

Interpretable Oracle Bone Script Decipherment via Progressive Glyph Analysis with LVLMs

Kaixin Peng* Mengyang Zhao* Haiyang Yu*,[◇] Teng Fu Bin Li[†]

College of Computer Science and Artificial Intelligence, Fudan University, Shanghai, China

{kxpeng25, tfu23}@m.fudan.edu.cn

{myzhao20, hyyu20, libin}@fudan.edu.cn

* Equal Contribution [◇] Project Lead [†] Corresponding Author

Abstract

Deciphering scanned Oracle Bone Script (OBS) images into corresponding modern Chinese characters is a crucial step in AI-assisted OBS decipherment. As a highly pictographic writing system, OBS often embeds semantic information in glyph forms, making its decipherment depend on more than visual similarity alone. However, how to effectively exploit such glyph-meaning cues and how to make the decipherment process interpretable remain challenging. To this end, we propose an LVLM-based framework that bridges glyphs and meanings through structured glyph analysis. Specifically, we introduce a progressive training strategy that guides the model from radical analysis to pictographic analysis and finally to mutual analysis, enabling it to capture semantic information embedded in OBS glyphs. The generated textual analyses serve as both interpretable reasoning traces and explicit evidence for retrieving corresponding modern Chinese characters from a dictionary. We also construct the Pictographic Decipherment OBS Dataset (PD-OBS), which provides detailed radical and pictographic analyses to support glyph-analysis learning and dictionary-based retrieval. Experiments on OBS benchmarks demonstrate the strong zero-shot capability of our method, especially its substantially improved Top-10 accuracy on unseen OBS, highlighting its ability to provide reliable and interpretable candidate references for AI-assisted archaeological and historical studies. The code and dataset will be released at <https://github.com/PKXX1943/PD-OBS>.

1 Introduction

Oracle Bone Script (OBS), one of the earliest mature writing systems, preserves valuable evidence of early human civilization. Recent studies have increasingly focused on mapping scanned OBS images to corresponding modern Chinese characters (Guan et al., 2024b; Jiang et al., 2024; Li et al., 2025a), a crucial task for AI-assisted OBS decipher-

ment that addresses a fundamental and challenging step in the broader process of OBS Decipherment. However, this task remains difficult due to the rarity, abstraction, and diversity of OBS glyphs, especially in real scenarios where models must generalize to unseen characters and provide interpretable evidence for expert verification. For easier reading, we explain core OBS terms including glyph, radical and pictographic analysis in Appendix A.1.

Existing deep learning-based methods have substantially advanced OBS decipherment. As summarized in Figure 1, current approaches mainly fall into three paradigms: classification-based, composition-based, and diffusion-based methods. Classification-based methods (Guo et al., 2022; Luo et al., 2023; Zheng et al., 2024; Gan et al., 2023; Lin et al., 2022; Jiang et al., 2023) perform well in closed-set settings, but they cannot handle unseen OBS classes. Composition-based methods (Shi et al., 2025; Wang et al., 2024b; Hu et al., 2024, 2025) improve zero-shot capability and provide coarse structural cues by decomposing glyphs into sub-components and reassembling predictions. Diffusion-based methods (Guan et al., 2024b; Li et al., 2023) further enhance recognition accuracy and zero-shot capability through conditional control and sampling strategies.

However, these advances still leave two key questions underexplored: how to effectively exploit the semantic clues embedded in OBS glyph forms, especially for unseen characters, and how to make the decipherment process interpretable. Prior studies (Qiao et al., 2024; Li et al., 2025a) have shown that radicals often convey basic semantic cues, while pictographic forms are often closely related to meaning. Motivated by this observation, we propose an LVLM-based framework that bridges OBS glyphs and meanings through structured glyph analysis. Specifically, we propose a progressive training strategy that guides the model from radical analysis to pictographic analysis and

finally to mutual analysis, allowing it to progressively capture the rich semantic information embedded in OBS glyphs. Based on the generated analysis contents, we further introduce a Radical-Pictographic Dual Matching mechanism to retrieve candidate decipherments from a dictionary. By grounding predictions in explicit analysis evidence, our method improves generalization while ensuring an interpretable and reliable decipherment process.

Such structured glyph analysis requires resources that go beyond image-label pairs, including fine-grained glyph-meaning annotations and a reliable retrieval dictionary. To this end, we construct the Pictographic Decipherment OBS Dataset (PD-OBS), which provides detailed radical and pictographic analyses for 3,173 OBS classes and expands them into a 47,157-character dictionary for evidence-based retrieval. By linking OBS glyphs, semantic analyses, and modern Chinese candidates, PD-OBS supports our analysis-to-retrieval framework and offers a reusable benchmark for future research on OBS decipherment and glyph analysis.

Our main contributions are as follows:

- We propose an interpretable LVLM-based framework for OBS decipherment that bridges glyphs and meanings through progressive radical, pictographic, and mutual analysis.
- We construct the PD-OBS dataset with multiple OBS images and related ancient and modern characters, annotated with detailed radical and pictographic analyses, providing a well-structured benchmark for OBS research.
- Our method demonstrates strong zero-shot decipherment ability, especially with substantially higher Top-10 accuracy on unseen OBS characters, enabling it to provide reliable candidate sets as interpretable references for AI-assisted archaeological and historical studies.

2 Related Work

2.1 Oracle Bone Script Datasets

With the expansion of digitized resources, numerous high-quality OBS datasets (Li et al., 2020; Hu et al., 2025; Han et al., 2020; Huang et al., 2019; Yue et al., 2022; Chen et al., 2025; Li et al., 2026, 2025b; Diao et al., 2025) have been released. Since Oracle-20K (Guo et al., 2016), data volume and quality have improved significantly. Currently, HUST-OBC (Wang et al., 2024a) and

EVOBC (Guan et al., 2024a) are the most widely adopted benchmarks: HUST-OBC collects 77,064 deciphered and 62,989 undeciphered images across 1,588 classes, derived from books, websites, and prior datasets; EVOBC contains 229,170 images spanning 13,714 classes across six historical script stages, collected from authoritative literature and websites. Both datasets serve as the primary data sources for constructing our PD-OBS dataset and as benchmarks for experimental evaluation.

2.2 Oracle Bone Script Decipherment

Classification model-based approaches (Meng, 2017; Zhou et al., 1995; Zheng et al., 2024; Lin et al., 2022; Jiang et al., 2023; Fujikawa et al., 2021; Dosovitskiy et al., 2021; Li et al., 2026, 2025b; Diao et al., 2025) for OBS decipherment have demonstrated performance comparable to or even surpassing that of human archaeologists in the closed-set setting. Enhanced Inception-V3 (Guo et al., 2022) employs convolutional attention modules to improve decipherment based on a CNN backbone, while the Pyramid Graph Transformer (Gan et al., 2023) integrates a pyramid-structured ViT with skeleton graph representations, attaining state-of-the-art closed-set results.

However, classification-based methods fail to handle unseen OBS classes, limiting their applicability to undeciphered characters. Wang *et al.* (Wang et al., 2024b) decompose OBS structural components and align them with modern Chinese radicals, though significant glyph structure discrepancies result in limited accuracy. The diffusion-based methods (Guan et al., 2024b; Li et al., 2023) achieves remarkable accuracy via conditional character generation, yet suffers from instability and a lack of interpretability. OracleSage (Jiang et al., 2024) is the first LVLM-based method for OBS analysis, but remains suboptimal and unreliable due to insufficient exploitation of glyph features and dictionary resources. These limitations highlight the need for a method that jointly addresses generalization, interpretability, and accuracy. To this end, we propose a novel LVLM-based framework, along with a large pictographic decipherment dataset, to advance OBS decipherment.

3 Pictographic Decipherment OBS Dataset

Dataset Collection. To alleviate the challenge of applying LVLMs to OBS decipherment, we in-

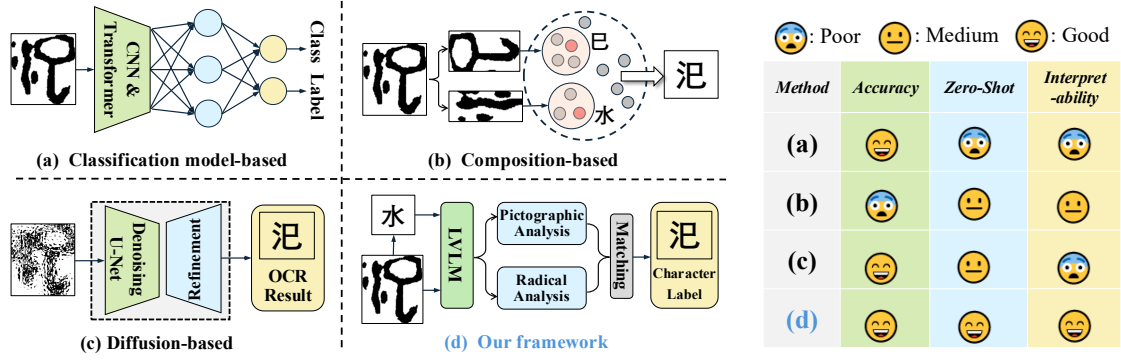


Figure 1: Qualitative summary of three existing paradigms and our paradigm for OBS decipherment. “Poor / Medium / Good” reflects trends in Table 1: Accuracy and Zero-shot correspond to the validation and zero-shot metrics. Interpretability is defined by the level of intermediate analysis: no interpretable cues (Poor), coarse structural hints (Medium), and fine-grained reasoning (Good).

troduce the Pictographic Decipherment OBS (PD-OBS) dataset to train LVLMs with the capability for analyzing radicals and pictographs, which is crucial for OBS decipherment. PD-OBS comprises 47,157 Chinese characters: 3,173 characters are associated with OBS images from HUST-OBC(Wang et al., 2024a) and EVOBC(Guan et al., 2024a); 10,968 characters are provided with ancient Clerical Script images from glyph repositories(Jerry); and all characters are accompanied by modern regular script images from Han Dian(Han Dian, 2004-2025). Each character is further annotated with detailed radical and pictographic analysis texts closely related to its semantic meaning. The original annotations are in Chinese due to the inclusion of numerous ancient and modern Chinese characters, but have been translated into English throughout this paper.

