

<https://doi.org/10.1038/s40494-026-02511-w>

Oracle bone inscriptions information processing: a comprehensive survey



Zijian Chen^{1,2}✉, Wenjie Hua³, Jinhao Li⁴, Yucheng Zhu⁵, Xiaona Zhi⁶, Zhiji Liu⁷, Tingzhu Chen⁸✉, Wenjun Zhang¹ & Guangtao Zhai^{1,2}✉

This paper presents a comprehensive, task-centric synthesis that reframes the development of Oracle bone inscriptions (OBI) information processing as a coevolution between methodologies and their underlying data substrate. Critical challenges and potential research directions are discussed for building trustworthy, efficient OBI information processing systems that function as true partners for frontline researchers, enabling them to contribute to key debates on the social structure and periodization of the pre-Qin period.

Oracle Bone Inscriptions (OBI) represent the earliest known system of ancient Chinese writing, dating primarily to the late Shang dynasty (c. 1300–1046 BCE), which were typically incised on turtle plastrons and ox scapulae for pyromantic divination¹. Studies on OBI offer unparalleled insights into the phonology, morphology, and semantics of archaic Chinese. Beyond their profound linguistic value, the divinatory texts constitute an invaluable primary source for understanding the history, religion, governance, and societal structure of China's Bronze Age.

The methodology of OBI processing has always been a unique and formidable set of challenges rooted in the artifacts' physical nature and the script's intrinsic properties. Foremost among these is the severe fragmentation of the OBI. The majority of bones and plastrons were broken after thousands of years of natural weathering and man-made destruction, scattering related texts across massive fragments. The monumental task of fragment rejoining, a complex, high-dimensional puzzle, is often a prerequisite for coherent semantic analysis, which, prior to the popularization of image processing technologies, relied exclusively on decades of painstaking manual assembly by oracle bone experts². Moreover, at the paleographic level, the script is characterized by a high degree of allography, with a single character possessing numerous morphological variants, and glyph similarity, where distinct oracle bone characters bear a strong visual resemblance³. Meanwhile, the specific noises in the OBI rubbings, including stroke-broken, bone-cracked, abnormal edges, and dense white regions⁴, further challenge the recognition of oracle bone characters. Lastly, the ultimate goal of OBI processing is to decipher them. Since the first discovery of oracle bones in 1899, over 4,500 character classes have been excavated, with only about one-third of them successfully deciphered or interpreted. One major obstacle lies in the scarcity of the corpus and the narrow contextual scope. The vast majority of OBI texts are divinatory records, which

are highly formulaic and restricted in semantic space. A related difficulty is the pictographic and ideographic nature of many graphs. While their *iconographic logic* is sometimes apparent, the original visual metaphor is often lost to time or stylized beyond intuitive recognition. This forces their decipherment to rely heavily on comparative paleography, i.e., tracing a glyph's evolution forward to more stable forms in bronze inscriptions or bamboo slips, which is often speculative⁵.

Driven by advances in artificial intelligence (AI) technology, especially the proliferation of large generative models, the field of OBI processing is continuously evolving and also changing traditional OBI experts' perspectives. In recent years, a large number of OBI papers that cross various techniques have been published (Fig. 1), which, to a certain extent, may obscure the truly pivotal points. Due to interdisciplinary issues, the key questions that the archeological linguists or historians focus on may have a large gap with computer scientists or AI algorithm experts. Therefore, conducting a thorough survey in the OBI field is of great significance, which will help gain a comprehensive understanding of the relevant literature and shed light on the future development direction.

Related surveys

Although a large number of OBI studies have been carried out, comprehensive review articles that synthesize and critically analyze these works remain conspicuously scarce. Table 1 summarizes the existing surveys in the field. Xing et al.⁶ provided an overview of OBI detection models, such as Faster R-CNN⁷, SSD⁸, YOLO⁹, RFBnet¹⁰, and RefineDet¹¹, to select an appropriate architecture for OBI detection that achieves the speed/memory/accuracy balance for a given platform. Diao et al.¹² gave a survey on ancient script image recognition methods in 2025. They categorized existing studies by script type and analyzed the corresponding recognition methods,

¹Institute of Image Communication and Information Processing, Shanghai Jiao Tong University, Shanghai, China. ²Shanghai Artificial Intelligence Laboratory, Shanghai, China. ³School of Chinese Language and Literature, Wuhan University, Wuhan, China. ⁴School of Computer Science and Technology, East China Normal University, Shanghai, China. ⁵USC-SJTU Institute of Cultural and Creative Industry, Shanghai Jiao Tong University, Shanghai, China. ⁶Institute of History, Chinese Academy of Social Sciences, Beijing, China. ⁷Center for the Study and Application of Chinese Characters, East China Normal University, Shanghai, China. ⁸School of Humanities, Shanghai Jiao Tong University, Shanghai, China. ✉e-mail: zijian.chen@sjtu.edu.cn; tingzhuchen@sjtu.edu.cn; zhaiguangtao@sjtu.edu.cn

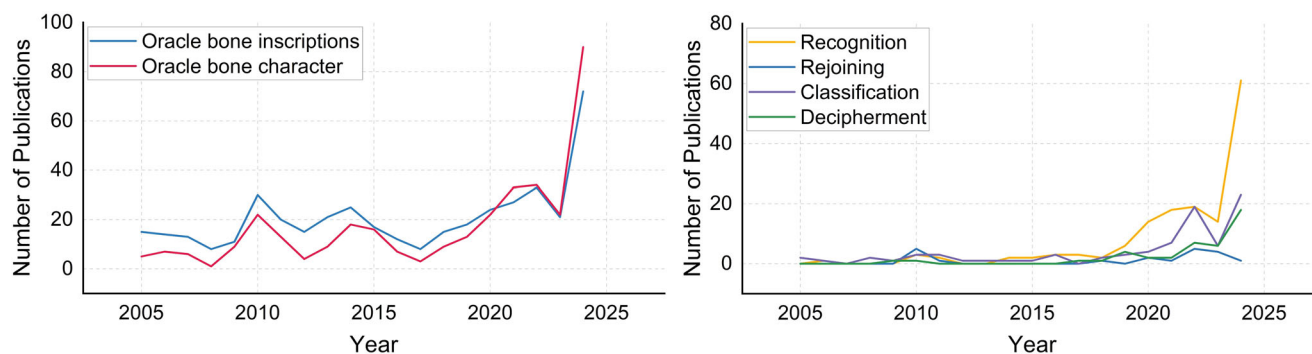


Fig. 1 | Trend of publications indexed by the Web of Science database, whose topic, titles, or abstracts mention the keyword “Oracle bone inscriptions”, “Oracle bone character” (left) or the combination of OBI and four mainstream tasks (i.e., recognition, rejoining, classification, and deciphering) (right). Results cover the last twenty years, from 2005 to 2024. We can observe a relatively stable

publication output before 2017. Between 2017 and 2022, as deep learning techniques gained widespread adoption, the number of OBI studies increased steadily. More recently, with the advancement of large language models, the volume of relevant articles has shown rapid acceleration, reflecting growing interdisciplinary interest in large language models across different OBI tasks.

highlighting both their differences and shared strategies. Challenges unique to ancient script recognition and recent solutions, including few-shot learning and noise-robust techniques, are also reviewed. Recently, Li et al.¹³ presented a structured survey of the current landscape of oracle character recognition (OrCR) research, including an overview of the primary benchmark datasets and digital resources available for OrCR, as well as a review of contemporary research methodologies with their respective utilities, limitations, and applicability.

To the best of our knowledge, the aforementioned articles are the only existing surveys related to the OBI research. However, they have narrowly focused on a single facet, i.e., the recognition task, overlooking the systematic integrity and intrinsic interdisciplinarity of OBI studies. Some key tasks and new emerging areas in OBI research are not included. For example, OBI dataset construction, OBI rejoining, OBI classification, OBI deciphering, and many related up-to-date AI techniques have all been intensively researched in recent years. These emerging topics represent new trends in OBI processing, but they are missing from almost all current OBI surveys.

Considering the limitations of the current OBI surveys described above, a comprehensive and up-to-date survey, including not only traditional OBI processing measures but also whole-process OBI processing tasks from excavation to digital preservation, is greatly needed for a better top-down understanding of the history, current state-of-the-art, and future trend of the field. This survey is written to fulfill that need.

Survey contribution and novelty

This survey makes the following unique contributions:

- **Comprehensive Review:** It provides a thorough review of the existing OBI processing techniques with surveyed OBI papers several times more than those of the previous surveys, highlighting both data and methodology perspectives across different tasks.
- **Integrating Emerging Technologies:** It situates current OBI processing methods within the context of emerging technologies such as deep learning, generative models, and large language models.
- **Identification of Challenges:** It discusses the key challenges and limitations in the current state of OBI research, especially the applicability of the current evaluation approaches.
- **Guiding Future Research:** Several promising research directions for future works are identified and detailed, encompassing a variety of aspects such as text-to-OBI generation, specific evaluation metric design, multimodal learning for OBI deciphering, and 3D reconstruction of OBI fragments.

Survey methodology and organization

To ensure a holistic and systematic review of the literature on OBI processing, a clear and structured survey methodology is important. The methodologies

for conducting a thorough survey, including literature search, selection criteria, taxonomy, analysis, and synthesis, are summarized as follows.

- **Literature search:** Considering the several major fields related to OBI processing, such as archeology, philology, computer science, signal processing, and linguistics, etc., we utilize mainstream academic databases such as Google Scholar, IEEE Xplore, ACM Digital Library, SpringerLink, and Web of Science as the searching platform. Besides, we focus on prominent international and Chinese journals in the fields of multimedia, computer vision, artificial intelligence, and historic preservation, such as IEEE Transactions on Image Processing, Pattern Recognition, npj Heritage Science, and conferences like ICLR, ACM MM, CVPR, ICCV, ACL, AAAI, etc. Moreover, we use a combination of keywords related to OBI, such as “oracle bone inscriptions”, “oracle character recognition”, “rejoining oracle bone inscriptions”, “oracle bone inscription datasets”, “deep learning in OBI processing”, and “deciphering OBI” to search the relevant papers.
- **Selection criteria:** The inclusion criteria include topic relevance, methodology, novelty, impact, etc. The exclusion criteria include language, study quality, publication type, redundancy, etc.
- **Taxonomy:** We categorize the papers according to the OBI datasets, OBI processing-related techniques, and task-specific OBI processing works.
- **Analysis:** We present a systematic review of OBI information processing studies, charting the evolution from traditional methodologies to modern AI-driven approaches. We conduct a task-by-task introduction and analysis while highlighting the distinct characteristics of the OBI datasets derived from these tasks.
- **Synthesis:** Finally, based on the current challenges in the digitization, preservation, and interpretation of paleography, and in conjunction with the technological trajectory of AI, we overview future research and development trends in this field.

We confine our survey pool to encompass technical papers published in the last 20 years, i.e., 2005–2025. Note that this does not limit the citation and reference of the work of early experts and scholars, considering the long-lasting characteristic of OBI research. We include more OBI tasks and related papers that are considered important but have not been included in existing OBI surveys. The taxonomy of this survey is shown in Fig. 2. First, we outline four key stages in the development of OBI research, each embodying distinct methodology designs, data formats, and technical challenges. This historical perspective not only situates the current state of the field but also clarifies the motivations for the directions explored in later sections of this survey. Second, we review the existing OBI datasets for specific tasks. Third, we summarize mainstream OBI preprocessing methodologies, including traditional image processing and generative model-based data augmentation, and review OBI processing methods for different

Table 1 | Comparison of related surveys along with the published year, the number of referenced OBI papers, topics, challenges, and key contributions

Surveys	Xing et al. ⁶	Diao et al. ¹²	Li et al. ¹³	Ours
Year	2019	2025	2025	2026
#OBI Paper	11	30	69	113
Topic	OBI Detection, Deep Learning Algorithm	Ancient Script Recognition, OCR, Deep Learning, Data Imbalance	OBI Recognition, Dataset	OBI Recognition, Rejoining, Classification, Retrieval, Deciphering, Dataset, Evaluation
Challenges	Data noise, fragment, distribution, and variant; Synthesis of OBI noise image	Data scarcity, distribution, and noisy imagery; Need for semantically-aware denoising methods	Intrinsic variability; Quantity limitations of available data; Degraded image quality	Data source barriers, quality, distribution; Corpus-level datasets; Dataset construction pipelines; Heterogeneous OBI data representation; Objective evaluation criteria
Contributions	Compare the performance of object detection algorithms in the OBI dataset	Review recent advances in ancient script recognition; Highlight key methodological distinctions; Emphasize challenge of noisy image recognition	Overview of oracle character recognition works; Outline the key challenges and commonly used datasets	Introduces a detailed taxonomy to categorize OBI information processing works; Discuss the applicability of the evaluation metrics; Highlight the recent trends, challenges and outline future directions

tasks, including OBI recognition, OBI rejoining, OBI classification, and OBI deciphering. Fourth, the evaluation of OBI processing methods in different tasks is discussed, with a comparison of their advantages and disadvantages. Fifth, we summarize challenges and outlooks for future trends in OBI processing, and finally, we conclude the whole paper.

Evolving paradigms of OBI processing and discussion

The evolution of OBI information processing has undergone a series of significant paradigm shifts (Fig. 3), reflecting a deeper integration of frontier AI techniques with expert-calibrated cognitive processes. In this section, we outline four key stages in the development of OBI processing studies, each embodying distinct mainstream methodologies, effectiveness analyses, and technical challenges. This historical perspective not only situates the current state of the field but also clarifies the motivations for the directions explored in later sections of this survey.

Stage 1: Traditional expert-led methodologies

In the early days, the processing of OBI was conducted mainly by manually collating and proofreading ancient books. Scholars such as Yirong Wang, Guowei Wang, Zhenyu Luo, Moruo Guo, Houxuan Hu, Xigui Qiu, Tianshu Huang, Dekuan Huang, and Zhenhao Song have made outstanding contributions in both theoretical^{14,15} and practical^{16,17} aspects of the field. At that time, after being unearthed, the publication of oracle bone materials typically involved reproducing the inscriptions through ink rubbings, tracings, or photographs, and then compiling these reproductions into printed volumes. Among them,¹⁸ remains the most comprehensive catalog to date, containing the largest number of oracle bones ever published, complete with plates, transcriptions, and indexes. Based on these, the deciphering and interpretation of the script constitute the primary task of early oracle bone studies.

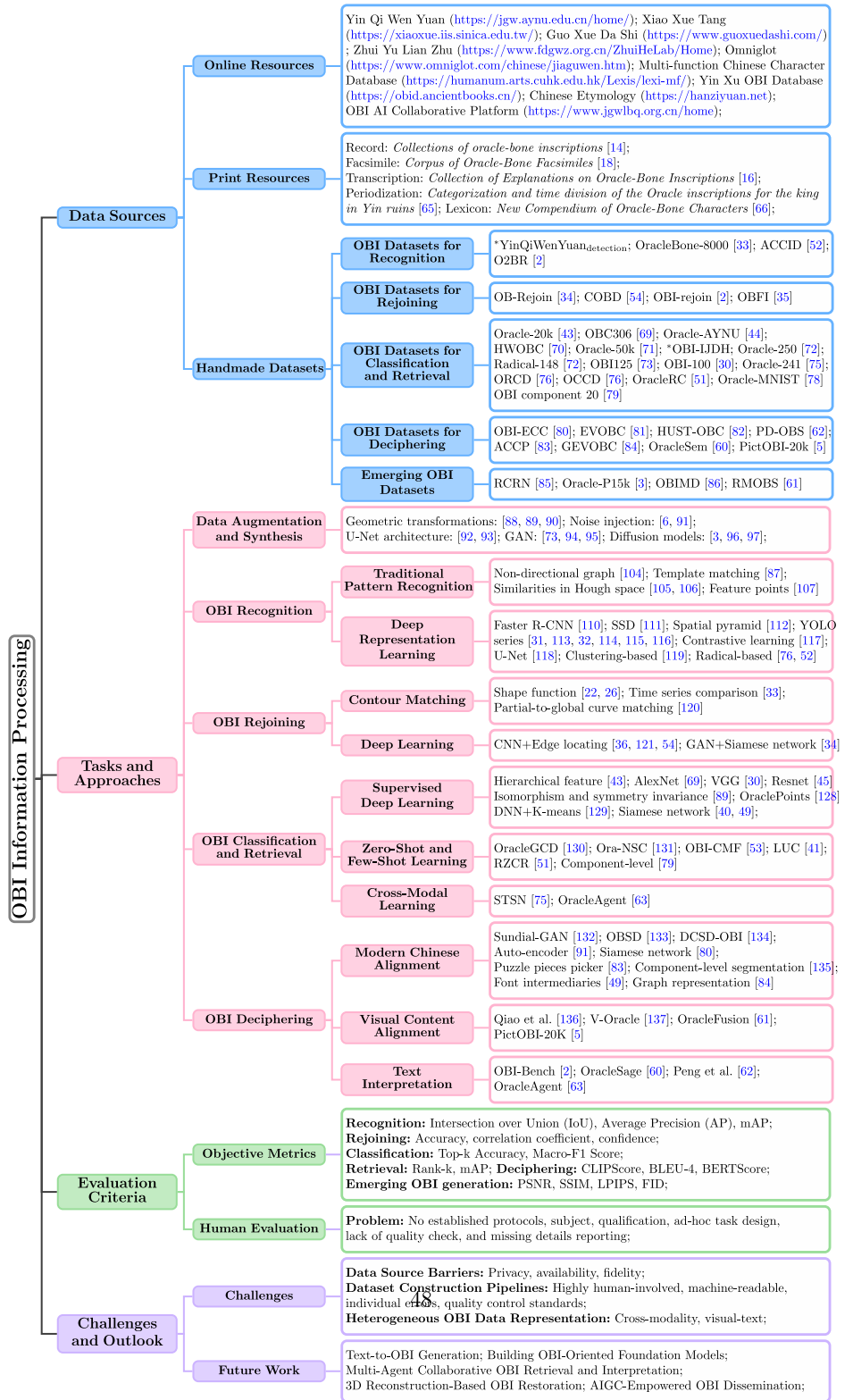
Beyond paleographic deciphering, significant progress has also been made in the chronological and periodization studies of oracle bone inscriptions. Zuobin Dong¹⁹ made a landmark contribution, dividing 273 years of oracle bone materials into five chronological phases. Later, modern scholars revised this framework by first classifying the inscriptions based on script features and other characteristics and then determining the temporal affiliation of each class, proposing a new *two-lineage* hypothesis to replace the traditional *five-period* model¹⁸. Other specialized studies, such as the research on the graphical system and the morphological formation system of OBI^{20,21}, encompass both intrinsic analyses of ancient character systems and diachronic explorations, contributing significantly to our understanding of the structural and evolutionary mechanisms of early Chinese writing.

However, the traditional expert-led methodology that dominated early oracle bone research, while foundational, was inherently constrained by human limitations. The primary drawback was a profound lack of efficiency. Progress was painstakingly slow, as every task, from deciphering individual characters and transcribing inscriptions to manually searching for and matching fragmented pieces, depended entirely on the laborious, time-consuming efforts of a few highly trained scholars. Furthermore, this approach suffered from low reproducibility and high subjectivity. Interpretations often relied heavily on an individual expert's intuition, memory, and accumulated private knowledge. This tacit knowledge was difficult to formalize or transfer, making it challenging for other researchers to independently verify findings or systematically build upon previous work, creating a significant bottleneck for scaling research and validating results objectively.

