Review of Research on Citation Objects in Scientific Papers: Postprint

Authors: Ma Na, Zhang Zhixiong, Yu Gaihong

Date: 2023-07-26T00:00:00+00:00

Abstract

[Objective/Significance] To improve the effectiveness of content-based citation analysis, this study reviews and synthesizes domestic and international research on citation objects, providing a reference for citation content analysis research. [Method/Process] Through investigating domestic and international research on citation objects, this paper systematically reviews research progress in conceptual definitions, classification systems, application domains, and automated recognition, summarizes current research limitations and gaps, and proposes future development directions. [Result/Conclusion] Citation objects evaluate the contribution and utilization value of academic literature from a semantic perspective, adding an important dimension to citation analysis methods. Research on citation objects requires advancement along three dimensions: theoretically, strengthening the investigation and analysis of multi-dimensional citation object features; technologically, exploring automated recognition methods based on large-scale corpora; and in application, developing research evaluation services based on citation objects.

Full Text

A Review of Research on Citation Objects in Scientific Papers

Ma Na[1,2], Zhang Zhixiong[1,2,3], Yu Gaihong[1,2]

[1] National Science Library, Chinese Academy of Sciences, Beijing 100190
[2] School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190
[3] Wuhan Library, Chinese Academy of Sciences, Wuhan 430071

Abstract

[**Purpose/Significance**] To enhance the effectiveness of content-based citation analysis, this paper investigates and summarizes domestic and international research on citation objects, providing a reference for citation content analysis studies.

[**Method/Process**] By surveying relevant research on citation objects, this paper reviews the conceptual definitions, classification systems, application domains, and automatic identification methods of citation objects, summarizes current research deficiencies, and proposes future development directions.

[**Result/Conclusion**] Citation objects evaluate the contribution and utilization value of academic research from a semantic perspective, adding an important dimension to citation analysis methods. Research on citation objects needs to be deepened in three directions: theoretically, by strengthening multi-dimensional feature analysis; technically, by exploring automatic identification methods based on large-scale corpora; and in application, by attempting research evaluation services based on citation objects.

Keywords: citation object; citation object identification; citation content analysis; research evaluation

1. Introduction

Scientific papers serve as important carriers of research innovation. During the writing process, authors cite others' research findings, reflecting the complex relationships among scientific works and forming the foundation of citation analysis. Traditional citation analysis methods consider only citation counts, failing to reveal the content-based connections and essential relationships between documents.

To address this limitation, scholars have begun analyzing citation content [1-4]. Citation content analysis involves mining and analyzing the semantic information of citation behavior, with existing research focusing on citation motivation identification [5-7], sentiment analysis [8-10], citation location analysis [11-12], citation content topic analysis [13-14], citation content extraction [15-17], and citation content analysis frameworks [18-19]. This approach compensates for the shortcomings of traditional citation analysis that measures quality by quantity, and some scholars believe that content-based citation analysis will become the next generation of citation analysis [1-2].

As citation analysis deepens toward content analysis, the concept of citation objects has emerged. Through investigation, we found that citation objects have multiple expressions, such as Concept Symbols, Knowledge Claims, Reference Terms, etc., with some researchers directly defining them as "citation objects" [10,19]. Despite different names, the essence remains the same. Simply put, a citation object refers to the specific scientific content being cited, directly revealing the essential connections between documents. Citation objects evaluate academic contributions and utilization value from the semantic level, adding an important dimension to citation analysis methods. Combining qualitative analysis of citation objects with traditional quantitative citation analysis based on citation frequency can improve research evaluation systems and offers a promising solution to the long-criticized metric-based evaluation approach.

Given the significance of citation objects, we conducted a comprehensive review of relevant research progress. Using English terms "citation" and "reference" combined with "object," "factoid," "terms," "concepts," "symbols," and "context," and Chinese keywords including "引文"(citation), "引用"(reference), "对象"(object), "内容" (content), "参考文献" (reference), "术语" (term), and "关键词" (keyword), we searched authoritative academic websites and databases, thoroughly reading and cross-referencing relevant literature and research reports to summarize advances in citation object research for scientific papers.

2. Conceptual Research on Citation Objects

Researchers have not yet unified the definition of citation objects, but they share a consistent understanding of their essence: objects in citation content that express the cited literature' s content, typically represented as nouns or noun phrases. Different scholars offer slightly different definitions. To trace the research lineage, we chronologically review the evolution of citation object concepts.