Dataset Annotation. The annotation process is conducted in three stages, as illustrated in Figure 2. First, radical labels, definitions, and explanations are retrieved from Shuowen Jiezi(Shen, 1963) (an ancient Chinese dictionary) and the authoritative dictionary Han Dian(Han Dian, 2004-2025). Second, the acquired radical labels and explanations are associated with each character’s modern, ancient, and OBS forms, with a powerful commercial LVLM further enriching the radical labels and summarizing the analysis contents based on the referenced glyph images. Finally, automated self-checking with commercial LVLM and manual review are performed to correct non-standard or inaccurate annotations.

Dataset Usage. PD-OBS is utilized in two key stages of our method. We construct multi-modal, multi-turn dialogue training samples by pairing

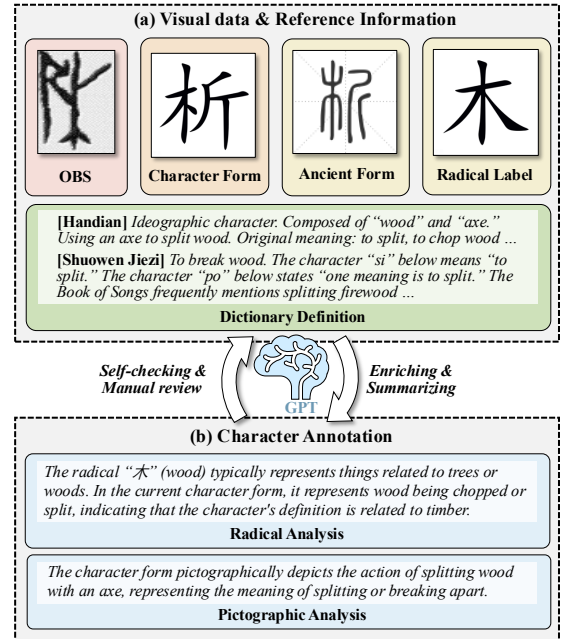


Figure 2: The demonstration of our data engine.

OBS images with corresponding modern character labels, enhancing the LVLM’s basic capacity to understand OBS glyphs. We also group all characters by radical tags and encode character annotation texts into feature vectors using BERT (Devlin et al., 2019), forming a Chinese character–pictograph analysis dictionary $\mathcal{D}$ for matching and verifying decipherment outputs. More details are presented in the Appendix A.2.

4 Method

4.1 Framework

Our framework is built upon Qwen2.5-VL-7B (Bai et al., 2025), sharing the same vision encoder and

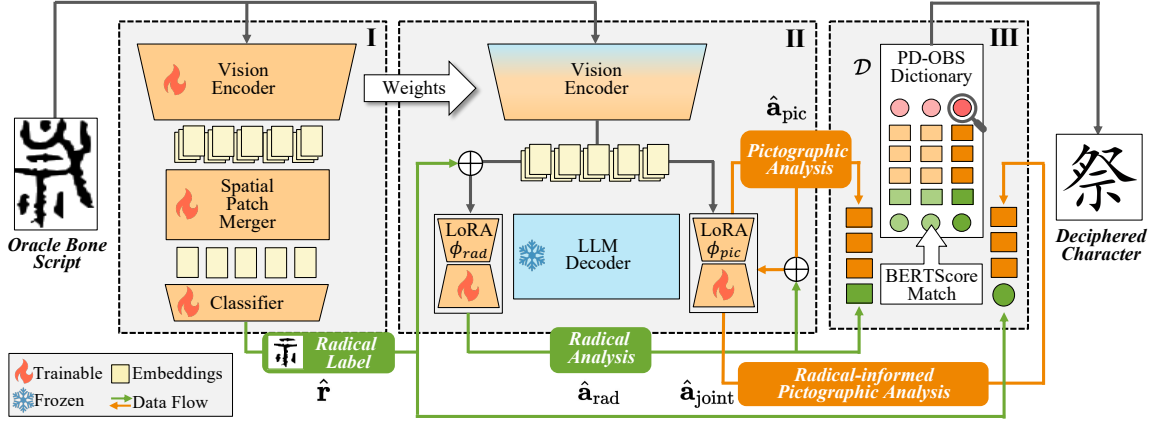


Figure 3: The framework of our method. **I** is used to indicate the Radical Recognition stage, and **II** is used to indicate the Radical-Pictographic Mutual Analysis stage, while **III** is used to indicate Radical-Pictographic Dual Matching.

LLM module. As illustrated in Figure 3, we introduce a spatial patch merger as the visual adapter and a classifier to predict the radical label. We also propose a radical LoRA module ϕ_{rad} and a pictographic LoRA module ϕ_{pic} (Hu et al., 2021) to analyze the corresponding information. Furthermore, we design a progressive training—starting from radical recognition (Sec. 4.2), followed by radical and pictographic analysis (Sec. 4.3), and culminating in mutual analysis—to gradually lead the model to bridge the gap between glyphs and meanings of OBS. Finally, we propose a novel radical-pictographic dual matching mechanism (Sec. 4.4) to select the appropriate modern Chinese characters from the database as the decipherment result. More details of spatial patch merger, batch sampling, and LoRA training are provided in **Appendix A.3**.

4.2 Radical Recognition

In this stage, we aim to adapt the vision encoder to the unique visual style of OBS and to predict radical labels that will serve as critical cues for downstream reasoning. For this purpose, we designed a spatial patch merger as a visual adapter, which downsamples the visual embeddings to obtain a representative feature vector for classification tasks. In addition, we design a triplet loss (Schroff et al., 2015) based on the Euclidean distance to explicitly improve the differentiation of feature vectors with different radicals.

Specifically, we implement a sampling strategy to ensure that each batch contains at least two samples for each radical class. During training, for each sample in the batch, we designate its feature vector $\mathbf{v}_n$ as an anchor sample, then select a positive sam-

ple $\mathbf{v}_n^+$ (i.e., a sample with the same radical label) and a negative sample $\mathbf{v}_n^-$ (i.e., a sample with a different radical label). Let $d_n^+ = \|\mathbf{v}_n - \mathbf{v}_n^+\|_2$ and $d_n^- = \|\mathbf{v}_n - \mathbf{v}_n^-\|_2$ denote the Euclidean distances to the positive and negative samples, respectively. The triplet loss is as follows:

$$\mathcal{L}_{\text{trip}} = \frac{1}{N} \sum_{n=1}^N \max(d_n^+ - d_n^- + \alpha, 0) \quad (1)$$

where N is the number of triplets in the batch and α is the margin hyperparameter.

Regarding the classifier, we use the cross-entropy loss $\mathcal{L}_{\text{ce}}$ to optimize it. Therefore, the whole loss function $\mathcal{L}_{\text{stage1}}$ of this stage can be shown as follows:

$$\mathcal{L}_{\text{stage1}} = \mathcal{L}_{\text{ce}} + \gamma \mathcal{L}_{\text{trip}}, \quad (2)$$

where γ is a hyperparameter used to balance the two items.

4.3 Radical-Pictographic Mutual Analysis

To bridge glyphs and meaning in OBS, we design a progressive glyph analysis process to facilitate the decipherment task. Specifically, we introduce a progressive training procedure, beginning with radical analysis, where the radical is predicted in the radical recognition stage. In both OBS and ancient Chinese characters, radicals often determine the basic semantics of characters, as illustrated by the Q1&A1 in Figure 4. Therefore, we train the model’s radical analysis capability with a large number of radical-analysis Q&A pairs constructed from the PD-OBS dataset. Next, we guide the model to perform a pictographic analysis for

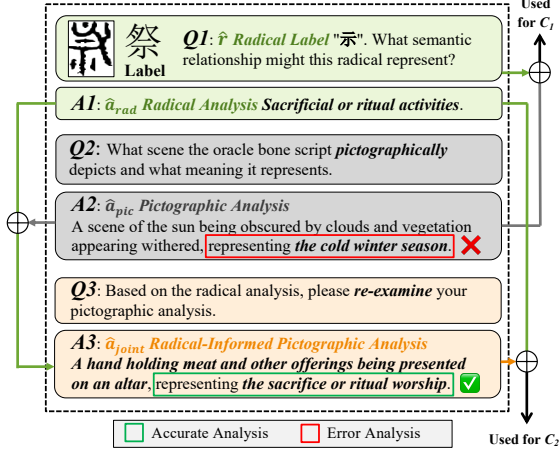


Figure 4: The workflow of radical-pictographic mutual analysis.

the entire character to analyze the meaning embedded in the full character glyph, as shown by the Q2&A2 in Figure 4.