Stage 2: Computer-aided digital image processing

This stage, from the 1990s to around 2012, marked a transformative era for digital image processing. Research in the OBI domain primarily focused on image enhancement²², restoration (e.g., histogram equalization and spatial filtering), and introduced local feature descriptors, such as scale-invariant feature transform (SIFT) and histogram of oriented gradients (HOG)²³ (<http://digital.library.wisc.edu/1793/67547>). These algorithms allowed computers to improve OBI image quality for downstream tasks and identify distinctive structures across images despite changes in scale, rotation, or illumination.

Fig. 2 | Taxonomy of the data sources, methodologies, evaluation criteria, challenges, and outlook in OBI information processing. The website reference of YinQiWenYuan_{detection} and OBI-IJDH are (<https://jgw.aynu.edu.cn/home/down/detail/index.html?sysid=3>) and (<http://www.ihpc.se.ritsumei.ac.jp/OBIdata/IJDH.zip>), respectively.



Wang et al. (<https://nlpr.ia.ac.cn/2006papers/gjhy/gh75%20.pdf>) utilized the wavelet transform to remove noise of different defective rubbings of oracle bone. In²⁴, the author proposed a mesh point feature extraction algorithm for neat oracle bone rubbings, which utilizes coarse mesh relative address, resulting in a decent recognition performance improvement. Liu et al.²⁵ presented a local boundary descriptor with a two-dimensional fragment contour matching algorithm based on angular sequence. Wang et al.²⁶ proposed an Euclidean distance-based contour matching approach for extracting and

tracking contour information from oracle shell images. Furthermore, at a time when open-source awareness was still in its infancy, demonstrating data ownership and integrity, especially for sensitive data such as OBI, was crucial for safeguarding cultural heritage. Liu et al.²⁷ designed a wavelet watermarking embedding and extraction pipeline for OBI images, which exhibits strong robustness to photo brush tampering and image clipping.

Nevertheless, algorithms at this stage are predominantly susceptible to noise and degradation, which often struggle to distinguish genuine character

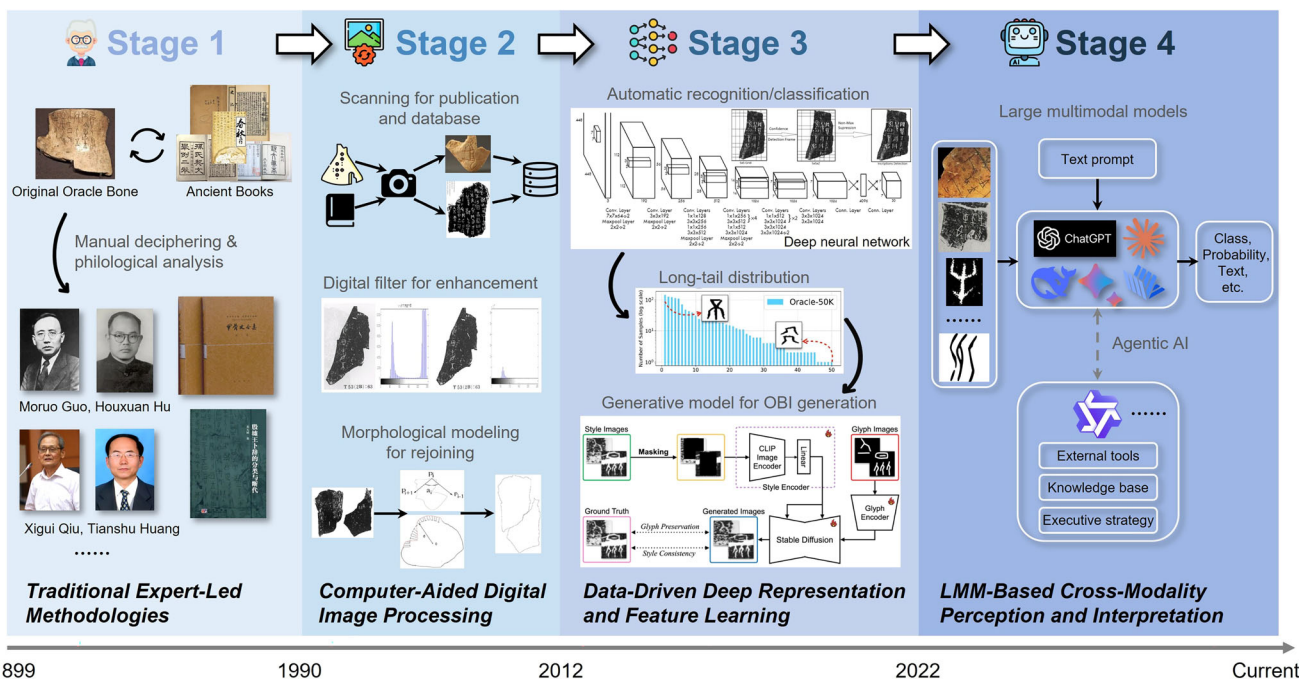


Fig. 3 | Evolution of OBI information processing methods from the last century to 2025 includes four paradigm shifts, i.e., the traditional expert-led methodologies, computer-aided digital image processing, data-driven deep representation enabling higher-level learning, and the latest cross-modality perception and interpretation marked by knowledge integration in large multimodal models. We show some representative research directions in each stage.

strokes and the stochastic background noise common in rubbings, such as cracks, erosion, and uneven ink distribution. Furthermore, this stage suffered from a profound semantic gap. These methods treated characters merely as rigid geometric patterns rather than semantic entities, relying on hand-crafted features that lacked the representational capacity to handle intra-class variance.

Stage 3: Data-driven deep representation and feature learning

The advent of deep convolutional neural networks (DCNNs)^{7,28,29} marked a pivotal shift in OBI processing: moving from expert-centric manual procedures to data-driven automatic frameworks. These models achieved strong performance in tasks such as OBI character recognition^{30–32}, rejoining^{33–36}, classification^{37–39}, and retrieval^{40–42}, laying the foundation for the subsequent construction of large-scale OBI databases.

This stage of research begins to combine low-level directional features with mid- and high-level encodings within a CNN framework to better capture the stroke characteristics of OBI characters^{43–45}, which is further improved by Transformer architectures^{46–48}. This shift enables the modeling of long-range structural dependencies, which is particularly crucial for OBI, where self-attention mechanisms allow models to infer complete glyph structures despite significant physical degradation or fragmentation. A similar multiscale feature integration scheme⁴⁹ was also proposed to align features across disparate domains to facilitate cross-modal OBI retrieval. Meanwhile, to combat the severe data scarcity of rare characters, adversarial data augmentation (e.g., AGTGAN⁵⁰, Oracle-P15k³) was employed to generate synthetic samples, transferring styles between handprinted glyphs and authentic rubbings. For zero-shot and open-set recognition, frameworks like RZCR⁵¹ and ACCID⁵² detect sub-character components (*radicals*) first and utilize knowledge graphs or DCNNs to infer the identity of novel characters based on structural composition rules. Then, with the emergence of self-supervised learning (e.g., pre-training on large-scale, unlabeled ancient script corpora), models can now acquire robust paleographic intuition without relying solely on scarce expert-annotated labels, effectively mitigating the long-tail distribution problem in OBI datasets⁵³. In the OBI rejoining task, researchers^{34,54} employed a dedicated deep neural network to serve as a binary classifier to verify texture continuity after preliminary edge matching.

Although significant progress has been made at this stage, there still exist many challenges. First, the distribution of OBI data is extremely imbalanced. Most undeciphered characters often appear only once or survive solely in fragmentary rubbings. Existing deep learning models readily overfit or completely fail on tail categories. Although generative AI-based data augmentation techniques have been exploited to perform style transfer between handprinted sketches and rubbings, the quality and diversity of synthetic samples still fall short of faithfully substituting genuine exemplars. Second, current recognition models are predominantly closed-set systems, which are prone to making mistakes when confronted with unknown characters. Endowing the classifier with a principled rejection capability and enabling it to infer unseen characters through reasoning constitutes the foremost technical bottleneck. Furthermore, models in this era predominantly rely on unimodal visual features, lacking comprehension of the text semantic content within divinatory inscriptions, which hinders further decipherment of OBI characters.

Stage 4: Large multimodal model-based cross-modality perception and interpretation

Recently, the rapid advancement of large multimodal models (LMMs)^{55,56} (<https://openai.com/gpt-5/> and <https://blog.google/products/gemini/gemini-3/>) has led to a paradigm shift in many fields^{57–59}, exhibiting unprecedented potential with capabilities in task automation, methodological innovation, and interactive application development. These models have fundamentally transformed OBI research by integrating multiple modalities, enabling cross-modality perception and interpretation for OBIs. The maturation of vision-language alignment (e.g., contrastive learning) allows the field to move beyond isolated character recognition toward contextual decipherment, where the model integrates visual glyph features with divinatory textual semantics, closely mimicking the multi-source evidentiary approach used by human experts. As a prominent example of this trend, OBI-Bench² has for the first time evaluated 23 LMMs in five main OBI processing tasks, including recognition, rejoining, classification, retrieval, and deciphering, and demonstrated the potential of LMMs in processing tasks demanding expert-level domain knowledge and cognition. Afterwards, specific visual-language OBI models fine-tuned from current foundation models, such as OracleSage⁶⁰, OracleFusion⁶¹, and Peng et al.’s model⁶²,

extract hierarchical visual information (*glyph structure*) to assist the reasoning reliability of LMMs for deciphering OBI characters visually and textually. The recently proposed OracleAgent⁶³, an agent system designed for structured management and retrieval of OBI-related information, integrates multiple OBI analysis tools linked by large language models (LLMs), achieving efficient character, document, interpretive text, and image retrieval. These approaches are effective for familiar or clean OBIs but struggle with noise-contaminated, original, cross-modal OBI inputs. Some pictograph-centric deciphering frameworks limit their applicability, since the major part of the undeciphered OBI is ideographic. Meanwhile, due to the lack of a large-scale, comprehensive OBI corpus, these LMM architectures do not perform well on surface-level pattern matching and static knowledge retrieval, let alone dynamic hypothesis generation and multi-step logical reasoning⁶⁴.

In this view, this limitation may catalyze the development of chain-of-thought (CoT) reasoning-based OBI deciphering. Furthermore, from a learning method perspective, reinforcement learning-enhanced multimodal reasoning demonstrated by DeepSeek-R1⁵⁵ can become another research direction to foster a new line of multimodal reasoning work for OBI processing. Taken together, the OBI processing evolution from conventional expert-led laborious pipelines to emerging multimodal processors outlines a clear trajectory toward more trustworthy, adaptive, all-round expert-friendly AI systems. In the following sections, we provide a detailed analysis of each stage, from data to methodology and evaluation metric, as well as the emerging research directions that shape the future of OBI processing.

Data sources for OBI information processing

Data source forms the foundation for conducting OBI research. OBI processing from excavation to digital preservation typically involves the processes of data augmentation, recognition, rejoining, classification, and the ultimate decipherment, as suggested by OBI-Bench². Basically, the OBI datasets are used to train and test automatic models in the respective tasks, whereas emerging OBI datasets are generally designed for some specific OBI processing problems and applications. Since our objective is to conduct a comprehensive search to encompass the full landscape of the field, we do not explicitly set specific criteria for dataset filtering. In general, we mainly imposed constraints on the temporal scope so that only datasets published in the past ten years (i.e., 2015-2025), representing the era of data-driven OBI processing with clear task orientation, are included. Besides, we excluded those internal or private datasets with no technical documentation or accessibility for research comparison. Fig. 2 provides an overview of 36 datasets proposed for four main OBI processing tasks: recognition, rejoining, classification/retrieval, and deciphering. Information, including the publication year, appearance type, data source, the number of samples and classes, image resolution, and availability, as well as the file format, is summarized. Detailed introduction of these datasets and their data sources can be found below. Figures 4 and 5 show some visualization examples.

Online resources

There are many online OBI resources available for public learning and research, which also serve as the main data sources for constructing OBI datasets and promoting computational paleography. Here, we briefly introduce nine representative web platforms. Figure 6 visualizes their interfaces.

- **Yin Qi Wen Yuan** (<https://jgw.aynu.edu.cn/home/>) is a comprehensive and open-access knowledge-sharing platform for OBI studies, which is maintained by Anyang Normal University. It serves as a crucial data source by integrating three main components, i.e., a character database, a collection of oracle bone inscriptions, and a library of academic literature, housing an extensive collection that includes 154 types of OBI publications, 239,902 images of oracle bones, and 35,585 published research works. It offers advanced visualization and search tools, allowing for tracing their appearances across various inscriptions and examining their form and context of use.
- **Xiao Xue Tang** (<https://xiaoxue.iis.sinica.edu.tw/>) is a major scholarly resource chronicling the evolution of Chinese character forms and sounds, which is jointly developed and maintained by the Department of Chinese Literature at National Taiwan University and the Digital Culture Center at Academia Sinica. This platform allows researchers to query and compare the various forms of a character as it evolved, beginning from its earliest appearances in OBI, making it an invaluable tool and data repository for paleographic analysis.
- **Guo Xue Da Shi** (<https://www.guoxuedashi.com/>) is a large-scale, comprehensive online database dedicated to Chinese classical studies and traditional culture. Its key utility for an OBI review lies in its extensive dictionary and character database. The site hosts digitized versions of critical paleographic reference works, most notably an oracle bone script dictionary, a compilation of bronze inscriptions, Shuowen Jiezi, and the Kangxi dictionary.
- **Zhui Yu Lian Zhu** (<https://www.fdgwz.org.cn/ZhuiHeLab/Home>), established by the Zhui He Lab at Fudan University, is a specialized database that stores a growing collection of over 6,600 instances of oracle bone fragment splicing. It provides essential metadata, including the original publication and fragment numbers of the joined pieces, the source of the splicing information, and the name of the scholar credited with the rejoining.
- **Omniglot** (<https://www.omniglot.com/chinese/jiaguwen.htm>) is an online encyclopedia of writing systems and languages. Its page on “Oracle Bone Script” serves as a tertiary, introductory data source. It provides a concise summary of the script’s history, its function in divination, and a comparative chart of selected character examples with their modern equivalents and meanings. It also functions as an aggregator of external links to more specialized academic websites.
- **Multi-function Chinese Character Database** (<https://humanum.arts.cuhk.edu.hk/Lexis/lexi-mf/>), hosted by the Humanum Computing Research Center at The Chinese University of Hong Kong, is a comprehensive lexicographical and etymological resource. It provides an extensive collection of ancient character forms, including a dedicated oracle bone component table and numerous OBI glyphs. The database links these ancient forms to their modern counterparts, offering detailed etymological analyses, component breakdowns, and phonetic information, making it an invaluable tool for tracing the morphological evolution of characters from the Shang Dynasty.
- **Yin Xu OBI Database** (<https://obid.ancientbooks.cn/>), contributed mainly by Prof. Nianfu Chen, contains 59,591 oracle bones and 143,856 divination inscriptions. Each item consists of a description section and an interpretation section, while recording the classification of the oracle bone inscriptions, the specific source and citation, as well as the group of the diviner.
- **Chinese Etymology** (<https://hanziyuan.net>), created by Richard Sears (also known as “Uncle Hanzi”), is a comprehensive digital etymological database, compiling a massive collection of ancient Chinese character forms, including over 31,000 oracle bone characters, over 24,000 bronze characters, and over 49,000 seal characters.
- **OBI AI Collaborative Platform** (<https://www.jgwlbq.org.cn/home>) is an advanced database and online research platform developed in collaboration with partners such as Tencent and academic institutions. It collects approximately 1,430,000 oracle bone single characters, 4,184 lemmata, and nearly 680,000 identified glyphs. It allows users to search for a character and retrieve its modern equivalent, meanings, and evolutionary path.

Print resources

Books and monographs on OBI have long served as the principal source through which the public studied these materials before database technologies matured. Below, we provide a brief overview of representative materials from five key perspectives, including record, hand copy, transcription, periodization, and lexicon.

Table 2 | An overview of existing oracle bone inscription datasets categorized by tasks

Task	Dataset	Year	Type	Source	# Samples	# Classes	Resolution	Available	F ²
Recognition	YinQiWenYuan _{detection}	2020	Rubbings	Book	9823	–	540 × 776	✓	jpg
	OracleBone-8000 ³³	2020	Rubbings	Book	128,770	–	–	✗	–
	ACCID ⁵²	2023	Handprinted	Book	15,085	2892	–	✗	–
	O2BR ²	2025	Original	Website	800	–	1529 × 1192	✓	jpg
Rejoining	OB-Rejoin ³⁴	2022	Rubbings	Book	998	249	–	✗	–
	COBD ⁵⁴	2023	Rubbings	–	960	–	–	✗	–
	OBI-rejoin ²	2025	Orig. & Rubb.	Website	483	200	265 × 313	✓	png
	OBF ³⁵	2025	Original	Website	5374	110	721 × 816	✓	jpg
Classification & Retrieval	Oracle-20k ⁴³	2016	Handprinted	Website	20,039	261	–	✗	–
	OBC306 ⁵⁹	2019	Rubbings	Book	309,551	306	382 × 478	✓	bmp
	Oracle-AYNU ⁴⁴	2019	Handprinted	–	39,062	2583	64 × 64	✗	–
	HWOBC ⁷⁰	2020	Handprinted	Website	83,245	3881	400 × 400	✓	png
	Oracle-50k ⁷¹	2020	Handprinted	Website	59,081	2668	50 × 50	✓	jpg
	OBI-IJDH	2020	Rubbings	–	655	29	64 × 96	✓	png
	Oracle-250 ⁷²	2020	Handprinted	Website	92,160	250	–	✗	–
	Radical-148 ⁷²	2020	Handprinted	Website	108,989	148	–	✗	–
	OBI125 ⁷³	2022	Rubbings	Book	4257	125	278 × 473	✓	jpg
	OBI-100 ³⁰	2022	Handprinted	Book	4748	100	–	✗	–
	Oracle-241 ⁷⁵	2022	Handpri. & Rubb.	Website	78,565	241	588 × 700	✓	dmp
	ORCD ⁷⁶	2022	Handpri. & Rubb.	–	6700	64	–	✗	–
	OCCD ⁷⁶	2022	Handprinted	Website	62,186	1,320	–	✗	–
	OracleRC ⁵¹	2023	Rubbings	Book	2005	202	–	✗	–
	Oracle-MNIST ⁷⁸	2024	Rubbings	Website	30,222	10	28 × 28	✓	idx
OBI component 20 ⁷⁹	2024	Handprinted	Website	10,257	20	413 × 401	✓	png	
Deciphering	OBI-ECC ⁸⁰	2022	Handprinted	Website	4860	972	105 × 105	✓	png
	EVOBC ⁸¹	2024	Handpri. & Rubb.	Website & Book	229,170	13,714	465 × 857	✓	png
	HUST-OBC ⁸²	2024	Handprinted	Website & Book	140,053	10,999	400 × 520	✓	png
	ACCP ⁸³	2024	Handpri. & Rubb.	Website & Book	346,344	88,901	103 × 129	✓	jpg & png
	OracleSem ⁶⁰	2024	Handprinted	Website & Book	–	1762	–	✗	–
	GEVOBC ⁸⁴	2025	Handprinted	Website & Book	3,780	765	105 × 105	✓	png
	PD-OBS ⁸²	2025	Handpri. & Rubb.	Website & Book	211,796	–	209 × 236	✓	jpg & png
	PictOBI-20k ⁵	2025	Handprinted	Website & Book	15,175	80	512 × 512	✓	jpg
Emerging	RCRN ⁸⁵	2022	Rubbings	Book	1606	362	520 × 668	✓	png
	OBIMD ⁸⁶	2024	Handpri. & Rubb.	Book	10,077	–	518 × 842	✓	jpg
	RMOBS ⁶¹	2025	Handprinted	Website	> 20,000	900	–	✗	–
	Oracle-P15k ³	2025	Handpri. & Rubb.	OBC306	14,542	239	128 × 128	✓	png

Since both the classification and retrieval tasks of OBI rely on the category labels of individual characters, we merge the datasets of these two tasks. We report the average resolution of the images in the ‘Resolution’ column. ‘F²’ denotes the file format. We differentiate recognition datasets from classification datasets according to whether bounding-box annotations are provided. The website references of YinQiWenYuan_{detection} and OBI-IJDH are (<https://jgw.aynu.edu.cn/home/down/detail/index.html?sysid=3>) and (<http://www.ihpc.se.ritsumei.ac.jp/OBIdatasetIJDH.zip>), respectively.