The earliest prototype of citation objects was proposed by H. Small in 1978, who detailed their connotation and representation. He argued that scientific content in citations represents authors' symbolic processing of cited literature, describing concepts or methods in cited papers through generalized "concept symbols" [20]. He believed these concept symbols more directly explained authors' citation behaviors than citation classification and held significant value.

When A. Ritchie et al. used terms from cited literature for index optimization, they called terms related to references in citation sentences "reference terms" and analyzed their representation forms [21]. V. Qazvinian et al., in their work on paper summarization, referred to non-overlapping phrases describing cited papers' main contributions in citation sentences as "facts" [22], later using "nuggets" [23] with unchanged meaning. Both "fact" and "nugget" represent citation objects.

M. H. Anderson et al., building on H. Small' s description, summarized "knowledge claims" from cited literature and mapped them to each citation sentence, with the mapped claims serving as citation objects [24]. In his doctoral dissertation "Sentiment Analysis in Scientific Paper Citations," Xu Deshan defined citation objects as the longest strings from cited literature appearing in citation sentences [10]. R. Jha et al., in automatic summarization research, used "factorids" about cited literature in citation sentences [25] to build summarization models. Although no explicit definition of factorid was provided, observation of their dataset reveals that factorids share the same meaning as Qazvinian' s "facts" —namely, noun terms that can express cited literature' s content.

Lu Wei et al., in their research on citation content annotation frameworks, defined citation objects as specific content objects mentioned by authors when citing literature [19]. A. Khalid et al. compared terms describing cited literature in citation sentences with topic words obtained through topic modeling algorithms, using "reference terms" to represent content related to cited literature, i.e., citation objects [26].

Synthesizing these research findings, different researchers share a consistent understanding of citation objects. We summarize the concept as follows: citation objects are generalized representations by citing authors of specific content objects in cited literature according to research needs, initiated by citing documents, with characteristics of generality and objectivity. Citation objects may be specific methods, tools, data, results, etc. (as in sentence a), or generalized summaries of methods, viewpoints, applications, or conclusions from cited literature (as in sentence b).

(a) Atomic scattering factors for the refinement of both structures were taken from Hanson, et al.

(b) Browne et al. (1993) proposed a series of statistical models of the translation process.

To further study the semantic-level citation relationships between scientific papers and their references, we propose the following basic model:

Citation Relationship = {Paper, Citation Sentence, Cited Literature, Citation Object, Citation Frequency (single, multiple), Citation Purpose, Sentiment Orientation (positive, neutral, negative)}

Where citation sentence refers to sentences in a paper describing cited literature; citation purpose represents the motivation for citing; sentiment orientation includes positive, negative, and neutral; citation frequency may be single or multiple; and citation object is the specific content object describing cited literature in the paper—the content-based connection point between paper and cited literature that directly indicates the cited literature' s utilization value and academic contribution.

Based on this model, we can clearly see the relationships between citation objects and citation purpose, sentiment orientation, etc., which helps further clarify the concept of citation objects.

3. Classification Research on Citation Objects

Classification research on citation objects proceeds from two perspectives: classification by representation form and classification by content.

Xu Deshan classified citation objects by form into named entity citation objects and abstract representation citation objects [10]. Named entity objects primarily include algorithms, terms, concepts, data, and software expressed as nouns or noun phrases, while abstract representation objects mainly represent cited literature' s viewpoints, theories, and methods as sentences or clauses. Named entity objects align with H. Small' s "concept symbols," and existing research primarily focuses on named entity objects [20-26].

Content-based classification is more complex, with no unified system yet established. Researchers categorize citation objects into 5-11 categories based on their understanding, analyzing frequency and proportion through manual annotation experiments to reveal characteristics more deeply and systematically.

M. Garzone' s classification system defines five types: "materials, equipment, tools," "formulas," "methods, procedures," "conditions, limitations," and "result analysis methods" [27]. R. Radoulov revised Garzone' s system [28], creating a classification that includes both objects representing citation purposes (e.g., "research background," "pioneering work," "related work") and specific content objects (e.g., "models," "products or materials," "data"). Lu Wei et al. further improved Radoulov's system [19], eliminating purpose-describing objects, adding concrete types like "tools" and "formulas/derivations," subdividing "methods" into "methods," "models," and "algorithms," and retaining "concepts," "data," and "results," ultimately forming an 11-category system.