In our observations, direct pictographic analysis of LVLM may yield erroneous results, likely due to the lack of prior knowledge or the overall abstract nature of the script, as shown in the A2 of Figure 4. Therefore, we design a mutual analysis as the final step, which informs the pictographic analysis with insights from radical analysis, resulting in more accurate character meanings. Specifically, we employ radical analysis $\hat{a}_{\text{rad}}$ and pictographic analysis $\hat{a}_{\text{pic}}$ as contextual information, prompting LVLM to re-examine pictographic analysis and generate the radical-informed pictographic analysis $\hat{a}_{\text{joint}}$ as illustrated by Q3&A3 in Figure 4. This enables the model to explicitly consider the basis of radical-implicated information, thereby mitigating the difficulty of directly analyzing the entire character. More qualitative examples of mutual analysis are provided in **Appendix A.6.1**

During training, we initialize the visual encoder with the weights from the previous stage, freezing the shallow layers to retain low-level features while fine-tuning the deeper layers for high-level semantic adaptation. In addition, we introduce a radical LoRA module ϕ_{rad} and a pictographic LoRA module ϕ_{pic} (Hu et al., 2021), the former for radical analysis while the latter for pictographic and mutual analysis. The training data consists of Q&A pairs from PD-OBS dataset, as shown in Figure 4, and the loss function employed is the cross-entropy loss commonly used in LVLM training.

Algorithm 1 Radical-Pictographic Dual Matching

Require: Dictionary $\mathcal{D}$
 $\mathcal{D} = \{(\mathbf{r}_i, \mathbf{a}_{\text{rad},i}, \mathbf{a}_{\text{pic},i}, \mathbf{a}_{\text{joint},i}, y_i)\}_{i=1}^N$
Require: Model output $(\hat{\mathbf{r}}, \hat{\mathbf{a}}_{\text{rad}}, \hat{\mathbf{a}}_{\text{pic}}, \hat{\mathbf{a}}_{\text{joint}})$
Require: Parameter k (Top k)
Require: $\mathcal{S}(\cdot, \cdot)$: Semantic similarity via BERT-Score.

- 1: // Filtered Matching
- 2: $\mathcal{D}_{\text{rad}} \leftarrow \{i \mid \mathbf{r}_i = \hat{\mathbf{r}}\}$
- 3: $C_1 \leftarrow \text{Top-}k \text{ in } \mathcal{D}_{\text{rad}} \text{ by } \mathcal{S}(\mathbf{a}_{\text{pic},i}, \hat{\mathbf{a}}_{\text{pic}})$
- 4: // Joint Matching
- 5: $C_2 \leftarrow \text{Top-}k \text{ in } \{1, \dots, N\} \text{ by } \mathcal{S}(\mathbf{a}_{\text{rad},i} \oplus \mathbf{a}_{\text{joint},i}, \hat{\mathbf{a}}_{\text{rad}} \oplus \hat{\mathbf{a}}_{\text{joint}})$
- 6: // Dual Matching
- 7: $C \leftarrow C_1 \cup C_2$
- 8: $R \leftarrow \text{Top-}k \text{ in } C \text{ by similarity scores}$
- 9: **return** $\{y_i \mid i \in R\}$

4.4 Radical-Pictographic Dual Matching

For each test character, previous stages produce four intermediate outputs: predicted radical label $\hat{\mathbf{r}}$, radical analysis $\hat{\mathbf{a}}_{\text{rad}}$, pictographic analysis $\hat{\mathbf{a}}_{\text{pic}}$, and radical-informed pictographic analysis $\hat{\mathbf{a}}_{\text{joint}}$. Single-path retrieval may suffer from error accumulation: an incorrect radical may exclude the target, while noisy pictographic analysis may retrieve wrong semantic neighbors. Thus, we propose Radical-Pictographic Dual Matching, an error-tolerant fusion mechanism over complementary radical and pictographic evidence.

As shown in Algorithm 1, given the dictionary $\mathcal{D} = \{(\mathbf{r}_i, \mathbf{a}_{\text{rad},i}, \mathbf{a}_{\text{pic},i}, \mathbf{a}_{\text{joint},i}, y_i)\}_{i=1}^N$, where y_i denotes a modern Chinese character, we first build a filtered matching set C_1 by selecting the Top- k entries with $\mathbf{r}_i = \hat{\mathbf{r}}$ according to $\mathcal{S}(\mathbf{a}_{\text{pic},i}, \hat{\mathbf{a}}_{\text{pic}})$, computed by BERT-Score (Zhang et al., 2020). We then build a joint matching set C_2 by selecting the Top- k entries from $\mathcal{D}$ according to $\mathcal{S}((\mathbf{a}_{\text{rad},i} \oplus \mathbf{a}_{\text{joint},i}), (\hat{\mathbf{a}}_{\text{rad}} \oplus \hat{\mathbf{a}}_{\text{joint}}))$. Finally, we merge and re-rank $C_1 \cup C_2$ to obtain Top- k decipherment results. Additional ablations and robustness analyses of Dual Matching are provided in **Appendix A.5**.

5 Experiments

5.1 Implementation Details

All training and evaluation experiments are conducted on 8 NVIDIA RTX 4090 GPUs. We initialize our model with the pretrained weights of Qwen2.5-VL-7B. During the radical recognition stage, we set the learning rate to 5e-4, batch size N to 8, and train for 5 epochs. The hyperparameters γ and α in the loss functions are set to 5 and 0.25, respectively. For the radical-pictographic mu-

Method	Validation (Seen-OBS Recognition)		Zero-shot (Unseen-OBS Decipherment)	
	HUST-OBC	EVOBC	HUST-OBC	EVOBC
<i>Classification model-based</i>				
InceptionV3 (Guo et al., 2022)	74.4 / 76.9	62.4 / 64.5	- / -	- / -
ViT (Dosovitskiy et al., 2021)	79.2 / 81.7	72.7 / 74.2	- / -	- / -
PyGT (Gan et al., 2023)	84.3 / <u>87.6</u>	78.1 / <u>81.2</u>	- / -	- / -
<i>Composition-based</i>				
PPP (Wang et al., 2024b)	76.8 / -	72.4 / -	13.6 / -	19.1 / -
<i>Commercial LVL</i>				
GPT-4.1 (OpenAI, 2024)	6.0 / 10.4	4.5 / 8.4	5.3 / 7.3	4.3 / 9.2
Qwen-VL-Max (Bai et al., 2025)	4.8 / 6.6	4.1 / 5.6	2.0 / 2.5	4.0 / 6.2
Gemini-2.5-Pro	6.3 / 13.9	5.1 / 10.4	5.0 / 8.6	6.2 / 15.0
GPT-5 (OpenAI, 2025)	7.2 / 16.1	5.3 / 12.5	6.4 / 9.8	6.0 / 14.3
<i>Diffusion-based</i>				
OBSD (Guan et al., 2024b)	66.8 / 72.9	71.2 / 77.9	18.3 / <u>27.5</u>	<u>30.4</u> / <u>50.5</u>
BBDM (Li et al., 2023)	55.8 / 59.5	60.3 / 62.1	8.0 / 14.1	19.5 / 29.5
Ours	<u>80.6</u> / 87.8	<u>76.3</u> / 81.7	<u>16.8</u> / 53.7	33.3 / 64.1
<i>Improvement</i>	-3.7 / +0.2	-1.8 / +0.5	-1.5 / +26.2	+2.9 / +13.6

Table 1: Each cell reports Top-1 / Top-10 accuracy (in %). **Zero-shot results are primary metric**, as they evaluate generalization to unseen OBS. The best and second-best results are respectively marked in bold and underlined. *Improvement* represents the performance gains achieved by our method compared to the existing best method.

tual analysis stage, we use a learning rate of $5e-5$, batch size of 4, and train for 4,000 steps. AdamW (Loshchilov and Hutter, 2019) is used as the optimizer. The radical LoRA ϕ_{rad} and pictographic LoRA ϕ_{pic} are configured with a dropout rate of 0.05 and 0.25, respectively, and both use rank and alpha values of 32.

5.2 Datasets and Evaluation Metrics

We perform experiments on the commonly used HUST-OBC (Wang et al., 2024a) and EVOBC (Guan et al., 2024a) datasets. Following prior OBS studies (Guan et al., 2024b; Li et al., 2023), we evaluate zero-shot generalization on held-out deciphered characters, which provides a reproducible approximation of unseen-character scenarios while avoiding the subjectivity and cost of expert validation on truly undeciphered cases. We select 200 character classes from each dataset as the zero-shot test sets. The remaining data are randomly split into training and validation sets in a 9:1 ratio to assess the OBS recognition capabilities.

We use Top-k accuracy as an evaluation metric, as in previous work (Guan et al., 2024b; Gan et al., 2023; Jiang et al., 2024), which is usually used in diverse classification tasks (Luo et al., 2023; Lin et al., 2022). To evaluate the interpretability of the model, we quantify the consistency between the analyzes content generated and the annotations.

Inspired by evaluation practices in image captioning (Galliena et al., 2025; Wang et al., 2022) and abstractive summarization (Liu et al., 2022), we adopt ROUGE-L (Lin, 2004), METEOR (Banerjee and Lavie, 2005), and BERT-Score (Zhang et al., 2020) to provide a complementary and holistic assessment of the generated analyzes.