- **Record:** The most famous record of OBI from the last century was compiled by Prof. Moruo Guo and is known as the “*Collections of oracle-bone inscriptions*”¹⁴. It is a monumental 13-volume compendium that represents a milestone in OBI studies. It systematically organizes 41,956 items, including rubbings, photographs, and transcriptions drawn from the first 80 years of oracle bone discoveries.
- **Facsimile:** “*Corpus of Oracle-Bone Facsimiles*”¹⁸, edited by Prof. Tianshu Huang, is the largest published archive of inscribed

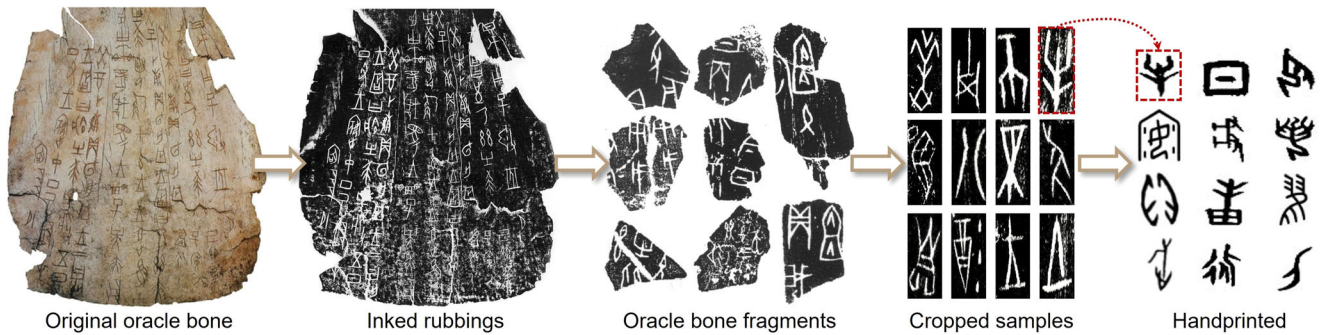


Fig. 4 | Visual manifestations of oracle bone inscriptions. The image displays five common appearances of OBI from excavation to synthesis, including original oracle bones, inked rubbings, oracle bone fragments, cropped samples, and handprinted OBI characters.

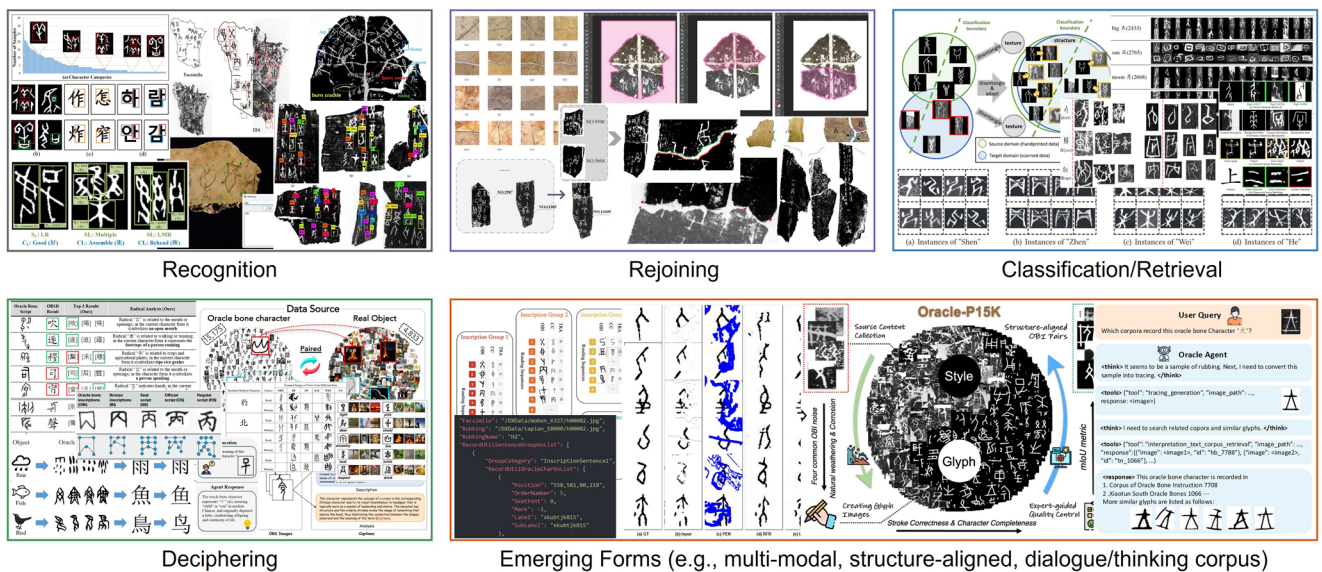


Fig. 5 | Examples of OBI data processed in different tasks. Some elements are adapted from the corresponding representative publications^{2,3,5,33–35,86,132}. Licensed under CC BY 4.0 (<http://creativecommons.org/licenses/by/4.0/>).

plastromancy remains, housing 70,659 fragmentary pieces that are all redrawn from the clearest extant rubrics or photographs. Each facsimile is tagged for class, optimal exemplar, previous joins, and duplicates, and is paired on the same opening with transcription and concordance indices, yielding a consultable reference for script evolution and divinatory syntax.

- **Transcription:** “*Collection of Explanations on Oracle-Bone Inscriptions*”¹⁶, contributed by Prof. Xingwu Yu, assembles over 4600 oracle-bone glyphs with all major scholarly glosses. It contains radical, glyph type summary table, and Pinyin indices, offering a systematic benchmark for subsequent studies of Shang script morphology.
- **Periodization:** “*Categorization and time division of the Oracle inscriptions for the king in Yin ruins*”⁶⁵, written by Prof. Tianshu Huang, applies rigorous typology to split the corpus into twenty font-based classes, embedding each within a two-lineage development theory. By integrating joins and calendrical algorithms, this book validates the relative sequence of *Zu*, *Bin Zu*, and *Li Lei* inscriptions, supplying a replicable methodology now exported to bronze chronology, geographical investigation, and Shang syntax studies.
- **Lexicon:** “*New Compendium of Oracle-Bone Characters*”⁶⁶, edited by Prof. Zhao Liu, digitizes thousands of high-resolution Shang glyph images by automated cropping and binarization, while Western-Zhou forms are hand-traced and then scanned. Integrating over 100 previous collections from Heji onward, the volume incorporates the latest decipherments and joins, offering the most comprehensive and

systematically arranged paleographic dataset available for quantitative and diachronic studies of early Chinese writing.

OBI Datasets for Recognition

In OBI recognition datasets, each image is densely annotated with axis-aligned or rotated boxes that localize every character region. By converting rubbings into precisely labeled imagery, they provide the training substrate for convolutional neural network (CNN)^{7,9}, Transformer⁶⁷, and SAM-based⁶⁸ detectors that can isolate characters amid complex fracture lines, background noise, and overlapping impressions. The resulting models not only accelerate the digital archiving for downstream tasks such as interpretation but also provide an objective, reproducible metric (e.g., mean average precision (mAP), intersection over union (IoU)) with which to benchmark new archivers against expert-drawn gold standards.

- **YinQiWen Yuan**_{detection} (<https://jgw.aynu.edu.cn/home/down/detail/index.html?sysid=3>): It is released on the Yin Qi Wen Yuan website, which contains 9823 rubbing images and 9134 annotation files that include the regional coordinates of individual characters in each image.
- **OracleBone-8000**³³: It contains 7824 OBI rubbing images from ref. 14 and 128,770 annotated character instances, with the characteristics of being highly imbalanced and sparse. A deep learning-based scene text detection algorithm is first employed to roughly predict the bounding box of every single character. Then the experts specify the sequence of each character at the sentence-level for fine-grained annotation.

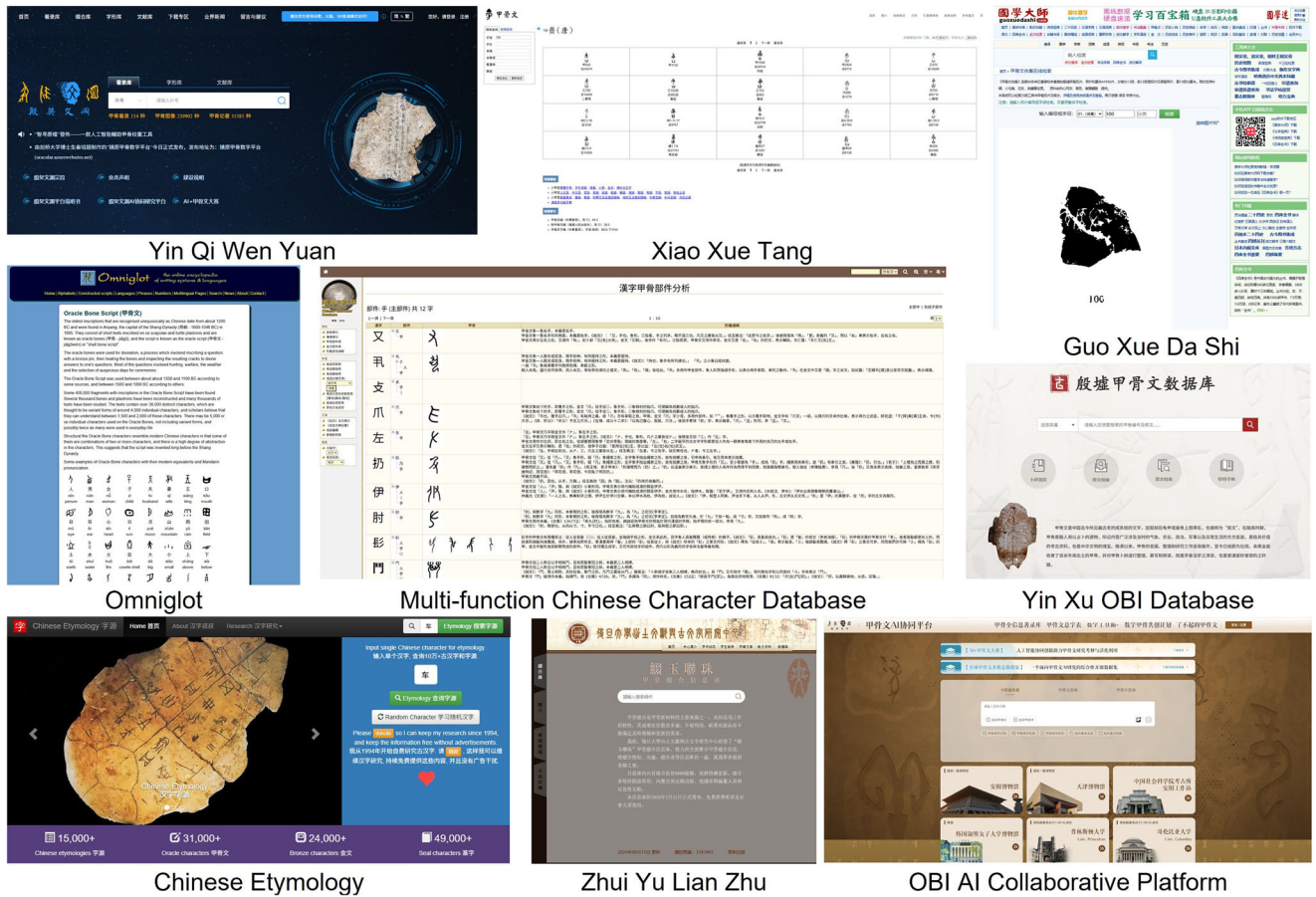


Fig. 6 | Screenshots of representative online OBI platforms. All nine platforms generally include functions such as glyph-based oracle bone character search and modern interpretation reference.

- ACCID⁵²: ACCID was constructed to satisfy both radical-level and character-level annotations to satisfy the requirements of the above-mentioned methods.
- O2BR²: O2BR consists of 800 original oracle bone images and 4211 bounding boxes. Its source contents were downloaded from the open museum led by the Institute of History and Philology in Academia Sinica under an appropriate license. Two domain experts and one senior OBI scholar were involved in the annotation and quality assessment processes.

OBI datasets for rejoining

Normally, an OBI rejoining dataset contains assembled high-resolution rubbings and expert-verified splice correspondences. By providing positive and negative edge labels between pieces, it enables machine-learning models to learn fracture contours, inscription continuity, and stratigraphic compatibility, dramatically reducing the search space for human specialists. Such a dataset also supplies calibrated benchmarks for evaluating automated rejoining algorithms or models, thereby accelerating the reconstruction of complete OBI and subsequent historical, linguistic, and paleographic analysis. Below, we list the existing OBI rejoining datasets.

- OB-Rejoin³⁴: Its rejoining results are manually retrieved from books such as “Collections of Oracle-Bone Inscriptions”, “Catalog of Oracle Bone Rejoinings”, and “The Fifth Collection of Oracle Bone Rejoinings”, as well as multiple digital repositories. It contains 998 rubbing images covering various writing styles with 249 pairs of known rejoinings. Specifically, only the top and bottom borders of the OB fragments are stroked in this work. Six domain experts and two senior OBI scholars

participated in the annotation process using a digital drawing tablet with commercial photo editing software.

- COBD⁵⁴: COBD contains 480 groups (960 images) of joinable oracle bone contour curve images and 1083 unvalidated OB rubbing images. Each group includes two OB rubbing images: the upper and lower OB fragment rubbing images.
- OBI-rejoin²: OBI-rejoin comprises 200 complete OBI pieces with 483 rejoinable fragments scraped from the website of the Pre-Qin Research Office at the Institute of History and the Chinese Academy of Social Sciences. Four OBI experts were involved in the annotation process. To reproduce the authentic conditions of oracle bone fractures, annotators manually tear the complete version on appropriate borders, utilizing their domain expertise to induce central or peripheral fractures. Its open-source nature provides a verifiable channel for the subsequent development of rejoining algorithms.
- OBFI³⁵: OBFI, with a total of 5,374 oracle bone images, consists of four subdatasets, i.e., ZLC, BZ, LS, and DJ, covering yellowing, high contrast, low contrast, and grayscale scenarios, respectively. It also includes more than 138,855 target area images, with 115,893 unrejoinable and 22,962 rejoinable target area images.

OBI datasets for classification and retrieval

Different from recognition datasets, OBI classification datasets are centered on holistic, image-level labels rather than localized annotations. Currently, the dominant classification and retrieval standards are formulated based on the semantic meaning of the oracle bone characters. Below, we introduce some representative datasets. Figure 7 shows the class distribution with a typical long-tail characteristic that exists in current OBI datasets.

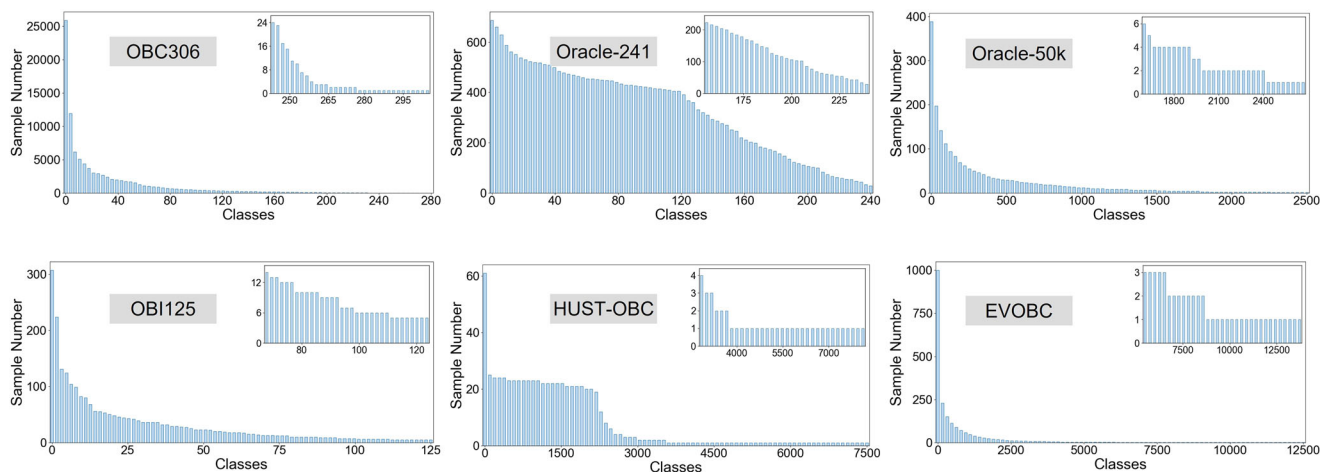


Fig. 7 | The long-tail distribution phenomenon in OBI datasets. First row from left to right: OBC306, Oracle-241, and Oracle-50k. Second row from left to right: OBI125, HUST-OBC, and EVOBC.

- **Oracle-20k**⁴³: The first large-scale OBI dataset, Oracle-20k, is composed of 20,039 oracle character instances belonging to 261 categories collected from the website *Chinese Etymology*. Its characters are a handprinted version extracted and modified by later generations from original OBI rubbings.
- **OBC306**⁶⁹: OBC306, collected from eight authoritative oracle bone publications, is the largest existing OBI dataset, which contains 309,551 samples classified into a total of 306 classes. However, the number of samples in each class varies significantly, ranging from 1 to 25,898, leading to a long-tailed distribution.
- **Oracle-AYNU**⁴⁴: Oracle-AYNU consists of 2583 categories covering 39,062 instances with 2 to 287 instances per category. All of them are binary images and are fixed to a size of 64×64 .
- **HWOBC**⁷⁰: HWOBC is a handprinted oracle bone character dataset, containing 83,245 single-character samples belonging to 3,881 categories. Each of them is fixed to 400×400 and follows a special font named AYJGW, which possesses a high intra-class similarity.
- **Oracle-50k**⁷¹: Oracle-50k, sourced from three digital OBI websites, consists of 59,081 instances divided into 2668 classes, where a long-tail distribution exists, making its classification problem naturally a few-shot learning problem.
- **OBI-IJDH** (<http://www.ihpc.se.ritsumei.ac.jp/OBIdatasetIJDH.zip>): It contains 655 OBI character images divided by 29 categories with mixed Japanese and English labels.
- **Oracle-250**⁷²: Oracle-250 consists of 250 most frequently used characters in OBI. Due to the uneven distribution of various categories, the authors adopt data augmentation algorithms to balance the sample size of the original data. At last, the total number of samples is 92,160.
- **Radical-148**⁷²: Radical-148 is constructed based on 148 main component types in OBI. The author invited 700 volunteers to handwrite some OBI components to balance the sample size, resulting in 108,989 samples.
- **OBI125**⁷³: OBI125 includes 4257 OBIs that belong to 125 classes, obtained by scanning and archiving the OBIs in the collection of the Shanghai Museum (Volume I)⁷⁴.
- **OBI-100**⁸⁰: OBI-100 contains 100 character classes, covering various types of characters, such as animals, plants, humanity, society, etc., with a total of 4748 samples.
- **Oracle-241**⁷⁵: Oracle-241, sampled from the Yin Qi Wen Yuan website, contains 78,565 handprinted and scanned rubbing characters of 241 categories.
- **ORCD**⁷⁶: ORCD is an oracle radical character dataset, containing 1288 rubbing samples and 5412 handprinted oracle radicals in 64 categories.
- **OCCD**⁷⁶: OCCD is collected from image synthesis technology and online handwritten OBI collection. The total number of samples is 62,186 images and 1320 categories. It contains 54,876 characters composed of single radicals and 7310 characters copied from existing oracle-combined characters in the font library.
- **OracleRC**⁵¹: OracleRC, sourced from ref. 77, includes 2005 OBI characters that can be categorized into 202 radicals and 14 structural relations. All radicals and structural relations were annotated by eight linguists.
- **Oracle-MNIST**⁷⁸: Oracle-MNIST dataset, originated from the Yin Qi Wen Yuan website, contains 30,222 28×28 grayscale oracle bone character images from 10 classes. It is relatively class-balanced, ranging from 3399 instances to 2,328 instances per class. Its MNIST-aligned format makes it easy to conduct benchmarking on existing classifiers and algorithms.
- **OBI component 20**⁷⁹: This dataset was designed for a component-level retrieval task, namely, using an OBI component to retrieve all characters containing this component. It comprises 9245 OBI characters and 1012 component images, covering 20 distinct components.

