medicine. 2025. [In Chinese] https://doi.org/10.7661/j.cjm.20250819.285

[37] Qi, X., Ding, W., Ju, H., et al. Epilepsy Identification Method Based on Multi-modal Multi-grained Fusion Network. Journal of Data Acquisition and Processing. 2024, 39(3), 710-723. [In Chinese] https://doi.org/10.16337/j.1004-9037.2024.03.018

[38] Liu, X., Xiang, A., Du, Y. Multimodal Rumor Detection Method Based on Multi-granularity Emotional Features of Image-Text. Journal of Frontiers of Computer Science and Technology. 2025, 19(4), 1021-1035. [In Chinese] https://doi.org/10.3778/j.issn.1673-9418.2406053

[39] Ying, X., Zhu, Y., Chen, H., Man, F., Zhang, Z. A Multimodal Named Entity Recognition Method Based on Multi-Granularity Progressive Fusion. Application Research of Computers. 2025. [In Chinese] https://doi.org/10.19734/j.issn.1001-3695.2025.03.0071

[40] HUANG J S, QI X Y, CHENG X Y, et al. MMF-NNs: Multi-modal Multi-granularity Fusion Neural Networks for brain networks and its application to epilepsy identification [J]. Artificial Intelligence in Medicine, 2024, 157: 102990. https://doi.org/10.1016/j.artmed.2024.102990

[41] YU Q S, SHAN J Y, MA J, et al. Multi-modal and Multi-view Cervical Spondylosis Imaging Dataset [J]. Scientific Data, 2025, 12: 1080. https://doi.org/10.1038/s41597-025-05403-z.

[42] PARK H, YOUM C, CHEON S M, et al. Using machine learning to identify Parkinson's disease severity subtypes with multimodal data. Journal of Neuro Engineering and Rehabilitation, 2025, 22: 126. https://doi.org/10.1186/s12984-025-01648-2

[43] Kang, P., Kang, Y., Bi, S., Zhao, F. Current Status of Unimodal and Multimodal Research on Identifying Depression Based on Inspection and Listening and Smelling Diagnosis. Information on Traditional Chinese Medicine. 2025. [In Chinese] [Online] Available from: https://link.cnki.net/urlid/23.1194.R.20251010.1428.020

[44] Li, R., Xu, Q., Wu, Y., Liu, X. Construction of Multimodal Knowledge Graph Based on Materia Medica. Modern Computer. 2022, 28(13), 10-17. [In Chinese] https://doi.org/10.3969/j.issn.1007-1423.2022.13.002

[45] Ma, L., Cai, H., Zhang, X., Fan, Y. Transformation and Upgrading of Chinese Medicine Industry in the Context of Digital Economy: Take Anguo Digital Chinese Medicine Capital as an Example. Journal of Hebei University of Technology (Social Sciences Edition). 2024, 16(3), 1-9. https://doi.org/10.14081/j.cnki.cn13-1396/g4.000351

[46] State Administration for Market Regulation, Standardization Administration of the People's Republic of China. Artificial intelligence—Code of practice for data labeling of machine learning. GB/T 42755-2023. Beijing, China: Standards Press of China; 2023.

[47] State Administration for Market Regulation, Standardization Administration of the People's Republic of China. Health informatics—Semantic network framework of traditional Chinese medicine language system. GB/T 38324-2019. Beijing, China: Standards Press of China; 2019.

[48] Alsaad R, Abd alrazaq A, Boughorbel S, et al. Multimodal Large Language Models in Health Care: Applications, Challenges, and Future Outlook [J]. JMIR Preprints, 2024, 26: e59505. https://doi.org/10.2196/59505

[49] Li, H., Wang, Y., Liu, X., et al. Large Language Models in Chinese Medicine: Potential Limitations. Modernization of Traditional Chinese Medicine and Materia Medica-World Science and Technology. 2025, 27(10). https://doi.org/10.11842/wst.20250212001