These classification systems reveal that standards and granularity significantly impact classification. We argue that citation object classification should reflect essential content characteristics rather than their role or purpose when cited. Classification standards should be independent and unique; otherwise, category confusion may arise. For example, a "dataset" as a citation object might serve to represent "research background" or "existing work" from the citing paper' s perspective—mixing two classification standards would cause the same object to belong to multiple categories, yielding poor consistency. Besides differing standards, existing systems vary greatly in granularity. For instance, "methods" are subdivided into three categories by Lu Wei et al. but treated uniformly by others. Fine-grained classification aids understanding but can cause confusion, such as "Expectation Maximization" being described as either "method" or "algorithm" across different papers [19].

4. Identification Research on Citation Objects

Citation object identification is an information extraction task based on unstructured data. Authors use flexible language patterns when citing, making even manual identification difficult in some cases. Many scholars acknowledge that automatic citation object identification is a challenging task [22-23,29]. Current identification methods fall into two categories: manual annotation and computer-based automatic identification. Existing research primarily uses manual identification, with limited automatic identification studies.

4.1 Manual Identification

Manual identification involves annotating citation objects in citation sentences through human effort. We review annotation methods by H. Small, A. Ritchie, V. Qazvinian, M. H. Anderson, Lu Wei, R. Jha, and others from dataset and annotation rule perspectives, as shown in Table 1 .

Table 1. Datasets and Annotation Rules for Manual Annotation

Researcher	Domain	Dataset	Annotation Rules
H. Small	Chemistry	Top 52 highly-cited papers from 294 chemistry journals and their citation content	Mark terms or phrases most relevant to cited literature
A. Ritchie et al.	Computational Linguistics	25 ACL papers and 954 citation sentences	Annotate terms describing cited literature content in citation sentences
V. Qazvinian et al.	Computational Linguistics	1 paper and 496 citation sentences	Read all citation sentences, extract main contributions; normalize annotations across annotators and map to citation sentences
M. H. Anderson	Management	20 LDA model papers and their citation content	Read papers to summarize "knowledge claims," combine with expressions in 496 citation sentences to identify 16 main citation objects
Lu Wei et al.	Computer Science	47 papers and 2,625 citation sentences	Annotate specific content objects describing cited literature in citation content
R. Jha et al.	Computational Linguistics	1 paper and 24 citation sentences	Annotate facts about cited literature in citation sentences (excluding those appearing only once)

Regarding datasets, since no open citation object datasets exist, scholars primarily use self-built datasets or full-text scientific literature datasets, mainly comprising cited literature and citation content. Due to manual annotation costs, dataset sizes are limited, with citation object quantities ranging from dozens to thousands. The Association for Computational Linguistics Anthology Network corpus (ANN) [29], widely used in citation content analysis, provides full-text papers from authoritative journals and top conferences in natural language processing with rich citation networks.

Regarding annotation rules, basic rules show high consistency—annotating content related to cited literature in citation sentences—but strategies vary by application need. H. Small, A. Ritchie, Lu Wei, and R. Jha directly annotate citation objects in each citation sentence. V. Qazvinian clusters all citation sentences for the same cited literature for unified annotation. M. H. Anderson summarizes all citation objects from all citation sentences first, then assigns them to each sentence. These strategies yield different data forms: the first produces rich expressions directly from citation sentences, while the latter two yield more standardized forms suitable for deep citation analysis.

4.2 Automatic Identification

Automatic identification via computer technology is in its infancy. R. Radoulov, Xu Deshan, and A. Khalid have explored classification, identification, and alternative identification approaches.

R. Radoulov used machine learning for citation object classification, incorporating paper structure features, citation location, part-of-speech features, syntactic features, and two types of cue words to train Naive Bayes classifiers for each object type [28]. Using 5,583 citation sentences, results showed that for nine categories, "specific background," "data," "methods," and "products" achieved good classification performance with over 70% accuracy and F-values between 0.7-0.5. The remaining five categories ("general background," "related work," "concepts," "historical notes," and "pioneering work") performed poorly with F-values around 0.3, which Radoulov attributed to insufficient data (these categories comprised only 10% of the dataset). This also reveals that citation objects concentrate in these four categories.