OBI datasets for deciphering

OBI decipherment datasets normally pair cropped OBI glyph images with their modern Chinese equivalents or textual glosses, providing the supervised signal needed to bridge the historical orthographic gap. By aligning undeciphered characters with known characters or semantic annotations, they enable sequence-to-sequence or contrastive-learning models to discover semantic and structural regularities, yielding candidate readings for previously opaque OBI characters.

- **OBI-ECC**⁸⁰: OBI-ECC, sourced from the Guo Xue Da Shi website, consists of images of the evolution of existing Chinese characters, including OBIs, bronze inscriptions, seal script, official script, and regular script. It has 972 categories, with each category containing 5 images.
- **EVOBC**⁸¹: EVOBC contains six evolutionary stages of the Chinese character, including oracle bone characters, bronze inscriptions, seal script, spring and autumn characters, warring states characters, and clerical script. Its oracle bone character part consists of 75,681 images representing 3077 distinct character categories.
- **HUST-OBC**⁸²: HUST-OBC encompasses a total of 77,064 images of 1588 individual deciphered scripts and 62,989 images of 9411 undeciphered characters. Its alignment with modern Chinese

characters facilitates the qualitative evaluation of interpretation models.

- **ACCP**⁸³: ACCP is expanded upon EVOBC and HUST-OBC to encompass characters from another two periods. Concretely, 9354 characters (1662 AD–1722 AD) from the Kangxi dictionary and 88,899 regular characters are added.
- **OracleSem**⁶⁰: OracleSem is a semantically enriched OBI dataset based on HUST-OBS and EVOBC, which incorporates rich semantic annotations that capture multiple layers of linguistic information. It provides detailed annotations for each character about its pictographic composition, structural organization, and semantic evolution pathway. The OracleSem dataset comprises 1762 OBI characters, with each character associated with 10 to 20 images.
- **GEVOBC**⁸⁴: GEVOBC is a graph-based evolutionary OBI dataset transformed from ref. 80. It contains 3780 Chinese character images classed into 756 groups, where each group contains five evolutionary shapes of an image. Each image is first processed with key point extraction, followed by node connection based on its strokes.
- **PD-OBS**⁶²: PD-OBS is a pictographic deciphering oracle bone script (OBS) dataset with multiple sources, comprising 47,157 Chinese characters annotated with OBS images and detailed radical-pictographic analysis texts. Among these, 3,173 and 10,968 characters are associated with OBS image and clerical script images, respectively.
- **PictOBI-20k**⁵: PictOBI-20k was designed to evaluate large multimodal models (LLMs) on the visual decipherment tasks of pictographic OBI characters, i.e., correlating OBI characters with real scenes. It comprises 15,175 different OBI character images from 80 pictographs and 4833 corresponding object images with various question-answer pairs constructed for evaluating LLMs.

Emerging OBI datasets

Beyond the aforementioned traditional OBI processing tasks, numerous studies have also proposed relevant datasets targeting specific issues in OBI research. For example, since most OBI images are in the rubbing form that suffers from severe image degradation caused by limited shooting conditions, constructing a high-quality oracle character dataset becomes rather challenging. Shi et al.⁸⁵ proposed RCRN, an OBI character image dataset containing 1467 and 139 noisy-clean image pairs, for OBI character image restoration. Later, Li et al.³ introduced Oracle-P15k, a structure-aligned OBI dataset with 14,542 images across 239 classes that covers four common noises in OBI rubbings for OBI character generation and denoising, significantly increasing the scale of OBI datasets for training generative models. Another development direction for OBI datasets is multimodality. Most existing OBI datasets only focus on or are annotated in one or a few dimensions, limiting the potential utility of their application. To this end, Li et al.⁸⁶ presented an oracle bone inscriptions multi-modal dataset (OBIMD), which includes annotation information for 10,077 pieces of oracle bones. Specifically, it contains detection boxes (*coordinates*), character categories, transcriptions, corresponding inscription groups, and reading sequences in the groups of each oracle bone character, providing comprehensive annotations. Similarly, RMOBS⁶¹ comprises over 20k entries across 900 deciphered OBS characters with pictographic components. Each entry consists of an oracle glyph, a concept, the key components, and a bounding box layout, facilitating fine-grained OBI decipherment model training. Overall, we anticipate that additional OBI datasets tailored to specific tasks will be released in the future, and we especially hope to see large-scale, open-source, and sentence-level OBI corpora that will facilitate the training of dedicated LLMs for the field of OBI and ancient text.

Observations

We further summarize the core characteristics for the aforementioned 36 OBI datasets based on labels, utilities, strengths, and limitations in Table 3. Please note that this does not imply any superiority or inferiority between these datasets. Instead, it provides researchers with some suggestions from

the perspectives of ease of use, verifiability, and reproducibility. Here, we make the following observations.

First, the current OBI dataset ecosystem is characterized by a trade-off between scale and granularity. Large-scale datasets such as OBC306⁶⁹ and HWOBC⁷⁰ provide sufficient data volume for deep learning but often lack the fine-grained structural or semantic annotations required for higher-order reasoning. Conversely, datasets like OBIMD⁸⁶ and OracleSem⁶⁰ offer radical-level and component insights but at a much smaller scale, often limited to a few hundred categories. Second, the long-tail distribution remains the most persistent limitation across the field, reflecting the historical reality of the divination corpus. Third, the lack of original bone images (as opposed to rubbings) in datasets like HUST-OBC⁸² and EVOBC⁸¹ limits the deployment of models in actual archeological excavations where rubbings have not yet been produced. Overall, the efficacy of OBI processing is fundamentally constrained by the data substrate quality. Researchers must navigate three primary visual challenges: authentic noise, resolution constraints, and distribution imbalance.

OBI processing tasks and approaches

In this section, we merely review those AI technique-based OBI processing approaches. From a methodological perspective, the early expert-led empirical methods were relatively uniform, i.e., literature consultation or genealogical comparison, as mentioned above.

OBI preprocessing: data augmentation and synthesis

The deployment of data-driven deep learning models in OBI analysis is fundamentally constrained by the quality and quantity of available training data. Unlike general computer vision tasks supported by massive, clean datasets, OBI research faces unique challenges that render data augmentation not merely beneficial, but strictly necessary. These challenges stem from three primary dimensions: physical degradation, image fidelity, and statistical distribution^{2,6,13,87}.

First, after enduring millennia of burial and exposure, OBI artifacts have suffered extensive natural erosion and anthropogenic damage. The majority of excavated materials exist as fragmented remnants rather than complete bones, resulting in severe semantic discontinuity and information loss, where crucial glyphs are often partially missing or obliterated. Second, in the early stages, the primary method for storing and disseminating OBI was the rubbing, which often suffers from low image fidelity. These images are frequently characterized by complex background noise and texture artifacts^{4,69}, as shown in Fig. 8. A critical challenge lies in the visual ambiguity between genuine character strokes and accidental surface cracks or chipping. Without robust augmentation to simulate these noise patterns, models struggle to disentangle semantic features from background interference. Third, current OBI datasets generally suffer from the long-tailed distribution characteristic, exhibiting extreme class imbalance. Direct training on such skewed data causes overfitting to head classes while failing to generalize to tail classes. To address these problems, researchers have resorted to refining and expanding the dataset through various methods such as geometric transformations^{88–90}, noise injection^{6,91}, U-Net architecture^{92,93}, and generative synthesis (e.g., generative adversarial networks (GAN)^{73,94,95} and diffusion models^{3,96,97}).

Traditional data augmentation methods employ rotation, scaling, flipping, and affine transformations to simulate variations in the spatial arrangement and writing styles of inscriptions. In^{88,90}, the authors further applied cutting, brightness changing, and contrast adjustment to increase the diversity of data. Besides, Gaussian noise addition is also used to simulate the distortion in the OBI rubbing images. However, such operations fail to introduce substantive semantic variations or novel stroke patterns, and heuristic noise injection is often too simplistic to model the complex, non-stationary degradation in real rubbings, potentially leading the model to learn irrelevant artifacts. Orc-Bert⁷¹ leveraged a self-supervised BERT model pre-trained on large unlabeled Chinese character datasets to generate sample-wise augmented samples. Afterwards, concatenated with Gaussian noise, the model further performed point-wise displacement to improve

Table 3 | Comparison between 36 OBI datasets for labels, utilities, strengths, and limitations

Task	Dataset	Annotation/ Label	Utility	Strengths	Limitations
Recognition	YinQiWenYuan _{dedection}	Bounding box; Binary OBI label	Multiple-character recognition on rubbings	Manually annotated; Widely-used	Noise-contaminated
	OracleBone-8000 ³³	Bounding box	Multiple-character recognition on rubbings	Manually annotated	Noise-contaminated; Imbalanced; Close-source
Rejoining	ACCID ³²	Bounding box; Class Label; Structural relation	Radical and single character recognition	Fine-grained radical-level annotation	Focus on single-character; Close-source
	O2BR ²	Bounding box	Multiple-character recognition on real OBI images	Focus on original OBI; Open-source	Relatively small scale
	OB-Rejoin ³⁴	Fragment-level outlines	Training data-driven rejoining model	Expert-level annotations	Low image quality; Sparse; Close-source
	COBD ³⁴	Fragment-level outlines	Training and testing rejoining models	Include curve trajectory sequence	Relatively small scale; Close-source
Classification & Retrieval	OBI-rejoin ²	Adjacency label	Testing rejoining models	Include real OBI and rubbings; Open-source	Relatively small scale; Coarse-grained label
	OBF ³⁵	Binary rejoinable label	Training and testing rejoining models	Large-scale; Multi-source; Open-source	Coarse-grained label
	Oracle-20K ⁴³	Class label	General character-level classification and retrieval	Early representative work	Severe long-tail distribution; Only handprinted
	OBC306 ³⁹	Class label	General character-level classification and retrieval Variant analysis	Largest sample size for classification; realistic noise	Severe long-tail distribution; Low image quality; Labeling errors
	Oracle-AYNU ⁴⁴	Class label	General character-level classification and retrieval	Larger than Oracle-20k	Severe long-tail distribution; Low resolution; Close-source; Only handprinted OBI
	HWOB ⁷⁰	Class label	General character-level classification and retrieval	Large-scale; Widely-used; Intra-class variants	Only handprinted OBI
	Oracle-50K ⁷¹	Class label	General character-level classification and retrieval	Large-scale; Widely-used	Severe long-tail distribution; Only handprinted; Low resolution
	OBI-IJDH	Class label	General character-level classification and retrieval	Open-source	Small scale
	Oracle-250 ⁷²	Class label	General character-level classification and retrieval Variant analysis	Diverse handprinted variants	Focus only on the top-250 most frequent characters; Close-source
	Radical-148 ⁷²	Radical label	General radical-level Classification and retrieval; Variant analysis	Diverse handprinted variants	Close-source
Oracle-MNIST ⁷⁸	OBI125 ⁷³	Class label	General character-level classification and retrieval; Variant analysis	Focus on rubbings	Severe long-tail distribution; Relatively small scale
	OBI-100 ⁷⁰	Class label	General character-level classification and retrieval; Variant analysis	Open-source	Limited coverage; Only handprinted; Small scale
	Oracle-241 ⁷⁵	Class label	Character-level classification; OBI generation; Style transfer	Multi-type; Open-source	Non-pixel-level alignment
	ORCD ⁷⁶	Radical label; Bounding box	OBI radical recognition and classification	Diverse handprinted radical	Relative small scale; Close-source
	OCCD ⁷⁶	Combined character structure	OBI radical detection	First Oracle combined character dataset	Close-source
	OracleRC ⁵¹	Radical- and stroke- level label	OBI classification	Fine-grained character decomposition	Long-tail distribution; Close-source
Oracle-MNIST ⁷⁸	Oracle-MNIST ⁷⁸	Class label	General character-level classification and retrieval Variant analysis	Open-source; Even distribution	Low resolution; Noise-contaminated
	OBI component 20 ⁷⁹	Component-level label	OBI retrieval	First work for component-level annotation	Relatively small scale; Long-tail distribution

Table 3 (continued) | Comparison between 36 OBI datasets for labels, utilities, strengths, and limitations

Task	Dataset	Annotation/ Label	Utility	Strengths	Limitations
Deciphering	OBI-ECC ⁶⁰	OBI character evolution process	Modern-Chinese-based deciphering	Open-source	Relatively small scale; Only handprinted
	EVOBC ⁶¹	OBI character evolution process	Modern-Chinese-based deciphering; OBI Generation	Multi-source; Open-sourcescale; Widely-used	No explicit limitations
	HUST-OBC ⁶²	Aligned modern Chinese	Modern-Chinese-based deciphering	Multi-source; Open-sourcescale; Widely-used	Long-tail distribution; Only handprinted
	ACCP ⁶³	OBI character evolution process; Class label; Radical sequence	Modern-Chinese-based deciphering	Fine-grained OBI decomposition Multi-source; Open-sourcescale	Expand upon HUST-OBC and EVOBC
	OracleSem ⁶⁰	Pictographic and semantic caption	Pictographic interpretations; Component analysis; Modern Chinese mapping	Fine-grained text interpretation	Close-source; Only handprinted
	GEVOBC ⁶⁴	Graph-based evolution process; Class label	Character evolutionary stage alignment	Novel graph data structure	Relatively small scale; Only handprinted
	PD-OBS ⁶²	Modern Chinese; Ancient form; Radical annotation	Pictographic interpretations; Radical analysis	Fine-grained OBI decomposition with interpretations	Only handprinted
	PictOBI+20k ⁵	Class label; OBC-Object pair; Reference points; Attention map	Visual decipherment; Attention consistency; LMM evaluation	LMM-oriented; Perception mechanism	Pictographic OBI only
Emerging	RCRN ⁶⁵	Synthetic noise variants	OBI denoising	Early representative work; Original-distorted image pair	Relatively small scale; Lack of real-world noise
	OBIMD ⁶⁶	Bounding box; Class label; Transcriptions; Inscription groups; Reading sequences	Multiple-character recognition; Rubbing denoising; Character matching; Completion; Reading sequence prediction	Multimodal; Good applicability; Open-source	Relatively small scale
	RMOBS ⁶¹	Radical-level bounding box; Semantic concept; Key component	Semantics understanding; Visual grounding	Fine-grained character annotation	Close-source; Only handprinted OBI
	Oracle-P15k ³	Class label; Image-glyph pair	OBI generation; Long-tail distribution; Style transfer	Structure-aligned pairs for generative AI and denoising	Relatively small scale compared to recognition sets

³The website references of YinQiWenYuan³detection and OBI-JDH are (<https://jgw.aynu.edu.cn/home/download/detail/index.html?sysid=3>) and (<http://www.jhpc.se.ritsumei.ac.jp/OBIdata/JDH.zip>), respectively.

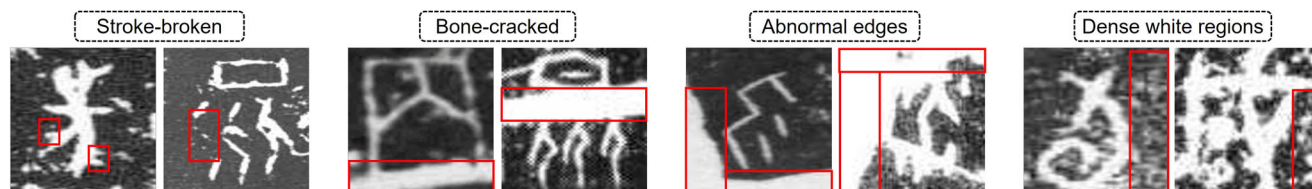


Fig. 8 | Common noise categories identified in OBI rubbing images. This figure illustrates four typical noise types: stroke-broken, bone-cracked, abnormal edges, and dense white regions.

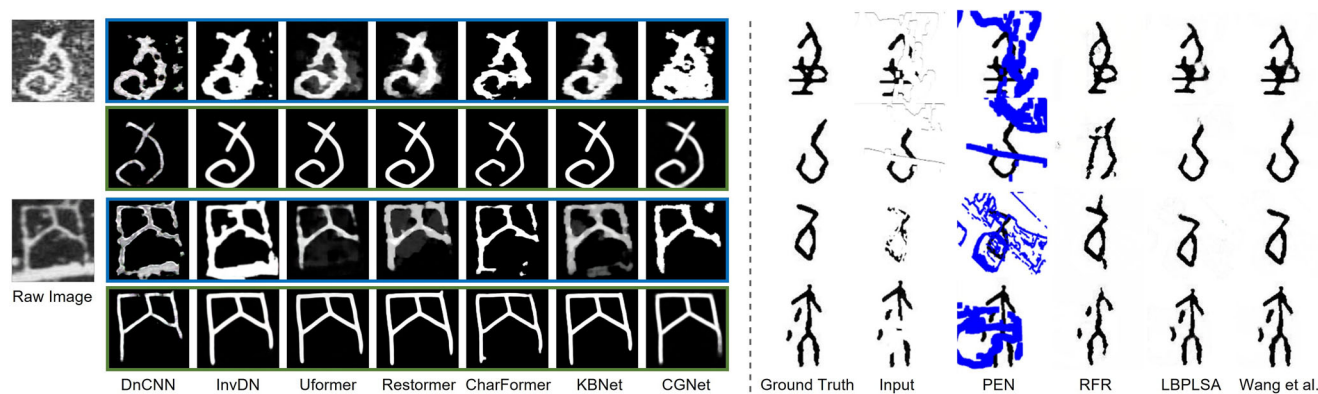


Fig. 9 | Left. Performance visualization of different OBI denoising methods trained on the RCRN (blue)⁸⁵ and Oracle-P15K (green) datasets, adapted from ref. 3. **Right:** Comparison of different OBI inpainting models on the irregular mask, adapted from ref. 99. Licensed under CC BY 4.0 (<http://creativecommons.org/licenses/by/4.0/>).

diversity. Subsequently, the emergence of generative models marked a paradigm shift in data augmentation. These models enabled the synthesis of high-quality, realistic samples with controllable attributes, going beyond simple geometric variations. Yue et al.⁷³ proposed a dynamic data augmentation strategy that utilizes the Wasserstein GAN-GP model⁹⁸ to learn and generate a new OBI dataset, followed by a dynamic selector, to keep the dataset content balanced. In⁹⁴, a CycleGAN-based OBI data generation strategy was proposed, which learns the mapping between the glyph image data domain and the real sample data domain. Similarly, Gao et al.⁹⁵ proposed a two-stage decomposition GAN to augment OBI samples exclusively from existing data, which learns an unidirectional mapping to transform low-quality samples into high-quality ones while ensuring enhanced realism and diversity. More recently, Li et al.³ proposed a controllable pseudo OBI image generator by incorporating glyphs and styles through a stable diffusion-based image translation architecture for expanding the tail data. Specifically, users can take those easily accessible rubbing images as a noise style and transfer the handprinted version of scarce OBI characters to realistic noisy versions.