5.3 Main Results

Decipherment Result. We conduct comprehensive comparisons as shown in Table 1. It should be clarified that in OBS decipherment research, the primary evaluation focus is the zero-shot setting (Guan et al., 2024b; Li et al., 2025a), which reflects a model’s ability to handle previously unseen characters—the core difficulty in real archaeological scenarios. By contrast, validation performance primarily indicates recognition capabilities for seen characters rather than decipherment ability, serving as a secondary auxiliary metric. We adopt InceptionV3 (Guo et al., 2022), ViT (Dosovitskiy et al., 2021), and PyGT (Gan et al., 2023) as classification model-based baselines, and OBSD (Guan et al., 2024b) and BBDM (Li et al., 2023) as diffusion-based methods. In addition, we include strong commercial LVLs, GPT-4.1 (OpenAI, 2024), Qwen-VL-Max (Bai et al., 2025), Gemini-2.5-Pro (Gemini Team, Google, 2025), and GPT-5 (OpenAI, 2025) for comparison. However, commercial

Method	HUST-OBC		EVOBC	
	Validation	Zero-shot	Validation	Zero-shot
Qwen2.5-VL-7B	0.355 / 0.358 / 0.694	0.309 / 0.301 / 0.651	0.341 / 0.350 / 0.683	0.337 / 0.348 / 0.679
Qwen-VL-Max	0.391 / 0.402 / 0.705	0.335 / 0.334 / 0.656	0.378 / 0.383 / 0.698	0.359 / 0.355 / 0.682
GPT-4.1	0.465 / 0.477 / 0.737	0.407 / 0.412 / 0.675	0.429 / 0.434 / 0.714	0.413 / 0.419 / 0.709
Gemini-2.5-Pro	0.486 / 0.501 / 0.745	0.421 / 0.419 / 0.712	0.436 / 0.447 / 0.713	<u>0.529 / 0.538 / 0.749</u>
GPT-5	<u>0.572 / 0.575 / 0.783</u>	<u>0.470 / 0.468 / 0.725</u>	<u>0.520 / 0.521 / 0.755</u>	0.498 / 0.501 / 0.740
Ours	0.914 / 0.907 / 0.946	0.550 / 0.525 / 0.794	0.887 / 0.884 / 0.937	0.576 / 0.586 / 0.849

Table 2: Interpretability performance comparison between different methods based on the ROUGE-L / METEOR / BERT-Score.

LVLMS perform poorly in both settings, with Top-1 accuracy consistently below 8%, highlighting their limited capability to understand OBS. On the validation set, although our method yields slightly lower Top-1 accuracy than the best classification model-based baseline (e.g., PyGT), it achieves the highest Top-10 accuracy, demonstrating superior capability in generating high-quality candidates. In the more important and challenging zero-shot scenario, our method exhibits notably strong performance: It remains competitive in Top-1 accuracy with the SOTA method OBSD and significantly outperforms all methods in Top-10 accuracy, surpassing the second-best method by 26.2% on HUST-OBC and 13.6% on EVOBC. These results confirm our method’s strong generalization and transferability to unseen OBS, highlighting its potential value in assisting the recognition of undeciphered OBS in archaeological research. Detailed baseline discussions and the commercial LVLMS evaluation protocol are provided in **Appendix A.4**.

Interpretability Performance. We employ Rouge-L (Lin, 2004), METEOR (Banerjee and Lavie, 2005), and BERT-Score (Zhang et al., 2020) to measure the similarity between the analysis text of Top-1 outputs and the ground truth text from the dictionary $\mathcal{D}$. We evaluate LVLMS, including Qwen2.5-VL-7B, Qwen-VL-Max, GPT-4.1, Gemini-2.5-Pro, and GPT-5, and compare their average performance with our method. As shown in Table 2, our method significantly outperforms the powerful commercial LVLMS GPT-5 and Gemini-2.5-Pro across three metrics. This result indicates that the analysis generated by our method is more reliable and informative.

5.4 Ablation Study

More detailed ablations on matching mechanisms, Top-K settings, dictionary scale, and incorrect radi-

Method	Valid.	ZS
Vision Encoder of Qwen	92.7	87.1
+ Our Radical Recognition	93.6	88.3
+ LoRA-based Recognition	80.1	69.8

Table 3: Radical recognition accuracy (in %) of different model variants on the HUST-OBC dataset with validation (Valid.) and zero-shot (ZS) settings.

Method	Validation		Zero-shot	
	Top-1	Top-10	Top-1	Top-10
Qwen2.5-VL-7B	1.4	1.4	0.2	0.2
+ Pictographic Analysis	52.4	52.4	1.6	1.6
+ Rad&Pic Mutual Analysis	60.3	61.4	5.2	5.4
+ Radical Recognition	64.2	64.2	6.6	6.6
+ Dual Matching	80.6	87.8	16.8	53.7
+ joint Matching C_2	69.1	80.5	14.9	45.8

Table 4: Top-1 and Top-10 accuracy (in %) of our model and its variants on HUST-OBC dataset.

cal prediction are reported in **Appendix A.5**.

Ablation on Radical Recognition Stage. We use the original vision encoder of Qwen2.5-VL-7B (Bai et al., 2025) as the baseline, and incorporate our radical recognition module or a LoRA-based recognition method. Our method introduces a spatial patch merger and the loss function $\mathcal{L}_{\text{trip}}$ on top of the baseline vision encoder, resulting in improvements of 0.9% and 1.2% accuracy on the validation and zero-shot settings, respectively, as shown in Table 3. The LoRA-based recognition method merges the recognition stage with the radical analysis process and training with LoRA-based fine-tuning. The results indicate that this method leads to a significant drop in radical recognition accuracy and introduces substantial errors in radical analysis, thus, we retain radical recognition as an independent stage in our method.

Data Type	OBS Image	GT	OBSD Top-3 Results	Our Top-3 Results	Radical Analysis (Ours)	Pictographic Character Form Analysis (Ours)
#Validation		吹 吹	[吹] [吹] [吹]	[吹] [喝] [唤]	Radical “口” is related to the mouth or openings; in the current character form it symbolizes an open mouth	A person opening their mouth and forcefully exhaling air, representing the action of blowing
		逐 逐	[逐] [逐] [逐]	[逐] [追] [追]	Radical “辶” is related to walking or running; in the current character form it represents the footsteps of a person running	A pig running with another person chasing behind it, representing the meaning of pursuit and driving away
		陟 陟	[陟] [陟] [陟]	[陟] [陞] [昇]	Radical “阝” indicates hillsides or elevated terrain; in the current character form it represents steps	A person walking step by step up stairs, representing the meaning of climbing or ascending
#Zero-shot		春 春	[春] [春] [春]	[春] [管] [晴]	Radical “日” indicates the sun; in the current character form it represents sunshine	A scene of plants growing and all things reviving under the radiant sunshine, representing the arrival of spring
		穆 穆	[穆] [穆] [穆]	[梨] [禾] [穆]	Radical “禾” is related to crops and agricultural plants; in the current character form it symbolizes ripe rice grains	A mature grain plant, representing the meaning of a bountiful harvest
		姝 姝	[姝] [姝] [姝]	[姝] [姝] [姝]	Radical “女” is related to feminine qualities; in the current character form it represents a working woman	A woman working in wheat fields, representing the virtuous qualities of women being diligent and thrifty in managing households
		司 司	[司] [司] [司]	[司] [后] [替]	Radical “口” is related to the mouth or openings; in the character form it symbolizes a person speaking	An ancient official reading a proclamation, expressing the meaning of management and execution
		叟 叟	[叟] [叟] [叟]	[叟] [叟] [叟]	Radical “又” indicates hands; in the current character form it appears as the image of a hand grabbing something	A hand holding a staff, representing the meaning of an elderly elder
		𠂔 𠂔	[𠂔] [𠂔] [𠂔]	[巡] [尢] [𠂔]	Radical “辶” is related to rivers, unrelated to the current character's pictographic meaning	A person walking around, representing the meaning of patrol or inspection

Figure 5: Visualization of the decipherment results and the interpretable content. Green rectangles and fonts indicate correct results and content, while red rectangles and fonts indicate errors. The leftmost item of the OBSD Top-3 results represents the corresponding modern Chinese character image generated by this method.

Ablation on the Analysis and Matching Strategies.

We take Qwen2.5-VL-7B (Bai et al., 2025) as the baseline and incrementally add each component, as shown in Table 4. Pictographic Analysis fine-tuning (+ Pictographic Analysis) greatly improves validation performance but still generalizes poorly to zero-shot scenarios. Introducing Radical-Pictographic Mutual Analysis (+ Rad&Pic Mutual Analysis) further improves validation accuracy, while the zero-shot gain remains limited. Using radical recognition (+ Radical Recognition) to guide mutual analysis brings similar incremental improvements. These results indicate that direct prediction still lacks sufficient generalization for unseen characters, which is a common challenge in zero-shot tasks (Yu et al., 2023). To mitigate this, we replace direct prediction with the Radical-Pictographic Dual Matching mechanism designed for OBS. The final model with Dual Matching (+ Dual Matching) significantly improves zero-shot performance. In addition, using only the joint matching branch C_2 (+ joint Matching C_2) from Algorithm 1 performs notably worse, further confirming the necessity of dual matching.