Xu Deshan identified citation objects when analyzing citation sentiment links by calculating string similarity between citation sentences and cited literature, extracting the longest common string as the citation object [10]. Due to difficulty obtaining full cited texts, matching was limited to titles. Using 61 scientific papers with 1,076 unique citation markers, the method correctly identified 705 citation objects (65.52% accuracy). Xu also attempted normalization by calculating weighted sums of morphological similarity, string length, and semantic similarity to determine if two citation objects for the same paper represented the same object—a process that directly affects application effectiveness.

A. Khalid identified topics from all citation sentences for a cited paper as an alternative to direct citation object identification [26]. From 113 ANN papers and 9,970 citation sentences, a computational linguist annotated terms related to cited literature (ambiguous cases were discarded). LDA topic modeling was then applied to each paper' s citation sentence set, with word clouds assigning high-frequency terms as topic labels, merging similar topics. Evaluation showed topic labels mostly matched manually annotated reference terms (i.e., citation objects), though exceptions occurred when citation sentences contained multiple objects.

Overall, automatic identification remains limited. We attribute this to two factors: (1) citation objects are authors' paraphrases of cited content, making them highly subjective; (2) lack of large-scale, high-quality evaluation corpora hinders research progress. Technically, citation object identification can be formulated as a sequence labeling problem, potentially applying models like HMM [30], CRF [31], LSTM-CNNs-CRF [32], and BERT [33] that have proven effective. For datasets, constructing large-scale, high-quality datasets from scientific literature—structured data comprising papers, cited literature, citation sentences, and citation objects—would provide clear structure and accessible samples. Given limited training data, semi-supervised [34] or transfer learning [35] strategies could reduce data requirements.

5. Applications of Citation Objects in Citation Analysis and Text Mining

Citation objects represent others' citations of research findings, possessing knowledge representation and dissemination features. Their utilization falls into two categories: (1) combining with citation frequency for citation analysis from knowledge propagation and inheritance perspectives, and (2) incorporating citation objects as important features into text mining algorithms to improve literature retrieval and automatic summarization. Current applications rely on manually annotated data, limiting scale and remaining exploratory.

Scientific literature' s main contributions can be revealed through academic community citations and descriptions. Citation objects naturally carry this information, offering inherent advantages for citation analysis and research evaluation. M. H. Anderson et al. analyzed a highly-cited management paper' s contributions using citation objects and citation frequency [24]. Their annotation process aggregated all citation sentences for the highly-cited paper, summarized main contributions, then mapped these back to each citation sentence as its citation object—simultaneously standardizing expressions for deeper analysis. By combining annotated citation objects with frequency counts, they identified the paper' s main contributions and compared its impact across seven disciplines, finding no significant differences. Management and computer science showed similar top three citation objects. This combined content-quantity approach compensates for traditional citation analysis limitations, offering new perspectives for scientometrics and research evaluation.

Beyond enhancing citation analysis, citation objects reflect other researchers' summaries of cited work, making them valuable for literature retrieval, automatic summarization, and sentiment analysis. A. Ritchie et al. demonstrated that reference terms in citation sentences improve index term construction [21]. V. Qazvinian et al. used citation objects to identify key contributions and remove duplicate content for more accurate and complete automatic summarization [22]. R. Jha et al. improved summarization by weighting citation objects by frequency [25]. Xu Deshan built citation object discourse links for citation sentiment analysis by first identifying and disambiguating objects [10].

6. Summary and Outlook

Citation objects reveal cited literature' s academic value from the content level, compensating for traditional quantity-based citation analysis and holding important theoretical and practical significance. This paper reviews citation object research from theoretical, application, and technical perspectives, showing that while concepts are clarifying and classification systems emerging, significant deficiencies remain in automatic identification and application. Overall, citation object research is in its early exploratory stage, far from achieving large-scale application goals, with much work remaining.

Theoretically, multi-dimensional feature analysis should be strengthened to clarify citation objects' essence. In early research stages, theoretical foundations require gradual improvement. As generalized knowledge representations, citation objects exhibit statistical features including part-of-speech, syntactic roles in citation sentences, positional relationships with citation markers, and language model characteristics. Linguistic analysis will deepen objective understanding. Additionally, analyzing citation object characteristics across disciplines is valuable. Limited by full-text availability, most existing research focuses on computer science; whether features generalize to other fields remains unknown, making cross-disciplinary comparative studies necessary.