In addition to expanding the number of training samples, researchers also propose to restore the OBI rubbing images through denoising and inpainting techniques to enhance their recognizability. Figure 9 visualizes some OBI denoising and inpainting results. Yang et al.⁹² formulated a large kernel convolutional attention-based U-Net framework to restore OBI images, which consists of two modified U-Net networks that perform the edge inpainting and overall image inpainting function, respectively. In ref. 99, Wang et al. introduced a two-stage GAN-based model that incorporates a dual discriminator structure to capture both global and local image information. A spatial attention mechanism and a multi-level fusion loss function are proposed to keep the accuracy of restoration. In ref. 93, the authors designed a U-Net-based cross-modal data homogenization module to unify heterogeneous data representations, which transfers rubbings into handwritings to bypass the problem of rubbing recognition. Diao et al.¹⁰⁰ presented a glyph extraction-driven image generation network for high-precision OBI restoration, leveraging character glyphs as supplementary information to address complex degradations while preserving the original glyph structures. Similarly, Zhang et al.⁹⁰ developed a structure-aware

diffusion model for OBI restoration, which contains an adaptive dynamic adjustment mechanism to perceive the hierarchical structure of the reconstructed image. Li et al.⁴ proposed OBIFormer, a fast attentive denoising framework for OBI images, which employs channel-wise self-attention, glyph extraction, and a selective kernel feature fusion module, achieving superior denoising results on multiple OBI datasets while being computationally efficient. Moreover, Orpaint⁹⁷ leverages the reverse generative capabilities of diffusion models and integrates the visual state space (VSS) block¹⁰¹, achieving significant reductions in time and computational costs compared to those U-Net architectures with multi-head self-attention, serving as an efficient and cost-effective solution for OBI inpainting.

OBI recognition

Generally, after being unearthed, an OBI fragment will undergo recognition to extract independent characters, thereby proceeding to downstream tasks. Here, we follow the standard object recognition definition^{102,103} that an OBI recognition task should include both oracle character localization and identification (Fig. 10). This also distinguishes this subsection from the OBI classification. Specifically, we introduce both traditional pattern recognition-based methods and modern deep learning-based methods.

Traditional pattern recognition. Early pattern recognition approaches are anchored in handcrafted feature extraction and classical classifiers, including structural descriptors, stroke-level graphs, directional histograms, and elastic matching. For example, Zhou et al.¹⁰⁴ regarded the OBI character as a non-directional graph and extracted its topological properties to make a two-stage recognition. Meng et al.⁸⁷ proposed a four-directional scan labeling to reduce the noise in the OBI rubbings, and developed a template matching-based OBI recognition method. In refs. 105,106, they further utilized line features and compared similarities in Hough space. Qu et al.¹⁰⁷ developed a recognition method by analyzing the topological features of OBI, such as topological feature points, connected domains, and genus. However, their reliance on laboriously tuned parameters and brittle stroke segmentation routines has limited scalability to the glyph variability exhibited by tens of thousands of unearthed fragments.

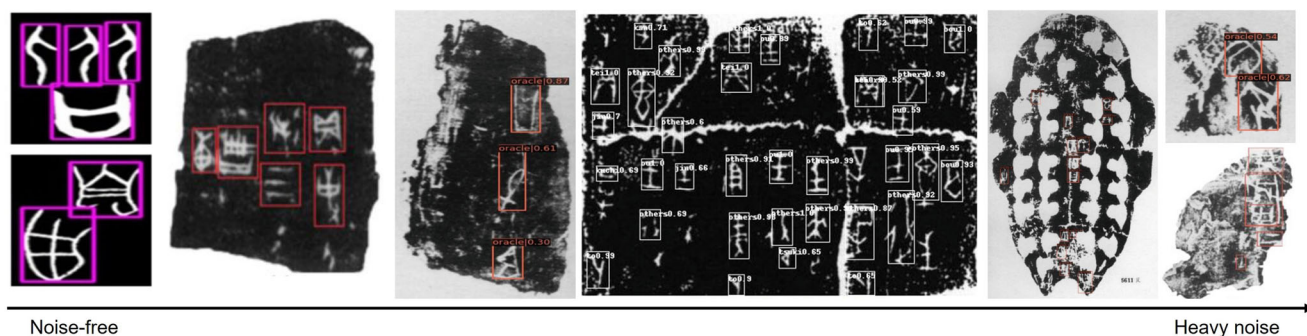


Fig. 10 | Results of OBI recognition across varying levels of noise severity. These data are adapted from refs. 66,90,138.

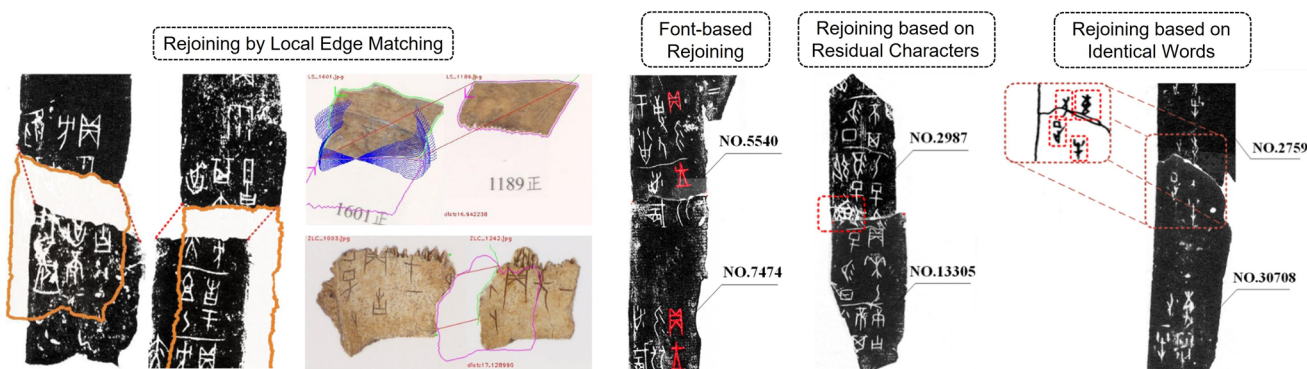


Fig. 11 | Visualization of representative OBI rejoining schemes, adapted from refs. 120,121. Licensed under CC BY 4.0 (<http://creativecommons.org/licenses/by/4.0/>).

Deep representation learning-based recognition. Driven by the rapid advances in deep learning-based generic object recognition, studies in the OBI domain have also shifted toward finer-grained goals that simultaneously localize and categorize every character instance in an image, generating bounding boxes together with confidence scores that quantify the likelihood of each character's presence^{6,13,108}. Over the past ten years, a large number of automatic OBI recognition approaches combining models such as Fast R-CNN¹⁰⁹, single-shot detector (SSD)⁸, YOLO⁹, and their variants, have been proposed.

In ref. 110, the authors proposed to detect OBI characters using Faster R-CNN⁷, while using multi-scale feature fusion to improve recognition capabilities. Meng et al.¹¹¹ remoulded SSD, enabling it to fit images with higher resolution and to adapt to the recognition of small characters. Liu et al.¹¹² redesigned the sizes and ratios of the anchor box according to the data characteristics by using K-means clustering. To stabilize features and suppress noise, a spatial pyramid block was proposed, offering more robust recognition results. Fujikawa et al.³¹ implemented YOLO-v3 tiny to perform a rough recognition first and used MobileNet to further inspect the unrecognized areas. Following the multiple optimized versions of YOLO-v3, Wang et al.¹¹³, Wen et al.³², Zhen et al.¹¹⁴, and Meng et al.¹¹⁵ further applied the YOLO-v4, YOLO-v5, and YOLO-v8 to OBI recognition tasks, respectively, significantly enhancing the recognition speed and accuracy. More recently, Xiong et al.¹¹⁶ proposed FDW-YOLO, an improved YOLO-v12 for OBI detection, which has a feature focusing diffusion pyramid network design to enhance the integration of multi-scale features. Together, a dynamic mixed convolution block is designed to reconstruct the network structure, helping the model adaptively extract and fuse effective information under complex backgrounds. The resulting FDW-YOLO surpasses vanilla YOLO-v12 in both *F1*-score and *mAP50*. Beyond these, Li et al.¹¹⁷ utilized oracle character prototypes and contrastive loss as supervisory signals, cyclically decoupling Oracle features and isolating information-rich character structure features, which effectively mitigates the influence of noise on OBI recognition. Different from mainstream approaches, Fu et al.¹¹⁸ performed OBI detection with the employment of information from

multilabel annotations rather than single-location information. A pseudo-label predictor is plugged after the backbone network for learning the particular structure prior to each inscription, thereby improving the recognition precision. In ref. 119, Tao et al. leveraged the OBI font library dataset as prior knowledge to enhance feature extraction in the detection network through clustering-based representation learning.

Furthermore, researchers have also studied the radical-based recognition task since radicals are important components for composing characters and expressing semantics. In ref. 76, Lin et al. combined the maximally stable extremal regions algorithm to generate single radical data annotation on OBIs and then integrated multi-scale features to implicitly extract radical features for radical detection. Diao et al.⁵² pre-defined 14 categories of character structures, such as single-body, surrounding, left-right, and up-down structures, and adopted a divide-and-conquer recognition strategy by decomposing characters into radicals for benchmarking zero-shot OBI character recognition methods.

Despite these advances, challenges still remain. In the full pipeline of OBI processing, recognition, as an intermediate step, is now governed more by data distribution, pre-processing fidelity, and morphological diversity than by model architecture. Currently, incremental recognition network refinement may no longer deliver the interpretability that OBI experts demand in practice.

OBI rejoining

Unlike other types of OBI processing tasks that have achieved unprecedented results with the advancement of AI technologies, rejoining studies have been painstakingly laborious up until now. Figure 11 shows four common rejoining bases. Existing OBI rejoining methods can be divided into two categories: contour-matching-based methods and deep-learning-based methods.

Contour matching-based methods. Early rejoining methods typically employ geometric heuristics and manually extract contour features for matching. In refs. 22,26, Wang et al. proposed an algorithm for

paragraph-by-paragraph search and matching of contour fragments for fracture phenomena, as well as a matching technique based on shape function operations. Specifically, contour fragments are described by the Freeman chain code, then matched through cascaded Fourier-moment-shape descriptors to propose joins. After the feature calculation, the oracle bone fragments to be rejoined and the candidate ones can be denoted by eigenvectors F_r and F_c . Then, a function is employed to calculate their similarity:

$$\text{Sim}(F_r, F_c) = \sum_{i=1}^{F_N} \frac{\min(F_r(i), F_c(i))}{\max(F_r(i), F_c(i))} \quad (1)$$

where F_N denotes the length of the eigenvector. Based on this similarity score, an OBI rejoining system can determine the connectivity between two fragments. Later, Tian et al.¹²⁰ further improved the curve matching process and introduced a partial-to-global curve matching algorithm. An OBI fragment is first labeled red by experts to distinguish the contour curves. Then, position correction is applied to unify dimensions and orientation. At the feature extraction stage, the inclination angle and horizontal distance are recorded as the basic local features of the curves. To obtain the global features descriptor, the authors further extract the global discrete points by collecting coordinate points, then denoising them with Gaussian smoothing. At last, a Pearson correlation analysis is adopted to compare the closeness of the two sets of characteristic vectors, where a curve fitting degree analysis algorithm is included to narrow down the number of candidate curves. Moreover, in ref. 33, Zhang et al. formulated the OBI rejoining task as a time series comparison task and transformed the stroked borderline curves into numerical *time series*. A simple tolerance difference similarity measure was devised for differentiating two time series.

Deep learning-assisted methods. Most of the deep learning-based OBI rejoining approaches share the same procedures as traditional methods in contour representation that utilize similar edge coordinate matching algorithms, while differing in the subsequent discrimination process. For example, in ref. 36, the authors proposed a deep rejoining scheme for automatic OBI rejoining, where an edge equal-distance rejoining method was first used to locate the matching position of the edges of two fragment images, and then a CNN was used to evaluate the similarity of texture in the target area image. Similarly, Zhang et al.¹²¹ proposed an internal similarity network to automatically rejoin the OBI fragment image. At the beginning, an edge equidistant matching algorithm was given to search for similar coordinate sequences of edge segment pairs. Then, an internal similarity pooling layer was employed to compute the internal similarity of the convolution feature gradient map. Recently, Zhang et al.³⁵ further enhanced the effectiveness of edge matching by proposing a longest similar edge segment algorithm and a complete image rejoining method. Yuan et al.⁵⁴ presented SFF-Siam, which includes a similarity feature fusion (SFF) module and a backbone feature extraction network to combine inputs and evaluate the similarity, respectively. Different from the aforementioned approaches, S³-Net³⁴ combines GAN and the Siamese network for data augmentation and measuring the similarity of contour curves to rejoin the OBI fragments.

Furthermore, we include a very related OBI duplicates discovery task, which can be regarded as a matching task involving fragments of the same structure, represented by Zhang et al.'s work¹²². They designed a progressive OBI duplicate discovery framework that combines low-level keypoint matching with text-centric content-based matching to refine and rank the candidate OBI duplicates with semantic awareness. Specifically, two unsupervised methods, Superpoint¹²³ and Lightglue¹²⁴, were used for keypoint extraction and mapping, respectively. Afterwards, affine transformations were applied in each image pair for coordinate alignment. Finally, after

localizing the OBI characters in each image, a Siamese network was used to compute the content similarity to discover duplicates.

OBI classification and retrieval

OBI Classification is normally to learn a mapping function $f_\theta : \mathcal{X} \rightarrow \mathcal{Y}$, predicting the character category y for a given input OBI image x . Given a query image, OBI retrieval aims to rank candidate images from a database based on their semantic or visual similarity to a query image. These tasks are particularly challenging due to the characteristics such as various appearances of OBI data and long-tailed distribution.

Supervised deep learning. The early OBI classification methods basically adopt the same models as those used in general image classification. As a pioneering work, Guo et al.⁴³ combined a Gabor-related low-level representation and a sparse-encoder-related mid-level representation for OBI characters, which is also complementary to CNN-based models. Huang et al.⁶⁹, Liu et al.⁴⁵, and Fu et al.³⁰ benchmarked traditional classification models including AlexNet¹²⁵, VGG¹²⁶, ResNet-50²⁸, ResNet-101²⁸, and Inception¹²⁷, on a large-scale OBI dataset (OBC306) and a self-curated damaged OBI dataset, which serves as baselines for the development of subsequent algorithms. Chen et al.³⁸ designed a two-step oracle materials classification, where *shield grain* and *tooth grain* are detected to divide multiple areas on an OBI image. A CNN is then used to extract the features of each region for final classification. Gao and Liang⁸⁹ utilized the isomorphism and symmetry invariance of OBIs to distinguish complicated oracle variants, thereby ensuring the classification accuracy. In addition, Jiang et al.¹²⁸ proposed OraclePoints, which represent OBI images as hybrid neural representations combining image and point-set features, making it easier and more effective to distinguish characters from noise. OraclePoints can be easily integrated with existing models, facilitating both classification and retrieval tasks.

Meanwhile, OBI retrieval approaches mainly focus on deep metric learning, which extends the classification paradigm by optimizing distance metrics. Liu and Wang¹²⁹ combined deep neural networks (DNN) and K-means clustering technology for the retrieval of OBI images. Ding et al.⁴⁰ adopted a Siamese network-based image retrieval method to learn feature representations of similar and dissimilar images. To overcome the vast morphological gap accumulated over millennia, Wu et al.⁴⁹ proposed a cross-font OBI image retrieval network, which employs a Siamese framework to extract deep features from character images with various fonts, exploring structure clues in different resolutions by multi-scale feature integration.

Zero-shot and few-shot learning. Essentially, OBI suffers from the problem of data limitation and imbalance, which therefore should be taken as a zero-shot or few-shot learning task. Wu et al.¹³⁰ introduced OracleGCD, a generalized category discovery (GCD)-based framework that simultaneously handles the recognition of known characters and the discovery and clustering of novel categories. The authors proposed a stroke-aware asymmetric view augmentation mechanism and a logit-based confidence-guided mechanism for feature enhancement and enabling supervised contrastive learning with virtual labels, respectively. In ref. 131, Hu et al. presented Ora-NSC, a semi-supervised approach for OBI fragment classification that integrates mean teacher with FixMatch, utilizing the exponential moving average strategy to enhance the accuracy and stability of pseudo-label prediction. Liu et al.⁵³ proposed a self-supervised learning approach with contrastive masked frequency modeling for OBI classification, named OBI-CMF, which can extract and learn both global and local features of OBI images. Besides, an inter-domain supervision is achieved by learning OBI features in the spatial and frequency domains, which significantly improves the robustness and generalization ability in OBI classification. Besides, mainstream OBI information systems are based on a collection of databases, which limits retrieval to known and paired OBI characters

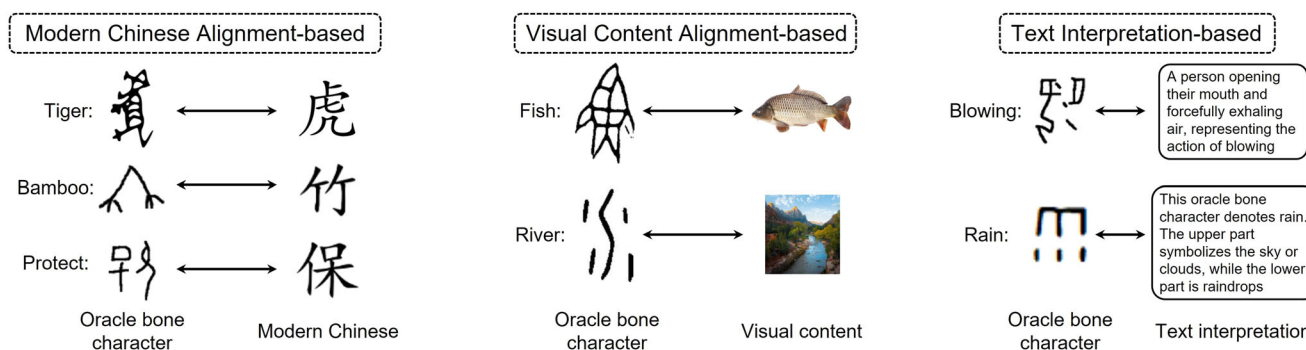


Fig. 12 | Illustration of mainstream OBI deciphering objectives. From left to right: modern Chinese alignment-based deciphering, visual content alignment-based deciphering, and text interpretation-based deciphering.

organized by experts. To reduce such reliance, Gao et al.⁴¹ presented a method named linking unknown characters (LUC) for searching arbitrary OBI characters, including both deciphered and undeciphered. Specifically, they designed a domain-aware embedding module to narrow the gap between glyph and scanned data, enhancing the utilization of radical prototypes of OBI characters. The core feature aggregation and dimension lifting layers are employed to get a unified and diverse representation for matching. Similar to the recognition task, researchers utilize radical information to facilitate OBI classification and retrieval. Diao et al.⁵¹ proposed a zero-shot OBI recognition framework via radical-based reasoning, which includes a radical information extractor and a knowledge graph-based character reasoner for OBI characters to search for their corresponding modern Chinese versions. Similarly, to support a finer-grained OBI deciphering, Hu et al.⁷⁹ proposed a component-level OBI retrieval task, i.e., using an OBI component to retrieve all OBI characters containing this component. A dual-stream attention-based model and two types of triplets based on components and characters were designed to model the relationships between components and characters.