5.5 Qualitative Results

To demonstrate the performance and interpretability of our method, Figure 5 compares our decipherment results with OBSD (Guan et al., 2024b). Although both methods predict correctly on validation

samples, our model generates more semantically aligned Top-3 candidates by leveraging analysis contents and shows stronger robustness in zero-shot cases where OBSD fails on complex or infrequent characters. Moreover, our radical and pictographic analyses provide human-consistent explanations, such as linking *nü* (woman radical) to feminine qualities and *ri* (sun radical) to sunshine, thereby justifying predictions and improving reliability.

Additional qualitative results, undeciphered-character predictions, and cross-dataset visualizations are shown in Appendices A.6 and A.7.

6 Conclusion

In this paper, we address scanned OBS image decipherment as a crucial step in AI-assisted OBS research. We propose an LVLM-based interpretable framework that bridges OBS glyphs and meanings through progressive radical, pictographic, and mutual analysis. To support structured glyph analysis, we construct PD-OBS, a dataset with detailed radical and pictographic annotations and a large-scale dictionary for evidence-based retrieval. Experiments demonstrate strong zero-shot candidate-generation ability for unseen OBS characters, especially substantially improved Top-10 accuracy. These results show the potential of our framework to provide reliable and interpretable candidate references for AI-assisted archaeological and historical studies.

Limitations

Our method has two main limitations. First, although we use a dictionary with broad coverage that includes rare characters as much as possible, candidate retrieval may still fail when an OBS corresponds to an extremely rare modern Chinese character beyond the dictionary coverage. In such cases, the model cannot provide a definitive prediction, but can still offer interpretable radical and pictographic analyses as references for experts. Second, the current framework requires fine-grained annotated data for supervised fine-tuning.

Acknowledgments

This work was supported by the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning.

References

- Shuai Bai, Keqin Chen, Xuejing Liu, Jialin Wang, Wenbin Ge, Sibao Song, Kai Dang, Peng Wang, Shijie Wang, Jun Tang, Humen Zhong, Yuanzhi Zhu, Mingkun Yang, Zhaohai Li, Jianqiang Wan, Pengfei Wang, Wei Ding, Zheren Fu, Yiheng Xu, and 8 others. 2025. *Qwen2.5-vl technical report*. *Preprint*, arXiv:2502.13923.
- Satanjeev Banerjee and Alon Lavie. 2005. Meteor: An automatic metric for mt evaluation with improved correlation with human judgments. In *Proceedings of the acl workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization*, pages 65–72.
- Zijian Chen, Wenjie Hua, Jinhao Li, Lirong Deng, Fan Du, Tingzhu Chen, and Guangtao Zhai. 2025. *Pictobi-20k: Unveiling large multimodal models in visual decipherment for pictographic oracle bone characters*. *Preprint*, arXiv:2509.05773.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. *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)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.
- Xiaolei Diao, Rite Bo, Yanling Xiao, Lida Shi, Zhihan Zhou, Hao Xu, Chuntao Li, Xiongfeng Tang, Massimo Poesio, Cédric M. John, and Daqian Shi. 2025. *Ancient script image recognition and processing: A review*. *Preprint*, arXiv:2506.19208.
- Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. 2021. *An image is worth 16x16 words: Transformers for image recognition at scale*. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net.
- Yoshiyuki Fujikawa, Hengyi Li, Xuebin Yue, Aravinda C V, Amar Prabhu G, and Lin Meng. 2021. *Recognition of oracle bone inscriptions by using two deep learning models*. *Preprint*, arXiv:2105.00777.
- Tommaso Galliana, Tommaso Apicella, Stefano Rosa, Pietro Morerio, Alessio Del Bue, and Lorenzo Natale. 2025. Embodied image captioning: Self-supervised learning agents for spatially coherent image descriptions. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*.
- Ji Gan, Yuyan Chen, Bo Hu, Jiayu Leng, Weiqiang Wang, and Xinbo Gao. 2023. *Characters as graphs: Interpretable handwritten chinese character recognition via pyramid graph transformer*. *Pattern Recognit.*, 137:109317.
- Gemini Team, Google. 2025. Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities. https://storage.googleapis.com/deepmind-media/gemini/gemini_v2_5_report.pdf.
- Haisu Guan, Jinpeng Wan, Yuliang Liu, Pengjie Wang, Kaile Zhang, Zhebin Kuang, Xinyu Wang, Xiang Bai, and Lianwen Jin. 2024a. An open dataset for the evolution of oracle bone characters: Evobc. *arXiv preprint arXiv:2401.12467*.
- Haisu Guan, Huanxin Yang, Xinyu Wang, Shengwei Han, Yongge Liu, Lianwen Jin, Xiang Bai, and Yuliang Liu. 2024b. *Deciphering oracle bone language with diffusion models*. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15554–15567, Bangkok, Thailand. Association for Computational Linguistics.
- Jun Guo, Changhu Wang, Edgar Roman-Rangel, Hongyang Chao, and Yong Rui. 2016. *Building hierarchical representations for oracle character and sketch recognition*. *IEEE Transactions on Image Processing*, 25(1):104–118.
- Ziyi Guo, Zihan Zhou, Bingshuai Liu, Longquan Li, Qingju Jiao, Chenxi Huang, Jianwei Zhang, and Tongguang Ni. 2022. *An improved neural network model based on inception-v3 for oracle bone inscription character recognition*. *Sci. Program.*, 2022.
- Wenhui Han, Xinlin Ren, Hangyu Lin, Yanwei Fu, and Xiangyang Xue. 2020. Self-supervised learning of orc-bert augmentator for recognizing few-shot oracle characters. In *Proceedings of the Asian Conference on Computer Vision (ACCV)*.

- Han Dian. 2004-2025. Han Dian: Online chinese dictionary. <https://www.zdic.net/>. Accessed: 2025.
- Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2021. *Lora: Low-rank adaptation of large language models*. *Preprint*, arXiv:2106.09685.
- Zhikai Hu, Yiu-ming Cheung, Yonggang Zhang, Zhang Peiying, and Tang Pui Ling. 2025. *Component-level segmentation for oracle bone inscription decipherment*. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(27):28116–28124.
- Zhikai Hu, Yiu-ming Cheung, Yonggang Zhang, Peiying Zhang, and Pui-ling Tang. 2024. *Component-level oracle bone inscription retrieval*. In *Proceedings of the 2024 International Conference on Multimedia Retrieval, ICMR '24*, page 647–656, New York, NY, USA. Association for Computing Machinery.
- Shuangping Huang, Haobin Wang, Yongge Liu, Xiaosong Shi, and Lianwen Jin. 2019. *Obc306: A large-scale oracle bone character recognition dataset*. In *2019 International Conference on Document Analysis and Recognition (ICDAR)*, pages 681–688.
- Jerry. CCAMC: A comprehensive online dictionary of ancient and modern Chinese characters. <https://www.ccamc.org>. A non-profit online dictionary covering Unicode CJK characters, including glyph evolution, phonological changes, and classical dictionary analyses. Accessed: May 2025.
- Hanqi Jiang, Yi Pan, Junhao Chen, Zhengliang Liu, Yifan Zhou, Peng Shu, Yiwei Li, Huaqin Zhao, Stephen Mihm, Lewis C Howe, and Tianming Liu. 2024. *Oraclesage: Towards unified visual-linguistic understanding of oracle bone scripts through cross-modal knowledge fusion*. *Preprint*, arXiv:2411.17837.
- Runhua Jiang, Yongge Liu, Boyuan Zhang, Xu Chen, Deng Li, and Yahong Han. 2023. *Oraclepoints: A hybrid neural representation for oracle character*. In *Proceedings of the 31st ACM International Conference on Multimedia, MM '23*, page 7901–7911, New York, NY, USA. Association for Computing Machinery.
- Bang Li, Qianwen Dai, Feng Gao, Weiye Zhu, Qiang Li, and Yongge Liu. 2020. *Hwobc-a handwriting oracle bone character recognition database*. *Journal of Physics: Conference Series*, 1651(1):012050.
- Bo Li, Kaitao Xue, Bin Liu, and Yu-Kun Lai. 2023. *Bbdrm: Image-to-image translation with brownian bridge diffusion models*. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1952–1961.
- Caoshuo Li, Zengmao Ding, Xiaobin Hu, Bang Li, Donghao Luo, AndyPian Wu, Chaoyang Wang, Chengjie Wang, Taisong Jin, Seven Shu, and 1 others. 2025a. *Oraclefusion: Assisting the decipherment of oracle bone script with structurally constrained semantic typography*. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 19893–19902.
- Jing Li, Xueke Chi, Qiufeng Wang, Kaizhu Huang, Da-Han Wang, Yongge Liu, and Cheng-Lin Liu. 2026. *A comprehensive survey of oracle character recognition: Challenges, datasets, methodology, and beyond*. *Pattern Recognition*, 169:111824.
- Jinhao Li, Zijian Chen, Runze Jiang, Tingzhu Chen, Changbo Wang, and Guangtao Zhai. 2025b. *Mitigating long-tail distribution in oracle bone inscriptions: Dataset, model, and benchmark*. *Preprint*, arXiv:2504.09555.
- Chin-Yew Lin. 2004. Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*, pages 74–81.
- Xiaoyu Lin, Shanxiong Chen, Fujia Zhao, and Xiaogang Qiu. 2022. *Radical-based extract and recognition networks for oracle character recognition*. *Int. J. Doc. Anal. Recognit.*, 25(3):219–235.
- Yixin Liu, Pengfei Liu, Dragomir Radev, and Graham Neubig. 2022. *BRIO: Bringing order to abstractive summarization*. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2890–2903, Dublin, Ireland. Association for Computational Linguistics.
- Ilya Loshchilov and Frank Hutter. 2019. *Decoupled weight decay regularization*. *Preprint*, arXiv:1711.05101.
- Yanlong Luo, Yiwen Sun, and Xiaojun Bi. 2023. *Multiple attentional aggregation network for handwritten dongba character recognition*. *Expert Syst. Appl.*, 213(Part):118865.
- Lin Meng. 2017. *Recognition of oracle bone inscriptions by extracting line features on image processing*. pages 606–611.
- OpenAI. 2024. Gpt-4.1 technical overview. <https://openai.com/research/gpt-4-1>. Accessed: 2025-09-18.
- OpenAI. 2025. GPT-5 System Card. <https://cdn.openai.com/gpt-5-system-card.pdf>. Version: Aug 13, 2025.
- Runqi Qiao, Lan Yang, Kaiyue Pang, and Honggang Zhang. 2024. *Making visual sense of oracle bones for you and me*. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12656–12665.
- Florian Schroff, Dmitry Kalenichenko, and James Philbin. 2015. *Facenet: A unified embedding for face recognition and clustering*. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, page 815–823. IEEE.