Technically, large-scale corpus-based automatic identification methods should be explored. Current automatic identification achievements are limited. We suggest formulating it as a sequence labeling problem using effective models like HMM, CRF, LSTM-CNNs-CRF, and BERT. Constructing large-scale, high-quality evaluation corpora from scientific literature is essential. Semi-supervised or transfer learning can address limited training data.

In application, automation technology can enhance semantic citation analysis and enable citation object-based research evaluation services. Citation objects embody papers' academic value and contributions—unique among citation content features. Combining qualitative citation object analysis with quantitative citation frequency methods to improve research evaluation systems represents the most fundamental and important application direction. Improved automatic identification will lay foundations for deep content-based citation analysis, making citation object-based evaluation services feasible while revealing technical improvement needs through applications. Large-scale citation objects can build semantic entity-level citation networks for rapid contribution analysis, mining classic theories, important algorithms, or widely-used datasets, exploring interdisciplinary intersections, and constructing knowledge flow paths.

References

[1] DING Y, ZHANG G, CHAMBERS T, et al. Content-based citation analysis: the next generation of citation analysis[J]. Journal of the Association for Information Science and Technology, 2014, 65(9): 1820-1833.
[2] Zhao Rongying, Zeng Xianqin, Chen Bikun. Full-text citation analysis—a new development in citation analysis[J]. Library and Information Service, 2014, 58(9): 129-135.
[3] Liu Liu, Wang Dongbo. A review of citation content analysis research[J]. Journal of the China Society for Scientific and Technical Information, 2017, 36(6): 637-643.
[4] Hu Zhigang. Full-text citation analysis[M]. Beijing: Science Press, 2016.
[5] GARFIELD E. Can citation indexing be automated?[J]. National Bureau of Standards, 1964, 269(114): 189-192.
[6] TEUFEL S, SIDDIHARTHAN A, TIDHAR D. Automatic classification of citation function[C]//Proceedings of the 2006 conference on empirical methods in natural language processing. Sydney: Association for Computational Linguistics, 2006: 103-110.
[7] MUNKHDALAI T, LALOR J, YU H. Citation analysis with neural attention models[C]//Proceedings of the seventh international workshop on health text mining and information analysis (LOUHI). Austin: Association for Computational Linguistics, 2016: 69-77.
[8] ATHAR A, TEUFEL S. Context-enhanced citation sentiment detection[C]//Proceedings of the 2012 conference of the North American chapter of the association for computational linguistics: human language technologies. Montreal: Association for Computational Linguistics, 2012: 597-601.
[9] YU B, FANELLI D. Classifying negative findings in biomedical publications[J]. Association for Computational Linguistics, 2014: 19-23.
[10] Xu Deshan. Sentiment analysis in scientific paper citations[D]. Beijing: National Science Library, Chinese Academy of Sciences, 2012.
[11] DING Y, LIU X, GUO C, et al. The distribution of references across texts: some implications for citation analysis[J]. Journal of informetrics, 2013, 7(3): 583-592.
[12] Hu Zhigang, Chen Chaomei, Liu Zeyuan, et al. From citation-based to citation-based: a new method for counting total citations[J]. Library and Information Service, 2013, 57(21): 5-10.
[13] LIU S, CHEN C. The differences between latent topics in abstracts and citation texts of citing paper[J]. Journal of the American Society for Information Science and Technology, 2013, 64(3): 627-638.
[14] Zhu Qingsong, Leng Fuhai. Identifying topics of highly-cited papers based on citation content analysis[J]. Journal of Library Science in China, 2014, 40(1): 39-49.
[15] ANGROSH M A, CRANEFIELD S, STANGER N. Conditional random fields based sentence identification: enhancing citation context ex-