Cross-modal learning. To mitigate the problem of the scarcity of OBI data, Wang et al.⁷⁵ proposed a structure-texture separation network (STSN) for joint disentanglement, transformation, adaptation, and recognition of different forms of OBI data, making full use of the majority of handwritten characters and performing cross-modal generation to enhance the discriminative ability of the learned features. The recent OracleAgent⁶³ further brought OBI retrieval to cross-modal vision-text scenarios. Benefit from the adopted agentic architecture and the extensive OBI knowledge base, OracleAgent can achieve diverse retrieval tasks such as character tracing and type finding using natural language.

OBI deciphering

Deciphering OBI remains a formidable challenge due to its abstract nature, structural diversity, and scarcity of glyphs, as well as the absence of contextual information. While the emergence of generative models has greatly changed the situation. Current OBI deciphering methods can be classified into three categories based on their objectives (Fig. 12), i.e., modern Chinese alignment-based methods, visual content alignment-based methods, and text interpretation-based methods.

Modern Chinese alignment-based deciphering. Different from the early recognition-based OBI deciphering, these methods normally employ generative models such as GAN or diffusion models to link OBI characters with modern Chinese. Chang et al.¹³² proposed Sundial-GAN, which, for the first time, utilizes the capability of multi-GANs to realize a simulation of the ancient character evolutionary process, i.e., transferring an OBI character to its modern Chinese versions. Following this, Guan et al.¹³³ trained a conditional diffusion model named OBSD that utilizes

seen categories of OBI characters as a conditional input to generate corresponding images of its modern counterpart. Benefiting from the proposed localized structural sampling technique, OBSD outperforms previous GAN-based frameworks in generation fidelity and recognizability. To make up for the lack of semantic alignment in OBSD, Wu et al.¹³⁴ presented DCSD-OBI, which incorporates both OBI images and modern Chinese text as dual conditions while integrating visual and semantic information to enhance structural reconstruction and semantic alignment during the reverse diffusion process.

In addition to the above generative model-based methods, there are also alternative approaches that utilize other feature representations for alignment. For example, in ref. 91, Zhang et al. took an unknown OBI character as a query and used an auto-encoder to retrieve suitable image representations in feature space, where subsequent fonts, such as seal character and modern Chinese, are also embedded. In ref. 80, the authors utilized the Siamese network in few-shot learning with binary cross-entropy loss to predict the evolution rules flow of Chinese characters. Wang et al.⁸³ proposed the Puzzle Pieces Picker (P3) method, which deconstructs OBI into foundational strokes and radicals first and then reconstructs them into their modern counterparts through a Transformer model, explicitly utilizing the structural relationships between ancient and modern characters. Similarly, Hu et al.¹³⁵ proposed a component-level OBI segmentation model that processes OBI characters and OBI components independently to perform accurate segmentation, thereby accelerating the understanding and deciphering process. In ref. 49, the authors utilized historical font intermediaries to overcome glyph gaps and enable effective decipherment of OBI characters. Inspired by the similarity of the elements throughout the evolutionary process, Jiao et al.⁸⁴ employed a graph-based representation for OBI and modern characters, wherein nodes correspond to key structural points of a character, and edges define their spatial relationships. They demonstrated a higher similarity of graph-based representation than image-based representation in the same character family, providing valuable data support for the semantic reasoning of undeciphered OBI characters. However, over thousands of years, originally pictorial lines may undergo straightening, simplification, or distortion for the sake of writing efficiency, causing the character to lose its original semantics. This happens especially in those undeciphered OBI characters, challenging the practical application of these approaches. Moreover, most failed generation cases targeting modern Chinese characters consist of meaningless strokes, offering little substantive assistance in speculating about the interpretation of OBIs.

Visual content alignment-based deciphering. Compared to pure character-aligned deciphering, visual content-aligned deciphering can rectify semantic drift, creating a direct cognitive bridge that bypasses the confusing errors introduced during thousands of years of stroke evolution. As a pioneering work, Qiao et al.¹³⁶ conducted a comprehensive human study to show whether participants could indeed make better

sense of an oracle glyph subjected to a proper visual guide. Meanwhile, a conditional visual generation task based on an oracle glyph and its semantic meaning was designed to circumvent model training in the presence of a fatal lack of oracle data. V-Oracle¹³⁷ applied principles of pictographic character formation and framed the OBI deciphering task as a visual question-answering (VQA) problem. The progressive three-stage training paradigm enables V-Oracle to perform advanced reasoning with pictographic explanations. Li et al.⁶¹ proposed OracleFusion, which leverages the LMM with enhanced spatial awareness reasoning to ensure a more structured and interpretable representation of oracle glyphs. They also introduced a glyph structure-constrained semantic typography method to generate object-form deciphering results with both semantic integrity and spatial accuracy. More recently, Chen et al. constructed a visual perception benchmark, PictOBI-20K⁵, to evaluate the visual decipherment ability of general-purpose LMMs on pictographic OBI characters. Besides, a subjective annotation was conducted to investigate the consistency of the reference point between humans and LMMs in visual alignment, offering valuable insights for optimizing the attention mechanism in visual decipherment.

Text interpretation-based deciphering. The most recognized and straightforward deciphering objective is to generate text interpretation. As the first initial attempt, Chen et al.² introduced OBI-Bench and conducted evaluations systematically on 23 LMMs for examining their potential OBI deciphering abilities. Experiments on OBI characters with different frequencies and formation principles (e.g., pictograph, ideogram, and radical-based variants) reveal the performance map of LMMs, pointing out the direction for subsequent research. At the same time, Jiang et al.⁶⁰ proposed OracleSage, a cross-modal framework built upon LLaVA-1.5¹³⁸ for OBI deciphering that integrates hierarchical visual understanding with graph-based semantic reasoning to enhance the plausibility of interpretations. In ref. 62, Peng et al. proposed an interpretable OBI decipherment method based on large vision-language models (LVLM), which combines radical analysis and pictograph-semantic understanding to bridge the gap between glyphs and meanings of OBI, further enhancing the reliability of the OBI deciphering process. More recently, Li et al.⁶³ presented OracleAgent, an LLM-empowered agentic system for structured management and retrieval of OBI-related information. Benefiting from the constructed large OBI knowledge base and the integration of multiple OBI analysis tools, OracleAgent can generate reliable character explanations and meet the diverse needs of public users.

Evaluation of OBI processing models

Objective metrics

The evaluation protocol of OBI processing models normally depends on the task type.

OBI recognition. For recognition tasks involving the localization of individual glyphs, the evaluation protocols largely follow standard object recognition benchmarks, e.g., PASCAL VOC¹⁰² and MS COCO¹⁰³. There are two commonly used metrics at the bounding box level⁵²:

- **Intersection over Union (IoU):** For a predicted bounding box B_p and a ground truth bounding box B_{gt} , the IoU is defined as:

$$IoU = \frac{Area(B_p \cap B_{gt})}{Area(B_p \cup B_{gt})} \quad (2)$$

- **Average Precision (AP) and mAP:** Precision (P) and Recall (R) are calculated based on True Positives (TP), False Positives (FP), and False Negatives (FN). The Average Precision (AP) for a specific class c is the

area under the Precision-Recall curve. In practice, this is often approximated by a discrete sum over rank positions. The mean average precision (mAP) is the mean of AP values over all C classes, i.e., all unique OBI characters:

$$AP_c = \int_0^1 P(R) dR, \quad mAP = \frac{1}{|C|} \sum_{c=1}^{|C|} AP_c \quad (3)$$

Proposal. Considering that the standard IoU is sensitive to pixel-level alignment of bounding boxes, which is unreliable given the degraded and fragmented nature of OBI carriers (turtle shells/bones), we suggest introducing stroke-weighting coefficients, i.e., Structure-Weighted IoU (SW-IoU). By utilizing skeletonization, the metric focuses on the character's core structural strokes rather than its weathered edges. Predictions that omit critical structural features should be penalized more heavily than those with only minor boundary offsets.

OBI rejoining. Regarding the OBI rejoining task at the fragment level, there are no explicit evaluation metrics. Most researchers use deep models to determine whether a combination is rejoinable and then calculate the classification accuracy^{33,34,54}, correlation coefficient¹²⁰, or confidence³⁶, followed by a practical rejoining job.

OBI classification. In the OBI character classification task, Top- k accuracy and Macro-F1 score are usually adopted:

- **Top- k accuracy:** Let y be the ground truth label and $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_k$ be the top- k predicted classes sorted by probability confidence. The indicator function $\mathbb{I}(\cdot)$ evaluates to 1 if the condition is true, and 0 otherwise.

$$Acc@k = \frac{1}{N} \sum_{i=1}^N \mathbb{I}(y_i \in \{\hat{y}_{i,1}, \dots, \hat{y}_{i,k}\}) \quad (4)$$

Where N is the total number of test samples.

- **Macro-F1 score:** To address the class imbalance that commonly exists in OBI datasets, the Macro-F1 averages the F1-score of each class independently. First, for each class $c \in C$:

$$F1_c = 2 \cdot \frac{P_c \cdot R_c}{P_c + R_c} \quad (5)$$

Then, the Macro-F1 is:

$$Macro-F1 = \frac{1}{|C|} \sum_{c=1}^{|C|} F1_c \quad (6)$$

Proposal. Since OBI exhibits significant "isomers" (one character with multiple variants), standard accuracy, which penalizes a model for identifying a valid historical variant if it doesn't match the specific character in the ground-truth, could introduce bias in multi-source classification scenarios. Therefore, we suggest implementing an OBI variant library to better reflect the model's semantic understanding ability, where if the model's prediction is a linguistically recognized variant of the ground truth, it is granted a partial or full score based on a predefined decay weight.

OBI retrieval. In OBI retrieval tasks, Rank- k , which measures the probability that a correct match appears within the top k retrieved results, is often adopted. Besides, unlike recognition, retrieval mAP considers the

rank of all relevant matches^{2,41,79,128}, especially when used in cases where a single OBI character has multiple variations.

- **Rank-*k***: For a set of *Q* queries:

$$\text{Rank-}k = \frac{1}{Q} \sum_{q=1}^Q \mathbb{I}(\text{rank}(q) \leq k) \quad (7)$$

where $\text{rank}(q)$ is the position of the first correct match in the retrieved list for query *q*.

- **mAP for Retrieval**: For a single query *q*, let Rel_k be an indicator that represents if the item at rank *k* is relevant, $P@k$ be the precision at rank *k*, and \mathcal{R}_{total} be the total number of relevant items.

$$AP_q = \frac{\sum_{k=1}^N (P@k \times \text{Rel}_k)}{\mathcal{R}_{total}}, \quad mAP = \frac{1}{Q} \sum_{q=1}^Q AP_q \quad (8)$$

OBI deciphering. As for the OBI deciphering models, their evaluation metrics can be broadly categorized into three groups based on the interpretation forms, i.e., alignment with modern Chinese characters, alignment with visual contents, and text interpretation. For the first type, OCR-related metrics such as recognition accuracy were commonly used^{60,83,132,134,135}. In ref. 133, the authors proposed OBS-OCR, a ResNet-101-based²⁸ classifier trained on 88,899 classes of modern Chinese characters for evaluating the deciphering results generated by diffusion models. Qiao et al.¹³⁷ standardized all OBI meaning descriptions into a multiple-choice format and used regex to match the LMMs' predictions for calculating their accuracy against the ground-truth answers. Moreover, Chen et al.⁵ proposed a visual character-object alignment accuracy metric and compared visual decipherment results directly with human performance. Similarly, OracleFusion⁶¹ used CLIPScore¹³⁹ to evaluate the semantic relevance between generated glyph morphing and the text prompt. Let *I* be the generated image, and *C* be the text prompt. The image and text are encoded into feature vectors $v = E_I(I)$ and $t = E_T(C)$, respectively. The CLIPScore is calculated as the scaled cosine similarity:

$$\text{CLIPScore}(I, C) = w \cdot \max\left(\frac{v \cdot t}{\|v\| \|t\|}, 0\right) \quad (9)$$

where *w* denotes a scaling factor (typically set to 2.5) used to align the score with human judgment. A higher CLIPScore indicates that the generated visual content aligns more accurately with the description. For pure textual interpretation assessment, researchers^{2,61,62} mainly utilize BLEU-4¹⁴⁰ and BERTScore¹⁴¹ to measure the consistency between the generated interpretation and expert-labeled golden descriptions. BLEU is a classic metric widely used in machine translation to evaluate the quality of generated text by measuring the overlap of *n*-grams between the candidate hypothesis and the reference text:

$$\text{BLEU} - 4 = BP \cdot \exp\left(\sum_{n=1}^4 w_n \log p_n\right) \quad (10)$$

where p_n is the modified *n*-gram precision. w_n is the weight for each *n*-gram, typically set as 1/4. *BP* is the brevity penalty, introduced to penalize generation that is significantly shorter than the reference:

$$BP = \begin{cases} 1 & \text{if } l_c > l_r \\ e^{(1-l_r/l_c)} & \text{if } l_c \leq l_r \end{cases} \quad (11)$$

where l_c is the length of the candidate text and l_r is the length of the reference corpus. Unlike BLEU, which relies on exact string matching, BERTScore evaluates semantic similarity by leveraging contextual embeddings from

pre-trained BERT models¹⁴². Given a reference sentence $x = \langle x_1, \dots, x_k \rangle$ and a candidate sentence $\hat{x} = \langle \hat{x}_1, \dots, \hat{x}_l \rangle$, their contextual embeddings are denoted as **x** and $\hat{\mathbf{x}}$. BERTScore computes the Recall (R_{BERT}), Precision (P_{BERT}), and F1-measure (F_{BERT}) based on cosine similarity:

$$\begin{aligned} R_{\text{BERT}} &= \frac{1}{|\hat{\mathbf{x}}|} \sum_{x_i \in \mathbf{x}} \max_{\hat{x}_j \in \hat{\mathbf{x}}} \mathbf{x}_i^T \hat{\mathbf{x}}_j, \\ P_{\text{BERT}} &= \frac{1}{|\mathbf{x}|} \sum_{\hat{x}_j \in \hat{\mathbf{x}}} \max_{x_i \in \mathbf{x}} \mathbf{x}_i^T \hat{\mathbf{x}}_j, \\ F_{\text{BERT}} &= 2 \cdot \frac{P_{\text{BERT}} \cdot R_{\text{BERT}}}{P_{\text{BERT}} + R_{\text{BERT}}} \end{aligned} \quad (12)$$

Proposal. Currently, both OCR-based and text interpretation-based evaluation strategies treat the characters to be deciphered as a whole, ignoring the meanings at radical- or component-level, hindering fine-grained analysis for the optimization of deciphering models. To this end, we suggest decomposing OBI characters into functional components and calculating a weighted coverage score for these critical features, i.e., the Radical Integrity Score (RIS).

$$\text{RIS} = \sum_{i=1}^n w_i \cdot \text{Score}(c_i) \quad (13)$$

where w_i represents the scholarly importance or knowledge graph-guided weight of the *i*th component (c_i) in the decipherment process.

OBI generation. Furthermore, some emerging OBI generation tasks, such as OBI restoration^{4,100} and synthesis³, still use traditional image quality assessment metrics, including PSNR, SSIM¹⁴³, LPIPS¹⁴⁴, and FID¹⁴⁵. However, they have been demonstrated to have a relatively low correlation with the actual OBI image quality, providing suboptimal, unreliable evaluations. The inherent gap between natural images and OBI images in low-level characteristics, distortion types, and high-dimensional feature spaces presents new challenges for developing evaluation metrics tailored to these tasks.

Human evaluation

Deciphering or interpretation, as the ultimate goal of OBI information processing, typically involves perceptual and linguistic data, requiring human perception, understanding, and validation. Nevertheless, there are no established evaluation practices, and researchers have been using different protocols^{2,61-63,133,137}. Typically, most works provide a rough description of the number of experts invited for annotation, without providing any other information. As shown in Fig. 13, prior studies faced several major problems, including qualification definition, ad-hoc task design, lack of quality check, and missing details reporting. This hinders the fair comparison of the proposed methods. Some relevant domains, such as subjective quality assessment in multimedia^{146,147}, text generation^{148,149}, image generation^{150,151}, and video generation^{57,152}, provide in-depth analysis and referable designs of human evaluation procedures. Inspired by these works, we anticipate developing a recognized evaluation protocol for OBI deciphering. Specifically, we propose six perspectives that require special attention, including transparency, rating quality, evaluation dimension, annotators, literature support, and whether it has been empirically validated, to facilitate reliable decipherment comparison based on human evaluation:

- **Transparency:** The degree to which the evaluation protocols, data flows, and scoring criteria are explicit and reproducible. We advocate for the full disclosure of prompt engineering protocols (for LMM-based evaluations) and the release of anonymized raw scoring datasets. This ensures that the research community can scrutinize the decision-making process and replicate the evaluation environment under identical constraints.

Current human evaluation for OBI decipherment

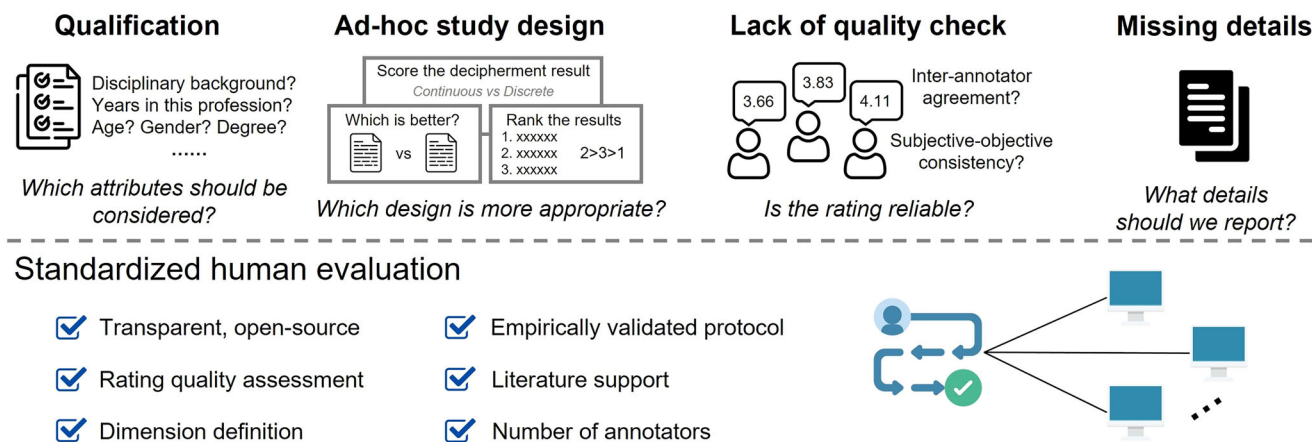


Fig. 13 | Some urgent points in the current human evaluation for OBI deciphering. We anticipate building a standardized evaluation procedure.