- Xu Shen. 1963. *Shuowen Jiezi*. Zhonghua Book Company, Beijing, China. Originally compiled c. 100 CE; standard modern edition published by Zhonghua Book Company, 1963.
- Fan Shi, Haiyang Yu, Bin Li, and Xiangyang Xue. 2025. [Cola: Chinese character decomposition with compositional latent components](#). *Preprint*, arXiv:2506.03798.
- Peng Wang, An Yang, Rui Men, Junyang Lin, Shuai Bai, Zhikang Li, Jianxin Ma, Chang Zhou, Jingren Zhou, and Hongxia Yang. 2022. [OFA: Unifying architectures, tasks, and modalities through a simple sequence-to-sequence learning framework](#). In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 23318–23340. PMLR.
- Pengjie Wang, Kaile Zhang, Yuliang Liu, Jinpeng Wan, Haisu Guan, Zhebin Kuang, Xinyu Wang, Lianwen Jin, and Xiang Bai. 2024a. [An open dataset for oracle bone script recognition and decipherment](#). *CoRR*, abs/2401.15365.
- Pengjie Wang, Kaile Zhang, Xinyu Wang, Shengwei Han, Yongge Liu, Lianwen Jin, Xiang Bai, and Yuliang Liu. 2024b. Puzzle pieces picker: Deciphering ancient chinese characters with radical reconstruction. In *Document Analysis and Recognition - IC-DAR 2024*, pages 169–187, Cham. Springer Nature Switzerland.
- Haiyang Yu, Xiaocong Wang, Bin Li, and Xiangyang Xue. 2023. [Chinese text recognition with A pre-trained clip-like model through image-ids aligning](#). In *IEEE/CVF International Conference on Computer Vision, ICCV 2023, Paris, France, October 1-6, 2023*, pages 11909–11918. IEEE.
- Xuebin Yue, Hengyi Li, Yoshiyuki Fujikawa, and Lin Meng. 2022. [Dynamic dataset augmentation for deep learning-based oracle bone inscriptions recognition](#). *J. Comput. Cult. Herit.*, 15(4).
- Biao Zhang and Rico Sennrich. 2019. [Root mean square layer normalization](#). *Preprint*, arXiv:1910.07467.
- Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q Weinberger, and Yoav Artzi. 2020. Bertscore: Evaluating text generation with bert. In *International Conference on Learning Representations*.
- Yushu Zhang and Tingjing Chen. 1958. *Kangxi Zidian*. Zhonghua Book Company, Beijing, China. Originally compiled 1710–1716 CE under imperial commission; modern reprint by Zhonghua Book Company, 1958.
- Yi Zheng, Yi Chen, Xianbo Wang, Donglian Qi, and Yunfeng Yan. 2024. [Ancient chinese character recognition with improved swin-transformer and flexible data enhancement strategies](#). *Sensors*, 24(7):2182.
- Xin-Lun Zhou, Xing-Cheng Hua, and Feng Li. 1995. [A method of jia gu wen recognition based on a two-level classification](#). In *Proceedings of 3rd International Conference on Document Analysis and Recognition*, volume 2, pages 833–836 vol.2.

A Appendix

A.1 Essential Terminology of OBS and Chinese Characters

Oracle Bone Script (OBS) is an early Chinese writing system inscribed mainly on bones and shells. In this paper, an OBS glyph refers to the visual form of an oracle-bone character image, while a modern Chinese character refers to its later standardized written counterpart used as the decipherment target. We use glyph form to describe the shape and structure of a character, rather than its pronunciation or textual context. Because many OBS characters preserve pictorial traces of concrete objects or actions, pictographic analysis refers to explaining what scene, object, or action the glyph visually depicts and what meaning it may imply.

In Chinese, the term "radical" encompasses two distinct concepts: radical components (pianpang) and the indexing radical (bushou). Radical components denote the constituent elements of compound Chinese characters, whereas the indexing radical refers to the single representative component used for character retrieval in dictionary compilation. In our work, "radical" specifically refers to the unique indexing radical of each character. Considering the importance of the indexing radical (bushou) in character retrieval due to their typical semantic function, we adopt the indexing radical (bushou) and its analysis as the key conditions for dictionary matching.

A.2 More Details of PD-OBS Dataset

A.2.1 Annotation of the PD-OBS

We use GPT-4.1 as our commercial LVLM tool. There is a substantial temporal gap between modern Chinese characters and oracle bone script. To mitigate this impact, we introduce an intermediate script layer based on ancient characters and use their philological analyses as auxiliary reference. This design provides GPT-4.1 with richer and more reliable evidence, enabling it to produce more faithful oracle-bone analyses and thereby support the construction of a higher-quality decipherment dataset. We instantiate this intermediate layer with Clerical Script for two reasons: (1) the historical emergence of Clerical Script is broadly contemporaneous with the compilation of the key reference *Shuowen Jiezi* (Shen, 1963), so that its glyph forms are naturally aligned with the dictionary analyses we obtain; and (2) compared with earlier scripts (such as bronze inscriptions), Clerical Script has a

substantially larger and more diverse surviving corpus, offering broader coverage of character types.

Each Radical and Pictographic Analysis annotation derived from GPT analysis and summarization is re-input into GPT alongside modern character forms, ancient character forms, and comprehensive dictionary analyses (including all reference information available from authoritative ancient sources like *Shuowen Jiezi* (Shen, 1963), *Kangxi Dictionary* (Zhang and Chen, 1958), and *Han Dian Dictionary* (Han Dian, 2004-2025)). The following self-checks are then performed: a. Whether the annotation description aligns with the character form; b. Whether the annotation description conflicts with dictionary analyses or contains information not covered by them; c. Whether the annotation summarizes the dictionary analysis content. All problematic annotations are regenerated and re-examined.

In addition, the final output undergoes manual correction to further ensure annotation quality. Specifically, among all annotations modified during the self-checking process, those exhibiting substantial semantic shifts between the original and revised versions are subject to manual review and correction, accounting for approximately 10% of the total samples. Furthermore, for samples where dictionary reference content is limited or entirely absent, we manually construct detailed and faithful annotations through consultation of additional authoritative sources, comprising approximately 0.3% of the total dataset.

A.2.2 Visualization of the PD-OBS Dataset

To provide a more intuitive demonstration of the proposed dataset PD-OBS, we have visualized several data cases in Figure 6. It is worth noting that, for the sake of clarity, we have translated the annotation information into English.

A.2.3 Annotation Participants

The annotations were completed by the authors. No external participants were recruited, and no payment or compensation was involved.

A.2.4 Licenses, Release, and Intended Use

PD-OBS is constructed from multiple existing resources, including HUST-OBC (Wang et al., 2024a), EVOBC (Guan et al., 2024a), Han Dian (Han Dian, 2004-2025), the CJKV Old and Ancient Chinese Character Glyph Database (Jerry), *Shuowen Jiezi* (Shen, 1963), and the *Kangxi Dictionary* (Zhang and Chen, 1958). HUST-OBC and