traction[C]//Proceedings of the 2014 Society for Industrial and Applied Mathematics international conference on data mining. Pennsylvania: Society for Industrial and Applied Mathematics, 2014: 361-369.
[16] SONDHI P, ZHAI C X. A constrained hidden Markov model approach for non-explicit citation context extraction[C]//Proceedings of the first Australasian Web conference-volume 144. Adelaide: Australian Computer Society, 2013: 59-68.
[17] Lei Shengwei, Chen Haihua, Huang Yong, et al. Research on automatic identification of citation contexts in academic literature[J]. Library and Information Service, 2016, 60(17): 78-87.
[18] ZHANG G, DING Y, MILOJEVIC S. Citation content analysis (CCA): a framework for syntactic and semantic analysis of citation content[J]. Journal of the American Society for Information Science and Technology, 2013, 64(7): 1490-1503.
[19] Lu Wei, Meng Rui, Liu Xingbang. A citation content annotation framework for citation relationships[J]. Journal of Library Science in China, 2014, 40(6): 93-104.
[20] SMALL H. Cited documents as concept symbols[J]. Social studies of science, 1978, 8(3): 327-340.
[21] RITCHIE A, TEUFEL S, ROBERTSON S. How to find better index terms through citations[C]//Proceedings of COLING/ACL workshop on how can computational linguistics improve information retrieval. Sydney: Association for Computational Linguistics, 2006: 25-32.
[22] QAZVINIAN V, RADEW D R. Scientific paper summarization using citation summary networks[C]//Proceedings of the 22nd international conference on computational linguistics. Manchester: Association for Computational Linguistics, 2008: 689-696.
[23] QAZVINIAN V, RADEW D R, OZGUR A. Citation summarization through keyphrase extraction[C]//Proceedings of the 23rd international conference on computational linguistics. Beijing: Association for Computational Linguistics, 2010: 895-903.
[24] ANDERSON M H, SUN Y T. What have scholars retrieved from Walsh and Unson (1991)? a citation context study[J]. Management learning, 2010, 41(2): 131-145.
[25] JHA R, ABU-JABARA A, RADEW D. A system for summarizing scientific topics starting from keywords[C]//Proceedings of the 51st annual meeting of the Association for Computational Linguistics. Sofia: Association for Computational Linguistics, 2013: 572-577.
[26] KHALID A, KHAN A F A, IMRAN B M, et al. Reference terms identification of cited articles as topics from citation contexts[J]. Computers and electrical engineering, 2018, 20(3): 1-12.
[27] GARZONE M, MERCER R E. Towards an automated citation classifier[C]//Advances in artificial intelligence. Berlin: Lecture notes in computer science, 2000: 337-346.
[28] RADOULOV R. Exploring automatic citation classification[D]. Waterloo: University of Waterloo, 2008.

[29] RADEW D R, MUTHUKRISHNAN P, QAZVINIAN V, et al. The ACL anthology network corpus[J]. Language resources and evaluation, 2013, 47(4): 1-26.
[30] BIKEL D M, MILLER S, SCHWARTZ R, et al. Nymble: a high-performance learning name-finder[C]//Proceedings of the fifth conference on applied natural language processing. Washington: Association for Computational Linguistics, 1997: 194-201.
[31] LAFFERTY J D, MCCALLUM A, PEREIRA F C N. Conditional random fields: probabilistic models for segmenting and labeling sequence data[C]//Eighteenth international conference on machine learning. Williamsport: Morgan Kaufmann Publishers Inc, 2001: 282-289.
[32] MA X, HOYY E. End-to-end sequence labeling via bi-directional LSTM-CNNs-CRF[C]//Proceedings of the 54th annual meeting of the Association for Computational Linguistics. Berlin: Association for Computational Linguistics, 2016: 1064-1074.
[33] DEVLIN J, CHANG M W, LEE K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding[EB/OL]. [2018-12-13]. https://arxiv.org/abs/1810.04805.
[34] PETERS M E, AMMAR W, BHAGAVATULA C, et al. Semi-supervised sequence tagging with bidirectional language models[C]//Proceedings of the 55th annual meeting of the Association for Computational Linguistics. Vancouver: Association for Computational Linguistics, 2017: 1756-1765.
[35] YANG Z L, SALAKHUTDINOV R, COHEN W W. Transfer learning for sequence tagging with hierarchical recurrent networks[C]//Proceedings of the 5th international conference on learning representations. Toulon: The DBLP Computer Science Bibliography, 2017.

Author Contributions:

Ma Na: Literature review and manuscript writing
Zhang Zhixiong: Manuscript structure and writing guidance
Yu Gaihong: Manuscript revision

Note: Figure translations are in progress. See original paper for figures.