- **Rating Quality:** The reliability and consistency of the subjective scores provided by human evaluators. To mitigate individual bias, we suggest employing Inter-Rater Reliability (IRR) metrics, such as Cohen’s κ or Fleiss’ κ , to quantify consensus.
- **Evaluation Dimension:** The multi-faceted criteria used to assess the performance of OBI decipherment or recognition. Rather than a binary “correct/incorrect” metric, evaluation should encompass: (a) Semantic Accuracy: Alignment with established philological interpretations; (b) Morphological Fidelity: The degree to which the interpretation respects the structural evolution of the glyph; (c) Contextual Logic: The coherence of the deciphered character within the syntactic structure of the original inscription.
- **Annotators:** The number, professional background, expertise level, and selection criteria for the human participants. Given the high entry barrier of OBI studies, we categorize annotators into domain experts (e.g., senior epigraphists) and trained practitioners (e.g., graduate students in Paleography). We suggest documenting their recruitment process, prior experience, and the calibration training they received to align their scoring standards before the formal evaluation.
- **Literature Support:** The grounding of the evaluation ground-truth in authoritative scholarly references. Every evaluation must be benchmarked against canonical corpora, such as “*Corpus of Oracle-Bone Facsimiles*”¹⁸ or “*Collections of Oracle Bone Inscriptions*”¹⁴. For characters with ongoing scholarly debate, the framework must specify whether it adopts the majority consensus or allows for parallel valid interpretations supported by peer-reviewed literature.
- **Empirical validation:** The evidence demonstrating that the proposed evaluation framework accurately reflects model capabilities in real-world scenarios. We suggest performing correlation analysis between human scores and downstream task performance, ensuring the framework’s methodological robustness and practical utility.

Challenges and Outlook

Challenges

Breaking down data source barriers. The efficacy of modern deep learning paradigms is inextricably linked to the availability of large-scale, high-quality annotated data. However, in the realm of OBI research, a significant paradox persists. Although the number of unearthed OBI artifacts and related materials is considerable, the availability of open-source, machine-readable datasets remains a critical bottleneck. As listed in Table 2, among the datasets proposed over the past ten years, nearly half are not actually accessible.

Currently, the vast majority of high-fidelity OBI data, including rubbings, high-resolution photographs, and 3D scans, is sequestered within

data silos. These resources are typically held by museums, research institutes, or private collectors and are restricted by strict intellectual property protections or physical access barriers. Consequently, many existing studies rely on proprietary datasets constructed ad hoc by individual research teams. While these private datasets facilitate isolated experiments, they severely undermine the scientific rigor of the field. To overcome these barriers, the community must move toward an open science framework. This entails not only the digitization of OBI artifacts but also the establishment of clear legal frameworks for data sharing that balance institutional rights with academic accessibility. Future advancements will depend on the creation of a comprehensive, open-access repository that provides standard databases for different downstream tasks. Without breaking down these data source barriers, the development of robust, generalized OBI foundation models will remain stagnant.

Optimizing dataset construction pipelines. Once a new research direction has been selected, the targeting dataset must be constructed or transformed into a format that a modern AI model can understand. This involves establishing robust and reliable data collection and processing pipelines tailored to the unique characteristics of OBI-related information. Traditional data pipelines are purely human-involved, which usually requires a significant amount of manpower and time for proof-reading and annotation, requiring highly specialized knowledge. These processes present significant challenges for researchers with a cross-disciplinary background in computer science or other disciplines. Any errors or academic disputes within it will undermine the reliability of the resulting model. For instance, the existing largest OBI dataset, OBC306⁶⁹, has been found to have 120 classification issues, involving over 26,000 characters, as well as 74 errors and omissions caused by data segmentation, affecting 80 characters¹⁵³, which introduces bias when serving as a validation dataset for the OBI classification task. Nowadays, as the amount of data required for training high-performance models continues to increase, building automatic data processing pipelines with quality control standards and verification procedures that are in line with the consensus has become an urgent issue that needs to be addressed.

Representing heterogeneous OBI data. Early OBI processing models almost exclusively supported unimodal data, typically image or text, which hinders their applicability in cross-modality scenarios. A significant portion of OBI knowledge is expressed in heterogeneous and interleaved formats, presenting a profound challenge for training general models. Specifically, each OBI character can possess multiple appearances, such as original excavated form, rubbing fragment, cropped, and handprinted forms (Fig. 4). Meanwhile, due to the diversity of bone carving, each character will also have

multiple fonts. From the perspective of historical continuity, each character may also have subsequent historical variations. For the OBI deciphering task, each character may need to incorporate contextual information and modern interpretations in multiple languages. Such heterogeneity is challenging to represent and process within a standard CNN or transformer architecture. This requires innovative approaches that integrate sequential language embedding and complex data structures into a high-dimensional knowledge space, which often involve multimodal or hybrid architectures.

Future work

Text-to-OBI generation. In the era of generative AI, future work may be conducted towards text-to-OBI image generation. Current OBI image generation works, including denoising⁴, inpainting⁹⁹, and synthesis³, are all image-to-image. Developing fine-grained text-guided image translation models or models that directly translate modern language to the target OBI images can facilitate the restoration of noise-contaminated OBI images, creative content production, and the improvement of long-tailed OBI data. The development of such specialized text-to-image (T2I) models will also be essential to serve as a shared tool for historical linguistics, cognitive semiotics, and digital humanities, forming a sustainable cultural production chain.

Building OBI-oriented foundation models. General-purpose LLMs equipped with powerful reasoning capabilities have demonstrated their potential in various OBI processing tasks², especially in deciphering rare characters. However, due to the lack of fine-grained domain knowledge, a considerable gap still exists compared to human experts. Future work should prioritize the development of OBI-oriented foundation models that unify all OBI processing tasks in heterogeneous forms and modalities with interactive natural language, facilitating frontline researchers. On the data side, collecting and constructing large-scale OBI corpus datasets as training material remains a fundamental challenge, as understanding ancient text and related publications requires considerable domain knowledge. Moreover, a significant portion of OBI data exists in image form, presenting a major challenge in how to extract and construct a corpus or move beyond traditional text-dominated corpora. Key criteria, such as alignment, completeness, task relevance, usability, and accessibility, must be integrated into data collection pipelines. Overall, an OBI foundation model enables the encoding of OBI attributes, including visual morphology, structural patterns, associated semantic meanings, and archeological context into a coherent, unified vector space, thereby powering sophisticated downstream tasks.

Multi-agent collaborative OBI retrieval and interpretation. The recent prevailing LLMs have shown remarkable capabilities in tasks such as creative content generation, understanding, and reasoning¹⁵⁴. However, they still suffer from intrinsic limitations such as hallucination, execution efficiency, and scaling laws. Building upon this, agentic AI utilizes LLMs as the control center or the orchestrator, integrating them with external tools and knowledge bases, enabling autonomous planning, dispatching, verification, and refinement of complex tasks^{63,155–157}. The existing studies have demonstrated that traditional models trained from a single modality struggle to handle cross-modal OBI information retrieval due to the large gap between different forms of OBI data. Furthermore, the inherent limitations of a single LLM knowledge base prevent it from effectively inferring the meanings of oracle bone characters. Hence, a key direction for future research is to develop a multi-agent collaborative framework for OBI information retrieval and interpretation, building a more effective, robust, and trustworthy knowledge network that drives the development of popularized OBI research systems and increases the interpretability of deciphering results, thereby enabling historians to evaluate multiple hypotheses.

3D reconstruction-based OBI restoration. The existing OBI studies, including datasets and methodologies, without exception, focus on 2D OBI contents sourced from publications or websites, which have a

significant discrepancy from the original, firsthand oracle bone artifacts. Currently, the field of 3D representation has experienced substantial growth, which allows for the creation of realistic models and environments. It may be applied to OBI processing to enhance the visual fidelity, success rate, and efficiency in both restoration and reconstruction. However, due to restrictions on the use of original OBI data and for the sake of protecting cultural relics, conducting large-scale and open-ended 3D data collection may not be practical. To this end, we envision the development of 3D OBI datasets guided by the authorities, consisting of voxel grids, point clouds, meshes, neural radiance fields (NeRF), or 3D Gaussian splats (3DGS)¹⁵⁸, and regard it as a potential objective shift in OBI processing. This requires not only technical infrastructure but also new incentive structures within the community that reward data curation and open source. Integrating both 2D and 3D OBI representations will enable more high-quality and aligned databases that drive the development of archeology and the digital protection of cultural relics.

AI-generated content-empowered OBI dissemination. As digital humanities deepen and intelligent technologies evolve, emerging media like AI-generated content (AIGC) are set to revolutionize the preservation of OBI. By reshaping how this seminal cultural heritage is perceived and expressed, AIGC can facilitate its creative transformation. The future of OBI educational outreach may hinge on a new pedagogical paradigm driven by AIGC and multimodal imaging, rooted in perceptual embodiment and immersive situational reconstruction. For example, utilizing AIGC to create Shang-style art from OBI glyphs or produce OBI variants from user photos. This design could become a core teaching method that effectively bridges cognitive gaps, activates cultural memory, and enables learners to intuitively grasp the construction logic of Chinese characters.

Conclusion

In this survey, we conducted a comprehensive and up-to-date review of OBI information processing, starting with an introduction to pivotal advancements and paradigm-shifting milestones in the field. Afterwards, datasets and approaches for OBI recognition, rejoining, classification, retrieval, and deciphering tasks are summarized and analyzed in sequence. Emerging topics in the realm of OBI processing, such as generative model-based OBI image generation and denoising, large multimodal models-based OBI interpretation, and agentic system-based cross-modal OBI retrieval, are also reviewed. We further discuss potential issues in data quality, cognitive differences between non-OBI researchers and OBI-specific experts, evaluation criteria, and outline future directions, including stylized OBI generation and OBI-oriented foundation model development. Overall, this survey offers a consolidated reference and a forward-looking roadmap for researchers in related areas while facilitating researchers within the field to find solutions and trends in their study.

Data availability

Relevant papers and resources are continuously updated at <https://github.com/OBI-Future/OBI-Survey>.

Received: 12 February 2026; Accepted: 29 March 2026;
Published online: 08 April 2026

References

1. Liu, K., Wu, X. et al. Radiocarbon dating of oracle bones of late Shang period in ancient China. *Radiocarbon* **63**, 155–175 (2021).
2. Chen, Z., Chen, T., Zhang, W. & Zhai, G. Obi-bench: Can LLMs aid in the study of ancient script on oracle bones? In *The 13th International Conference on Learning Representations* (2025).
3. Li, J. et al. Mitigating long-tail distribution in oracle bone inscriptions: Dataset, model, and benchmark. In *Proceedings of the 33rd ACM International Conference on Multimedia*, 7729–7738 (2025).
4. Li, J., Chen, Z., Chen, T., Liu, Z. & Wang, C. Obiformer: A fast attentive denoising framework for oracle bone inscriptions. *Displays* **89**, 103059 (2025).

5. Chen, Z. et al. Pictobi-20k: Unveiling large multimodal models in visual decipherment for pictographic oracle bone characters. *arXiv preprint arXiv:2509.05773* (2025).
6. Xing, J., Liu, G. & Xiong, J. Oracle bone inscription detection: a survey of oracle bone inscription detection based on deep learning algorithm. In *Proceedings of the International Conference on Artificial Intelligence, Information Processing and Cloud Computing*, 39, 1–8 (2019).
7. Ren, S., He, K., Girshick, R. & Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.* **28**, 91–99 (2015).
8. Liu, W., Anguelov, D. et al. SSD: Single shot multibox detector. In *European Conference on Computer Vision*, 21–37 (Springer, 2016).
9. Redmon, J. & Farhadi, A. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767* (2018).
10. Zhang, S., Wen, L., Bian, X., Lei, Z. & Li, S. Z. Single-shot refinement neural network for object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4203–4212 (2018).
11. Liu, S., Huang, D. et al. Receptive field block net for accurate and fast object detection. In *Proceedings of the European Conference On Computer Vision (ECCV)*, 385–400 (2018).
12. Diao, X., Bo, R. et al. Ancient script image recognition and processing: A review. *arXiv preprint arXiv:2506.19208* (2025).
13. Li, J. et al. A comprehensive survey of oracle character recognition: Challenges, datasets, methodology, and beyond. *Pattern Recognit.* **169**, 111824 (2025).
14. Guo, M. *Collections of oracle-bone inscriptions* (The Zhong Hua Book Press, 1982).
15. Huang, D. *Palaeography* (Shanghai Classics Publishing House, Shanghai, 2019).
16. Yu, S. *Lexicon and Explanations of Oracle Bone Characters* (Zhonghua Book Company, Beijing, 2009).
17. Sun, Y. *Examples of Inscriptions on Bone and Tortoise Shells (Qiwen Juli) and Mingyuan* (Zhonghua Book Company, Beijing, 2016).
18. Huang, T. *Corpus of Oracle-Bone Facsimiles* (Peking University Press, 2022).
19. Dong, Z. *Examples of Chronological Research on Oracle Bone Inscriptions* (Institute of History and Philology, Academia Sinica, Beijing, 1933).
20. Liu, Z. *Studies on the Morphological Formation of Ancient Chinese Characters* (Fujian People's Publishing House, Fuzhou, 2011).
21. Chen, T. *A Re-examination of the Graphical System of Oracle Bone Characters of the Yin and Shang Dynasties*. Ph.D. thesis (2007).
22. Wang, A., Ge, Y., Ge, W., Zhou, H. & Wang, D. System design for computer-aided rejoining of bones/tortoise shells with inscriptions based on contour matching. In *2010 IEEE 12th International Conference on Communication Technology*, 841–844 (IEEE, 2010).
23. Liao, C.-W. & Huang, J. S. A transformation invariant matching algorithm for handwritten chinese character recognition. *Pattern Recognit.* **23**, 1167–1188 (1990).
24. Guo, L. Characters feature extraction based on neat oracle bone rubbings. *TELKOMNIKA Indonesian J. Electr. Eng.* **11**, 5427–5434 (2013).
25. Liu, Y.-G., Wang, T.-L. & Wang, J.-P. The application of the technique of 2d fragments stitching based on outline feature in rejoining oracle bones. In *2010 International Conference on Multimedia Information Networking and Security*, 964–968 (IEEE, 2010).
26. Wang, A., Ge, Y. & Liu, G. Research on key technologies of the computer aided rejoining of oracle bone inscriptions. In *2010 2nd IEEE International Conference on Information and Financial Engineering*, 180–183 (IEEE, 2010).
27. Yongge, L. & Yanqiang, G. Study on digital watermarking in the oracle bone inscriptions rubbings. In *2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 925–928 (IEEE, 2009).
28. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778 (2016).
29. Liu, Z. et al. A convnet for the 2020s. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 11976–11986 (2022).
30. Fu, X., Yang, Z., Zeng, Z., Zhang, Y. & Zhou, Q. Improvement of oracle bone inscription recognition accuracy: A deep learning perspective. *ISPRS Int. J. Geo-Inf.* **11**, 45 (2022).
31. Fujikawa, Y. et al. Recognition of oracle bone inscriptions by using two deep learning models. *Int. J. Digit. Hum.* **5**, 65–79 (2023).
32. Wen, X., Setthawong, R. & Sun, H. Oracle bone inscriptions recognition based on deep learning. In *2025 IEEE 2nd International Conference on Electronics, Communications and Intelligent Science (ECIS)*, 1–7 (IEEE, 2025).
33. Zhang, C., Zong, R., Cao, S., Men, Y. & Mo, B. Ai-powered oracle bone inscriptions recognition and fragments rejoining. In *Proceedings of the 29th International Conference on International Joint Conferences on Artificial Intelligence*, 5309–5311 (2021).
34. Zhang, C. et al. Data-driven oracle bone rejoining: A dataset and practical self-supervised learning scheme. In *Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining*, 4482–4492 (2022).
35. Zhang, Z. et al. Deep rejoining model and dataset of oracle bone fragment images. *npj Herit. Sci.* **13**, 66 (2025).
36. Zhang, Z., Wang, Y.-T., Li, B., Guo, A. & Liu, C.-L. Deep rejoining model for oracle bone fragment image. In *Asian Conference on Pattern Recognition*, 3–15 (Springer, 2021).
37. Qiao, Y. & Xing, L. Automatic classification method for oracle images based on deep learning. *IEIE Trans. Smart Process. Comput.* **12**, 87–96 (2023).
38. Chen, S., Xu, H., Weize, G., Xuxin, L. & Bofeng, M. A classification method of oracle materials based on local convolutional neural network framework. *IEEE Computer Graph. Appl.* **40**, 32–44 (2020).
39. Zhu, M. et al. Oracle bone inscriptions image recognition and classification method based on deep learning. In *2024 IEEE 7th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, vol. 7, 1960–1965 (IEEE, 2024).
40. Ding, J., Wang, J., Aysa, A., Xu, X. & Ubul, K. Oracle bone inscription image retrieval based on improved ResNet network. In *International Conference on Pattern Recognition*, 47–62 (Springer, 2024).
41. Gao, F. et al. Linking unknown characters via oracle bone inscriptions retrieval. *Multimed. Syst.* **30**, 125 (2024).
42. Li, B. et al. An open benchmark for oracle bone rubbing image retrieval. *npj Herit. Sci.* **13**, 292 (2025).
43. Guo, J., Wang, C. et al. Building hierarchical representations for oracle character and sketch recognition. *IEEE Trans. Image Process.* **25**, 104–118 (2016).
44. Zhang, Y.-K., Zhang, H., Liu, Y.-G., Yang, Q. & Liu, C.-L. Oracle character recognition by nearest neighbor classification with deep metric learning. In *2019 International Conference on Document Analysis and Recognition (ICDAR)*, 309–314 (IEEE, 2019).
45. Liu, M., Liu, G., Liu, Y. & Jiao, Q. Oracle bone inscriptions recognition based on deep convolutional neural network. *J. Image Graph.* **8**, 114–119 (2020).
46. Zheng, Y., Chen, Y., Wang, X., Qi, D. & Yan, Y. Ancient Chinese character recognition with improved swin-transformer and flexible data enhancement strategies. *Sensors* **24**, 2182 (2024).
47. Chen, J., Zhang, H. & Zhang, Y. Decoding the past: Solving challenging oracle bone characters recognition problem by integrating vision transformer and generative adversarial image restoration technique. In *IEEE 8th International Conference on Vision, Image and Signal Processing*, 1–7 (IEEE, 2024).