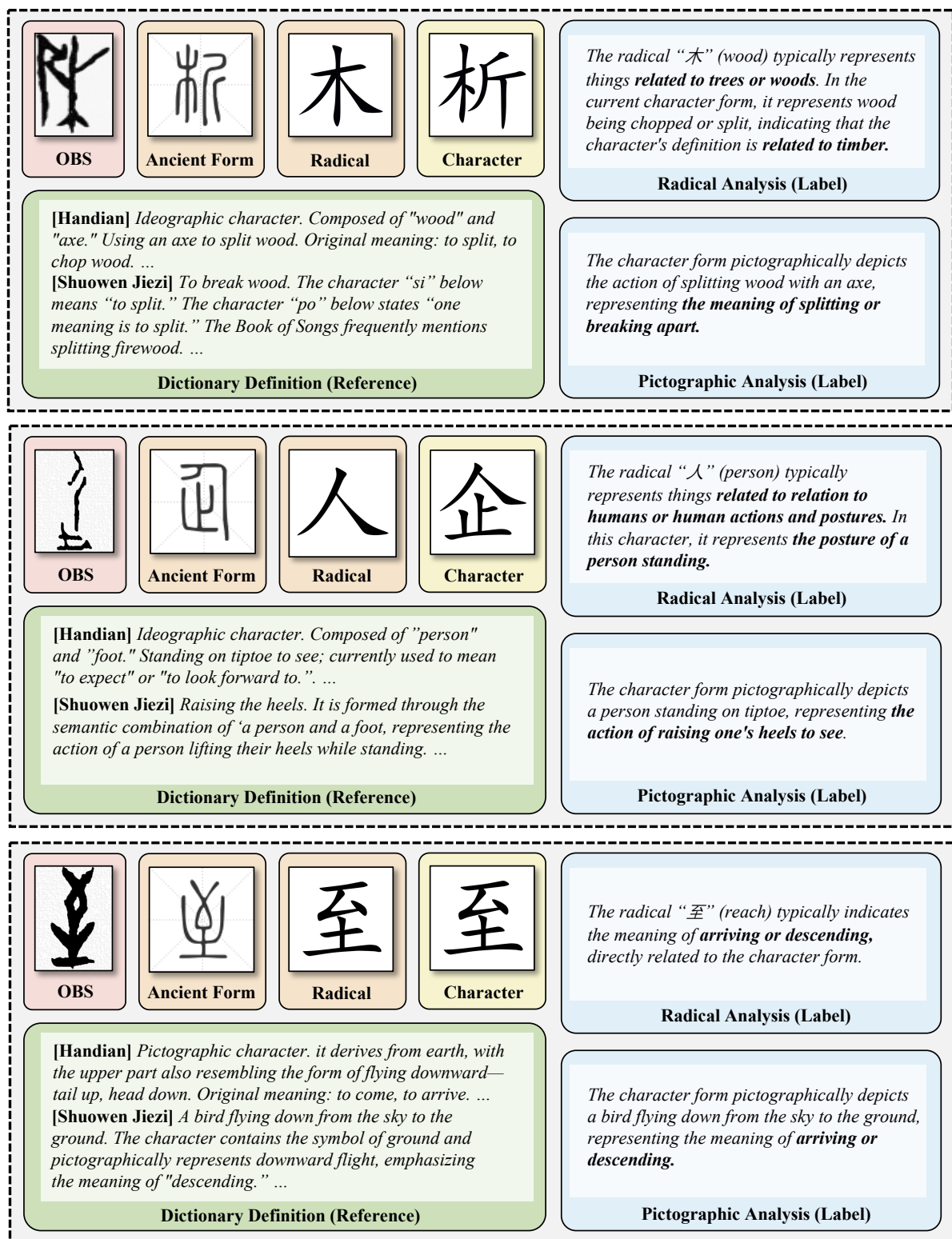


Figure 6: Visualization of the PD-OBS dataset.

EVOBC are publicly released academic datasets and are reused with attribution in accordance with their stated licenses or terms. Han Dian and the CJKV Old and Ancient Chinese Character Glyph Database are used as online reference resources for non-commercial academic research. Shuowen Jiezi and the Kangxi Dictionary are classical Chinese dictionaries whose underlying texts are in the public domain, although modern digitized editions, database structures, or webpage presentations may be subject to their own terms.

Our goal is to make PD-OBS as openly reusable as possible for academic research while respecting the licenses and terms of the source artifacts. For resources that permit redistribution or derivative academic use, we will release the corresponding processed data with proper attribution. For resources whose redistribution terms are unclear or more restrictive, we will provide source references, character identifiers, metadata, and reconstruction scripts where possible, rather than redistributing restricted raw contents. The newly created components of PD-OBS, including our radical labels, pictographic analyses, metadata, data splits, and evaluation files, will be released under the CC BY-NC 4.0 license for non-commercial academic research. Third-party materials included in or referenced by PD-OBS retain their original licenses and attribution requirements. Therefore, PD-OBS is designed as a fully reproducible and openly usable resource for academic research, with redistributable components released directly and source-restricted components made recoverable through references, metadata, and reconstruction scripts.

PD-OBS is intended to support academic research on ancient script analysis and AI-assisted decipherment, and data derived from research-only resources should not be used outside research contexts. Since PD-OBS consists of Chinese character glyphs and philological annotation texts rather than human-subject data, it does not contain personally identifiable information (PII).

A.3 More Implementation Details

A.3.1 Radical Recognition Stage

The Spatial Patch Merger. In this stage, we introduce the Spatial Patch Merger for feature down-sampling. First, the OBS images are uniformly resized to 224×224 , and patch embedding is performed with the default patch size of 14×14 following Qwen2.5-VL, ensuring that the Vision En-

coder outputs 256 tokens representing the visual features of each OBS image. The Spatial Patch Merger rearranges these tokens according to their original spatial positions and then merges tokens using the patch-merger module from Qwen2.5-VL, which consists of an RMSNorm layer (Zhang and Sennrich, 2019) and an MLP. Each merge operation reduces the number of tokens to one fourth while keeping the channel dimension unchanged. After four merge iterations, a unique global visual token is obtained, which is then transformed into a feature vector through an MLP.

Batch Sampling. To facilitate the computation of the triplet loss function $\mathcal{L}_{\text{trip}}$, we designed a specialized batch sampling method. First, OBS samples are grouped by radical label indices. Then, we randomly traverse all radical labels with a step size of half the batch size, sampling two samples per label. We mark sampled samples to prevent duplicate sampling, allowing repeated sampling only when the remaining samples for a radical label are odd in number.

A.3.2 Training Details

Radical LoRA. When training the radical LoRA module ϕ_{rad} , our training data Q&A pairs include all OBS with radical analysis, even though some of them are not strongly correlated with their character semantics. This helps train the model to determine the role of radical information in the overall semantic meaning of characters.

Pictographic LoRA. When training the pictographic LoRA module ϕ_{pic} , we first perform a warm-up phase for one epoch using pictographic analysis Q&A pairs. The model outputs from this warm-up phase, which usually contain incorrect pictographic analyses, are then combined with radical analysis data from PD-OBS and the original Q&A pairs to construct multi-turn dialogue data. In the formal training stage, we mix these multi-turn dialogues with the pictographic analysis Q&A pairs as the training set. This strategy is crucial for equipping the model with the capability for Radical-Pictographic Mutual Analysis.

During the above two training stages, we freeze the parameters of the first 16 layers of the visual encoder and train the last 16 layers of the visual encoder, the patch merger, and the LoRA modules.

In this section, we will provide additional explanations focusing primarily on the Radical Recognition stage and training details. Due to space constraints, these details were not elaborated upon in

the main text.

A.4 Experimental Details and Baseline Discussion

A.4.1 Discussion of Baselines

In this section, we further analyze existing methods, including those listed and unlisted in Table 1 of main text. When reporting the performance of the baseline and our method, we used five different random seeds and reported the average of the results.

Classification-based Methods. As shown in Table 1, the PyGT (Gan et al., 2023) slightly outperforms our approach on the validation set. This can be attributed to PyGT framing the task as a closed-set classification problem, where a fixed number of character classes significantly simplifies the task. Consequently, classification-based methods like PyGT may be considered incapable of achieving true decipherment and have limited potential for deciphering unknown characters; instead, they are better suited for OBS recognition tasks.

Diffusion-based Methods. Although our Top-1 zero-shot accuracy on the HUST-OBC dataset (Wang et al., 2024a) is slightly lower than that of the OBSD method (Guan et al., 2024b), our approach exhibits a distinct advantage in Top-10 accuracy. Notably, in contrast to the instability of results typically observed with diffusion-based methods, our framework offers higher interpretability, enhancing the reliability of the outputs. In addition, our method encompasses key archaeological procedures, including pictographic analysis, radical analysis, and dictionary verification, thereby offering a more professional and interpretable solution.

Non-open-source Methods. As an LVLM-based method, OracleSage (Jiang et al., 2024) reported zero-shot Top-1 and Top-10 accuracies of 20.2% and 40.9%, respectively, on a dataset composed of HUST-OBC and EVOBC. OracleFusion (Li et al., 2025a) proposes structurally constrained semantic typography for oracle bone script. It fuses component/radical layout with semantic cues to produce interpretable typography that supports decipherment. Despite relevance, we have not yet succeeded in fully reproducing these works, as the source code for both methods remains unpublished and certain experimental details are difficult to obtain. This makes it challenging to incorporate them into our comparative methodology.

A.4.2 The Evaluation of Commercial LVLM

To evaluate the performance of commercial LVLMs in OBS decipherment task, we randomly select five samples from the training set as test cases following OracleSage (Jiang et al., 2024), enabling these models to generate predictions through in-context learning. Each sample comprises three elements: an Oracle bone script image, its pictographic analysis, and the deciphered result. Each prediction consists of the ten most confidently predicted modern Chinese characters alongside their corresponding analysis content. To mitigate the impact of data randomness, we conducted five evaluations and calculated the average of the performances as the final results.

A.4.3 Commercial LVLMs with PD-OBS Matching

To examine whether the performance gain mainly comes from dictionary access, we further evaluate a dictionary-access variant of commercial LVLMs as shown in Table 5. Specifically, we first prompt each commercial LVLM to generate radical and pictographic analysis texts for the input OBS image, and then use the generated analysis texts as queries for the same PD-OBS dictionary matching mechanism. This setting gives commercial LVLMs access to the same retrieval dictionary, while replacing our progressively trained glyph-analysis model with their zero/few-shot analysis outputs.