48. Yang, Z., Han, Z., Aysa, A., Ibrahim, G. & Ubul, K. Long-tailed oracle character recognition based on convolutional neural networks and vision transformers. In *ICASSP 2025–2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1–5 (IEEE, 2025).
49. Wu, Z., Su, Q., Gu, K. & Shi, X. A cross-font image retrieval network for recognizing undeciphered oracle bone inscriptions. In *International Conference on Intelligent Computing*, 196–208 (Springer, 2025).
50. Huang, H., Yang, D. et al. Aagtgan: Unpaired image translation for photographic ancient character generation. In *Proceedings of the 30th ACM international conference on multimedia*, 5456–5467 (2022).
51. Diao, X. et al. Rzcr: zero-shot character recognition via radical-based reasoning. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence*, 654–662 (2023).
52. Diao, X. et al. Toward zero-shot character recognition: a gold standard dataset with radical-level annotations. In *Proceedings of the 31st ACM International Conference on Multimedia*, 6869–6877 (2023).
53. Liu, J. et al. Obi-cmf: Self-supervised learning with contrastive masked frequency modeling for oracle bone inscription recognition. *npj Herit. Sci.* **13**, 55 (2025).
54. Yuan, J., Shanxiong, C., Weize, G., Maling, P. & Lihua, J. Sff-siam: a new oracle bone rejoining method based on siamese network. *IEEE Comput. Graph. Appl.* **43**, 22–32 (2023).
55. Guo, D. et al. Deepseek-r1 incentivizes reasoning in LLMs through reinforcement learning. *Nature* **645**, 633–638 (2025).
56. Xu, J. et al. Qwen3-omni technical report. *arXiv preprint arXiv:2509.17765* (2025).
57. Chen, Z. et al. Gaia: Rethinking action quality assessment for AI-generated videos. *Adv. Neural Inf. Process. Syst.* **37**, 40111–40144 (2024).
58. Chen, Z. et al. Just noticeable difference for large multimodal models. *arXiv preprint arXiv:2507.00490* (2025).
59. Chen, Z. et al. Can large models fool the eye? A new Turing test for biological animation. *arXiv preprint arXiv:2508.06072* (2025).
60. Jiang, H. et al. Oraclesage: Towards unified visual-linguistic understanding of oracle bone scripts through cross-modal knowledge fusion. *arXiv preprint arXiv:2411.17837* (2024).
61. Li, C. et al. Oraclefusion: Assisting the decipherment of oracle bone script with structurally constrained semantic typography. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 19893–19902 (2025).
62. Peng, K., Zhao, M., Yu, H., Fu, T. & Li, B. Interpretative oracle bone script decipherment through radical and pictographic analysis with lvmms. *arXiv preprint arXiv:2508.10113* (2025).
63. Li, C. et al. Oracleagent: A multimodal reasoning agent for oracle bone script research. *arXiv preprint arXiv:2510.26114* (2025).
64. Li, Y. et al. Perception, reason, think, and plan: A survey on large multimodal reasoning models. *arXiv preprint arXiv:2505.04921* (2025).
65. Huang, T. *Categorization and time division of the Oracle inscriptions for the king in Yin ruins* (Science Press, 2007).
66. Liu, Z. *New Compendium of Oracle-Bone Characters* (Fujian People's Publishing House, 2009).
67. Zhang, Q. & Yang, Y.-B. Rest: An efficient transformer for visual recognition. *Adv. neural Inf. Process. Syst.* **34**, 15475–15485 (2021).
68. Kirillov, A. et al. Segment anything. In *Proceedings of the IEEE/CVF international conference on computer vision*, 4015–4026 (2023).
69. Huang, S., Wang, H., Liu, Y., Shi, X. & Jin, L. Obc306: A large-scale oracle bone character recognition dataset. In *2019 International Conference on Document Analysis and Recognition (ICDAR)*, 681–688 (IEEE, 2019).
70. Li, B. et al. Hwobc-a handwriting oracle bone character recognition database. *J. Phys.: Conf. Ser.* **1651**, 012050 (2020).
71. Han, W., Ren, X., Lin, H., Fu, Y. & Xue, X. Self-supervised learning of orc-bert augmentor for recognizing few-shot oracle characters. In *Asian Conference on Computer Vision*, 652–668 (2020).
72. Xuzheng, L., Hengjin, C. & Li, L. Recognition of oracle radical based on the capsule network. *CAAI Trans. Intell. Syst.* **15**, 243–254 (2020).
73. Yue, X., Li, H., Fujikawa, Y. & Meng, L. Dynamic dataset augmentation for deep learning-based oracle bone inscriptions recognition. *ACM J. Comput. Cult. Herit.* **15**, 1–20 (2022).
74. Pu, M. *Oracle Bone Inscriptions in the Collection of Shanghai Museum (Volume I)* (Shanghai Lexicographical Publishing House, 2009).
75. Wang, M., Deng, W. & Liu, C.-L. Unsupervised structure-texture separation network for oracle character recognition. *IEEE Trans. Image Process.* **31**, 3137–3150 (2022).
76. Lin, X., Chen, S., Zhao, F. & Qiu, X. Radical-based extract and recognition networks for oracle character recognition. *Int. J. Doc. Anal. Recognit. (JDAR)* **25**, 219–235 (2022).
77. Wu, Z. *Shang and Zhou Bronze Inscriptions and Image Integration* (Shanghai Classics Publishing House, 2012).
78. Wang, M. & Deng, W. A dataset of oracle characters for benchmarking machine learning algorithms. *Sci. Data* **11**, 87 (2024).
79. Hu, Z., Cheung, Y.-m., Zhang, Y., Zhang, P. & Tang, P.-l. Component-level oracle bone inscription retrieval. In *Proceedings of the 2024 International Conference on Multimedia Retrieval*, 647–656 (2024).
80. Wang, M. et al. Study on the evolution of Chinese characters based on few-shot learning: From oracle bone inscriptions to regular script. *Plos One* **17**, e0272974 (2022).
81. Guan, H. et al. An open dataset for the evolution of oracle bone characters: Evobc. *arXiv preprint arXiv:2401.12467* (2024).
82. Wang, P. et al. An open dataset for oracle bone script recognition and decipherment. *arXiv preprint arXiv:2401.15365* (2024).
83. Wang, P. et al. Puzzle pieces picker: Deciphering ancient Chinese characters with radical reconstruction. In *International Conference on Document Analysis and Recognition*, 169–187 (Springer, 2024).
84. Jiao, Q. et al. A graph-based evolutionary dataset for oracle bone characters from inscriptions to modern chinese scripts. *npj Herit. Sci.* **13**, 369 (2025).
85. Shi, D. et al. Rcrn: Real-world character image restoration network via skeleton extraction. In *Proceedings of the 30th ACM International Conference on Multimedia*, 1177–1185 (2022).
86. Li, B. et al. OBIMD: A multi-modal dataset for contextual interpretation of oracle bone inscriptions. *Scientific Data* (2026).
87. Meng, L., Fujikawa, Y., Ochiai, A., Izumi, T. & Yamazaki, K. Recognition of oracular bone inscriptions using template matching. *Int. J. Comput. Theory Eng.* **8**, 53–57 (2016).
88. Meng, L., Kamitoku, N. & Yamazaki, K. Recognition of oracle bone inscriptions using deep learning based on data augmentation. In *2018 metrology for archaeology and cultural heritage (MetroArchaeo)*, 33–38 (IEEE, 2018).
89. Gao, J. & Liang, X. Distinguishing oracle variants based on the isomorphism and symmetry invariances of oracle-bone inscriptions. *IEEE Access* **8**, 152258–152275 (2020).
90. Zhang, Y. et al. Generating oracle bone inscriptions based on the structure aware diffusion model. *npj Herit. Sci.* **13**, 461 (2025).
91. Zhang, G., Liu, D., Smyth, B. & Dong, R. Deciphering ancient Chinese oracle bone inscriptions using case-based reasoning. In *International Conference on Case-Based Reasoning*, 309–324 (Springer, 2021).
92. Yang, H. et al. Large kernel convolutional attention based u-net network for inpainting oracle bone inscription. In *Chinese Conference on Pattern Recognition and Computer Vision (PRCV)*, 117–129 (Springer, 2023).
93. Wang, N. et al. Multi-modal ancient scripts recognition via deep learning with data homogenization and augmentation. *npj Herit. Sci.* **13**, 522 (2025).

94. Wang, W. et al. Improving oracle bone characters recognition via a cycleGAN-based data augmentation method. In *International Conference on Neural Information Processing*, 88–100 (Springer, 2022).
95. Gao, J. et al. Oracle bone heritage data augmentation based on two-stage decomposition GANs. *npj Herit. Sci.* **13**, 230 (2025).
96. Ding, Z. et al. Radical-conditioned diffusion model for oracle bone character generation and analysis. In *The 7th International Symposium on Advanced Technologies and Applications in the Internet of Things*, 39–50 (2025).
97. Meng, Z. et al. Orpaint: a zero-shot inpainting model for oracle bone inscription rubbings with visual mamba block. *Sci. China Inf. Sci.* **68**, 189102 (2025).
98. Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V. & Courville, A. C. Improved training of Wasserstein GANs. *Adv. neural Inf. Process. Syst.* **30**, 5769–5779 (2017).
99. Wang, S., Guo, W., Xu, Y., Liu, D. & Li, X. Coarse-to-fine generative model for oracle bone inscriptions inpainting. In *Proceedings of the 1st Workshop on Machine Learning for Ancient Languages (ML4AL 2024)*, 107–114 (2024).
100. Diao, X. et al. Oracle bone inscription image restoration via glyph extraction. *npj Herit. Sci.* **13**, 321 (2025).
101. Liu, Y. et al. Vmamba: Visual state space model. *Adv. Neural Inf. Process. Syst.* **37**, 103031–103063 (2024).
102. Everingham, M., Van Gool, L., Williams, C. K., Winn, J. & Zisserman, A. The pascal visual object classes (voc) challenge. *Int. J. Computer Vis.* **88**, 303–338 (2010).
103. Lin, T.-Y. et al. Microsoft Coco: Common objects in context. In *European conference on computer vision*, 740–755 (Springer, 2014).
104. Zhou, X.-L., Hua, X.-C. & Li, F. A method of jia gu wen recognition based on a two-level classification. In *Proceedings of 3rd ICDAR*, vol. 2, 833–836 (IEEE, 1995).
105. Meng, L. Recognition of oracle bone inscriptions by extracting line features on image processing. In *JCPRAM*, 606–611 (2017).
106. L., Meng. Two-stage recognition for oracle bone inscriptions. In *International conference on image analysis and processing*, 672–682 (Springer, 2017).
107. Qu, H., Liu, J. & Wu, J. Oracle-bone inscriptions recognition based on topological features. *Comput. Sci. Appl.* **9**, 1111–1117 (2019).
108. Zhao, Z.-Q., Zheng, P., Xu, S. & Wu, X. Object detection with deep learning: A review. *IEEE Trans. neural Netw. Learn. Syst.* **30**, 3212–3232 (2019).
109. Girshick, R. Fast R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision*, 1440–1448 (2015).
110. Liu, Z. et al. Oracle character detection based on improved faster R-CNN. In *2021 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, 697–700 (IEEE, 2021).
111. Meng, L., Lyu, B. et al. Oracle bone inscription detector based on SSD. In *International Conference on Image Analysis and Processing*, 126–136 (Springer, 2019).
112. Liu, G., Xing, J. & Xiong, J. Spatial pyramid block for oracle bone inscription detection. In *Proceedings of the 2020 9th International Conference on Software And Computer Applications*, 133–140 (2020).
113. Wang, N., Sun, Q., Jiao, Q. & Ma, J. Oracle bone inscriptions detection in rubbings based on deep learning. In *2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, vol. 9, 1671–1674 (IEEE, 2020).
114. Zhen, Q., Wu, L. & Liu, G. An oracle bone inscriptions detection algorithm based on improved Yolov8. *Algorithms* **17**, 174 (2024).
115. Meng, X., Pu, H. & Meng, F. Automatic segmentation of oracle bone inscriptions using YOLOv8. *Procedia Comput. Sci.* **242**, 1074–1081 (2024).
116. Xiong, R., Liu, J., Zhang, S. & Liao, Z. Fdw-yolo: An improved yolov12 for oracle bone inscriptions detection. In *International Conference on Neural Information Processing*, 252–267 (Springer, 2025).
117. Li, B. et al. Oracle character prototype-guided cyclic disentanglement for oracle bone inscriptions detection. In *International Conference on Pattern Recognition and Artificial Intelligence*, 212–226 (Springer, 2024).
118. Fu, X., Zhou, R., Yang, X. & Li, C. Detecting oracle bone inscriptions via pseudo-category labels. *Heritage Science* **12** (2024).
119. Tao, Y., Fu, X., Pang, H., Yang, X. & Li, C. Clustering-based feature representation learning for oracle bone inscriptions detection. *npj Herit. Sci.* **13**, 296 (2025).
120. Tian, Y., Gao, W., Liu, X., Chen, S. & Mo, B. The research on rejoining of the oracle bone rubbings based on curve matching. *Trans. Asian Low.-Resour. Lang. Inf. Process.* **20**, 1–17 (2021).
121. Zhang, Z., Guo, A. & Li, B. Internal similarity network for rejoining oracle bone fragment images. *Symmetry* **14**, 1464 (2022).
122. Zhang, C. et al. Obd-finder: Explainable coarse-to-fine text-centric oracle bone duplicates discovery. *arXiv preprint arXiv:2505.03836* (2025).
123. DeTone, D., Malisiewicz, T. & Rabinovich, A. Superpoint: Self-supervised interest point detection and description. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 224–236 (2018).
124. Lindenberger, P., Sarlin, P.-E. & Pollefeys, M. Lightglue: Local feature matching at light speed. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 17627–17638 (2023).
125. Krizhevsky, A., Sutskever, I. & Hinton, G. E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **25**, 1097–1105 (2012).
126. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. In *the 3rd International Conference on Learning Representations* (2015).
127. Szegedy, C., Ioffe, S., Vanhoucke, V. & Alemi, A. Inception-v4, inception-resnet and the impact of residual connections on learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31 (2017).
128. Jiang, R. et al. Oraclepoints: A hybrid neural representation for oracle character. In *Proceedings of the 31st ACM International Conference on Multimedia*, 7901–7911 (2023).
129. Guoying, L. & Yangguang, W. Oracle character image retrieval by combining deep neural networks and clustering technology. *IAENG Int. J. Comput. Sci.* **47**, 109–116 (2020).
130. Wu, H., Li, K., Zhang, X.-Y., Wang, Q. & Wang, D.-H. Oraclegcd: Generalized category discovery for oracle bone scripts. In *International Conference on Document Analysis and Recognition*, 459–475 (Springer, 2025).
131. Hu, Y., Gao, W., Si, Y., Zhang, Y. & Liu, J. Ora-nsc: A novel semi-supervised approach for oracle bone fragment classification with imbalanced classes. In *Proceedings of the 7th ACM International Conference on Multimedia in Asia*, 1–8 (2025).
132. Chang, X., Chao, F., Shang, C. & Shen, Q. Sundial-gan: A cascade generative adversarial networks framework for deciphering oracle bone inscriptions. In *Proceedings of the 30th ACM International Conference on Multimedia*, 1195–1203 (2022).
133. Guan, H. et al. Deciphering oracle bone language with diffusion models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 15554–15567 (2024).
134. Wu, R., Wang, S., Li, X. & Liu, D. A text image dual conditional stable diffusion model for oracle bone inscription decipherment. *npj Herit. Sci.* **13**, 453 (2025).
135. Hu, Z., Cheung, Y.-m., Zhang, Y., Peiying, Z. & Ling, T. P. Component-level segmentation for oracle bone inscription decipherment. In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 39, 28116–28124 (2025).

136. Qiao, R., Yang, L., Pang, K. & Zhang, H. Making visual sense of oracle bones for you and me. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 12656–12665 (2024).
137. Qiao, R. et al. V-oracle: Making progressive reasoning in deciphering oracle bones for you and me. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 20124–20150 (2025).
138. Liu, H., Li, C., Wu, Q. & Lee, Y. J. Visual instruction tuning. *Adv. neural Inf. Process. Syst.* **36**, 34892–34916 (2023).
139. Hessel, J., Holtzman, A., Forbes, M., Le Bras, R. & Choi, Y. Clipscore: A reference-free evaluation metric for image captioning. In *Proceedings of the 2021 Conference on Empirical Methods In Natural Language Processing*, 7514–7528 (2021).
140. Papineni, K., Roukos, S., Ward, T. & Zhu, W.-J. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, 311–318 (2002).
141. Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q. & Artzi, Y. Bertscore: Evaluating text generation with BERT. In the *8th International Conference on Learning Representations* (2020).
142. Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference Of The North American Chapter of the Association for Computational Linguistics: Human Language Technologies, volume 1 (long and short papers)*, 4171–4186 (2019).
143. Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600–612 (2004).
144. Zhang, R., Isola, P., Efros, A. A., Shechtman, E. & Wang, O. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE Conference on Computer Vision And Pattern Recognition*, 586–595 (2018).
145. Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B. & Hochreiter, S. Gans trained by a two time-scale update rule converge to a local Nash equilibrium. *Advances in Neural Information Processing Systems* **30** (2017).
146. Chen, Z. et al. Band-2k: Banding artifact noticeable database for banding detection and quality assessment. *IEEE Trans. Circuits Syst. Video Technol.* **34**, 6347–6362 (2024).
147. Chen, Z. et al. Study of subjective and objective naturalness assessment of AI-generated images. *IEEE Trans. Circuits Syst. Video Technol.* **35**, 3573–3588 (2025).
148. Chiang, W.-L. et al. Chatbot Arena: an open platform for evaluating llms by human preference. In *Proceedings of the 41st International Conference on Machine Learning*, 8359–8388 (2024).
149. Karpinska, M., Akoury, N. & Iyyer, M. The perils of using Mechanical Turk to evaluate open-ended text generation. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 1265–1285 (2021).
150. Otani, M. et al. Toward verifiable and reproducible human evaluation for text-to-image generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 14277–14286 (2023).
151. Liang, Y. et al. Rich human feedback for text-to-image generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 19401–19411 (2024).
152. Zhang, Z. et al. Benchmarking multi-dimensional aigc video quality assessment: A dataset and unified model. *ACM Trans. Multimed. Comput. Commun. Appl.* **21**, 1–24 (2025).
153. Chen, T. Yinjiwenyuan obc306 dataset correction. *Sinogram Culture* 93–95 (2024).
154. Zhang, Z. et al. Aibench: Towards trustworthy evaluation under the 45'law. *Displays* **91**, 103255 (2026).
155. Maldonado, D., Cruz, E., Torres, J. A., Cruz, P. J. & Benitez, S.dP. G. Multi-agent systems: A survey about its components, framework and workflow. *IEEE Access* **12**, 80950–80975 (2024).
156. Chen, Z., Sun, Y., Tian, Y., Zhang, W. & Zhai, G. Maceval: A multi-agent continual evaluation network for large models. *arXiv preprint arXiv:2511.09139* (2025).
157. Chen, Z., Zhang, W. & Zhai, G. Evaluating from benign to dynamic adversarial: A squid game for large language models. *arXiv preprint arXiv:2511.10691* (2025).
158. Wang, Z. 3d representation methods: A survey. *arXiv preprint arXiv:2410.06475* (2024).

Acknowledgements

This work was supported in part by the National Social Science Fund of China under Grant 24BY005, in part by the Shanghai Office of Philosophy and Social Science under Grant 2023BY003, in part by the New Generation Artificial Intelligence-National Science and Technology Major Project under Grant 2025ZD0124104 in collaboration with Shanghai Artificial Intelligence Laboratory, and in part by the Shanghai Jiao Tong University's Key Project on Intelligent Humanities under Grant ZHWK2506.

Author contributions

Z.C. and T.C. conceived and initiated this research project. Z.C. and W.H. performed the literature search and synthesized the historical data on Oracle Bone Inscriptions. Dataset collection was performed by Z.C., W.H., and J.L. Z.C. drafted the initial manuscript. Y.Z. wrote the initial draft for the future work part. X.Z., Z.L., and W.Z. critically revised the manuscript for intellectual content. The final manuscript underwent thorough proofreading and review by X.Z., Z.L., T.C., W.Z., and G.Z. to ensure accuracy and clarity. T.C. and G.Z. conceived and supervised the project. All authors discussed the final version and approved the manuscript for submission.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Zijian Chen, Tingzhu Chen or Guangtao Zhai.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

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

© The Author(s) 2026