As shown in Table 5, applying the PD-OBS matching mechanism to commercial LVLM outputs does not improve over direct prediction in Table 1. This suggests that the main bottleneck is not dictionary access itself, but whether the generated radical/pictographic analyses are aligned with the PD-OBS annotation space. Unlike general-purpose LVLMs, our model is progressively trained for radical, pictographic, and mutual analysis, making its generated evidence more suitable for reliable dictionary matching.

A.5 Additional Ablation and Robustness Analysis

A.5.1 Hyperparameter Sensitivity

To validate the impact of hyperparameters α and γ , we conducted a sensitivity analysis as shown in the Table 6. We observe that small α makes the triplet constraint insufficiently strong, leading to weaker separation between visually similar radicals, whereas an excessively large α introduces optimization instability and degrades performance.

Method	Validation (Seen-OBS Recognition)		Zero-shot (Unseen-OBS Decipherment)	
	HUST-OBC	EVOBC	HUST-OBC	EVOBC
GPT-4.1 (OpenAI, 2024) + Matching	4.8 / 8.6	4.0 / 6.9	4.5 / 6.1	4.0 / 8.1
Qwen-VL-Max (Bai et al., 2025) + Matching	3.0 / 4.9	3.3 / 4.5	1.5 / 2.1	2.9 / 4.5
Gemini-2.5-Pro + Matching	5.9 / 11.2	5.0 / 9.7	4.8 / 7.9	5.5 / 13.8
GPT-5 (OpenAI, 2025) + Matching	6.4 / 14.5	5.1 / 11.7	6.1 / 9.0	5.8 / 12.8

Table 5: Performance of commercial LVLMs when their generated radical/pictographic analysis texts are used for PD-OBS dictionary matching. Each cell reports Top-1 / Top-10 accuracy (%).

In addition, since γ controls the relative contribution of the triplet loss, α and γ need to be adjusted jointly to ensure balanced optimization and prevent either loss from dominating. The empirical trends in the table support the choice of $\alpha=0.25$ and $\gamma=5$ used in the main experiments.

Table 6: Sensitivity of α (rows) and γ (columns). Radical recognition accuracy (in %) on the HUST-OBC dataset with validation (Valid.) and zero-shot (ZS) settings

$\alpha \backslash \gamma$	1	5	10
0.1	92.0 / 86.5	92.2 / 86.8	92.7 / 87.1
0.25	93.1 / 87.7	93.6 / 88.3	91.4 / 87.9
0.5	91.2 / 86.1	89.4 / 84.5	88.9 / 83.2

A.5.2 Ablation on Dual Matching

To further validate the effectiveness of Radical-Pictographic Dual Matching, we compared it with different matching mechanisms, namely Filtered matching using C_1 and Joint matching using C_2 as defined in Algorithm 1. As shown in Table 7, both mechanisms perform worse than Radical-Pictographic Dual Matching. This is primarily due to potentially inaccurate pictographic analysis in Filtered matching using C_1 , while Joint matching using C_2 may result in more failure cases on samples where radicals are weakly correlated with character meaning.

A.5.3 Robustness to Radical Recognition Errors

To further validate the robustness of the model’s multi-stage learning, we conducted experiments to investigate its performance when radical recognition errors occur, as shown in the Table 8. These results confirm that even when the radical recognition stage is incorrect, the model still succeeds in many cases. Specifically, when radical prediction errors occur, the proposed model still maintains Top-10 accuracy rates of 32.20% and 47.70% on HUST-OBC and EVOBC, respectively.

Table 7: Decipherment accuracy (in %) under different matching mechanisms.

Matching Mechanism	HUST-OBC		EVOBC	
	Valid.	ZS	Valid.	ZS
Filtered Matching using C_1	66.6	8.8	64.6	10.3
Joint Matching using C_2	69.1	14.9	73.1	27.6
Rad&Pic Dual Matching	80.6	16.8	76.3	33.3

Table 8: Decipherment Top-10 accuracy in the zero-shot setting with correct & wrong radical recognition result. The total numbers of zero-shot test samples are 1328 for HUST-OBC and 1616 for EVOBC.

Dataset	Correct / Wrong	Top-10 Acc. (Correct)	Top-10 Acc. (Wrong)
HUST-OBC	1157 / 171	56.90%	32.20%
EVOBC	1461 / 155	65.80%	47.70%

A.5.4 Top-K Parameter Analysis

As shown in Table 9, the zero-shot accuracy of our method significantly improves as we incorporate more characters into the candidate set, achieving a remarkably high accuracy at the Top-50. This result demonstrates that the Radical-Pictographic Dual Matching mechanism can effectively select the appropriate character sets. Although expanding the candidate set may increase the workload, it may enhance the probability of successful decipherment by providing human experts with a more comprehensive reference.

A.5.5 Dictionary Scale Study

In practice, only a limited subset of over 100,000 Chinese characters has achieved widespread circulation and possesses well-defined meanings. Therefore, selecting an appropriate candidate dictionary for oracle bone script decipherment is crucial. Under the PD-OBS dataset setting, the matching dictionary comprises 47,157 Chinese characters documented in the Kangxi Dictionary (a Chinese dictio-

Table 9: Decipherment accuracy (in %) under different Top-k settings. The Valid. and ZS indicate validation and zero-shot settings.

Top-k	HUST-OBC		EVOBC	
	Valid.	ZS	Valid.	ZS
Top-1	80.6	16.8	76.3	33.3
Top-5	86.0	39.3	79.8	56.0
Top-10	87.8	53.7	81.7	64.1
Top-50	92.1	74.2	88.0	80.2
Top-100	94.4	82.5	91.2	89.7

Table 10: Decipherment Top-10 accuracy (in %) under different dictionary scales. The Valid. and ZS indicate validation and zero-shot settings.

Dict. Scale	HUST-OBC		EVOBC	
	Valid.	ZS	Valid.	ZS
7000	59.6	31.9	65.7	48.2
10000	73.7	39.3	79.5	59.8
20902	86.5	51.8	83.9	66.4
27928	88.3	54.1	82.2	64.5
47157	87.8	53.7	81.7	64.1

nary).

We constructed four additional subset dictionaries: 7,000 commonly used Chinese characters, 10,000 commonly used characters including known oracle bone script decipherment results, 20,902 Unicode-supported characters, and 27,928 characters encompassing known oracle bone script decipherment results and Unicode Extension characters. We evaluated the Top-10 decipherment accuracy of our method under different dictionary scales.

As demonstrated in Table 10, larger candidate dictionaries do not necessarily yield superior results; instead, candidate dictionaries with scales ranging from 20,000 to 30,000 characters prove more suitable. This indicates a trade-off relationship between decipherment accuracy and potential recall rate. Despite a slight performance loss, we adopted the Kangxi Dictionary as the PD-OBS dictionary and reported the experimental results based on it in the main text, owing to its reliability and authority.

A.6 Qualitative Analysis

A.6.1 Mutual Analysis and Dual Matching

Figure 7 visualizes the detailed process of Radical-Pictographic Mutual Analysis and Radical-Pictographic Dual Matching. The left part presents the Radical-Pictographic Mutual Analysis content:

Table 11: Decipherment performance Top-1 / Top-10 (in%) on cross-dataset benchmarks.

Method	OBI-125 (rubbing)	Oracle-50K (Valid.)	Oracle-50K (ZS)
OBSD	43.1 / 49.7	54.8 / 62.4	20.5 / 25.6
Ours	46.2 / 60.3	55.3 / 72.5	19.7 / 34.8

Radical Analysis, Pictographic Analysis, and the Radical-Informed Pictographic Analysis. The right part displays the Top-3 candidate characters from Filtered Matching C_1 and Joint Matching C_2 . The final Top-k matching results are obtained by aggregating and re-ranking the candidate characters from C_1 and C_2 according to their respective BERT-Score value (Zhang et al., 2020). In the cases **I** and **II**, mutual analysis successfully rectifies errors originating from pictographic analysis, demonstrating its effectiveness and rationality. In the case **III**, the final deciphering result remained unaffected despite the introduction of unreasonable content during the radical analysis phase. This resilience comes from the error tolerance achieved through the Filtered Matching C_1 , which is enabled only by pictographic analysis. In Case **IV**, even when radical recognition fails, the model analyzes the role of radicals in character and utilizes pictographic and mutual analysis to mitigate the impact of erroneous information, thereby obtaining the accuracy of decipherment results.

A.6.2 Interpretable Content

Figure 8 displays the interpretable outputs from three LVLM-based methods, including our method. We employed BERT-Score (Zhang et al., 2020) to calculate the similarity between model outputs and ground truth annotations, evaluating the reliability of these interpretable contents as a qualitative supplement to Table 2. In Case **I**, where all models provided correct decipherment results, our method’s analysis demonstrated the highest similarity with both PD-OBS annotations and authoritative Chinese dictionary definitions (Dictionary Definition). In Cases **II**, **III**, and **IV**, our model demonstrates more precise interpretability compared to models that predict correctly, and more reasonable explanations and deciphering compared to models that predict incorrectly.

A.6.3 More visualization results

Figure 10 visualizes additional analysis and decipherment results of our method, with the OBSD

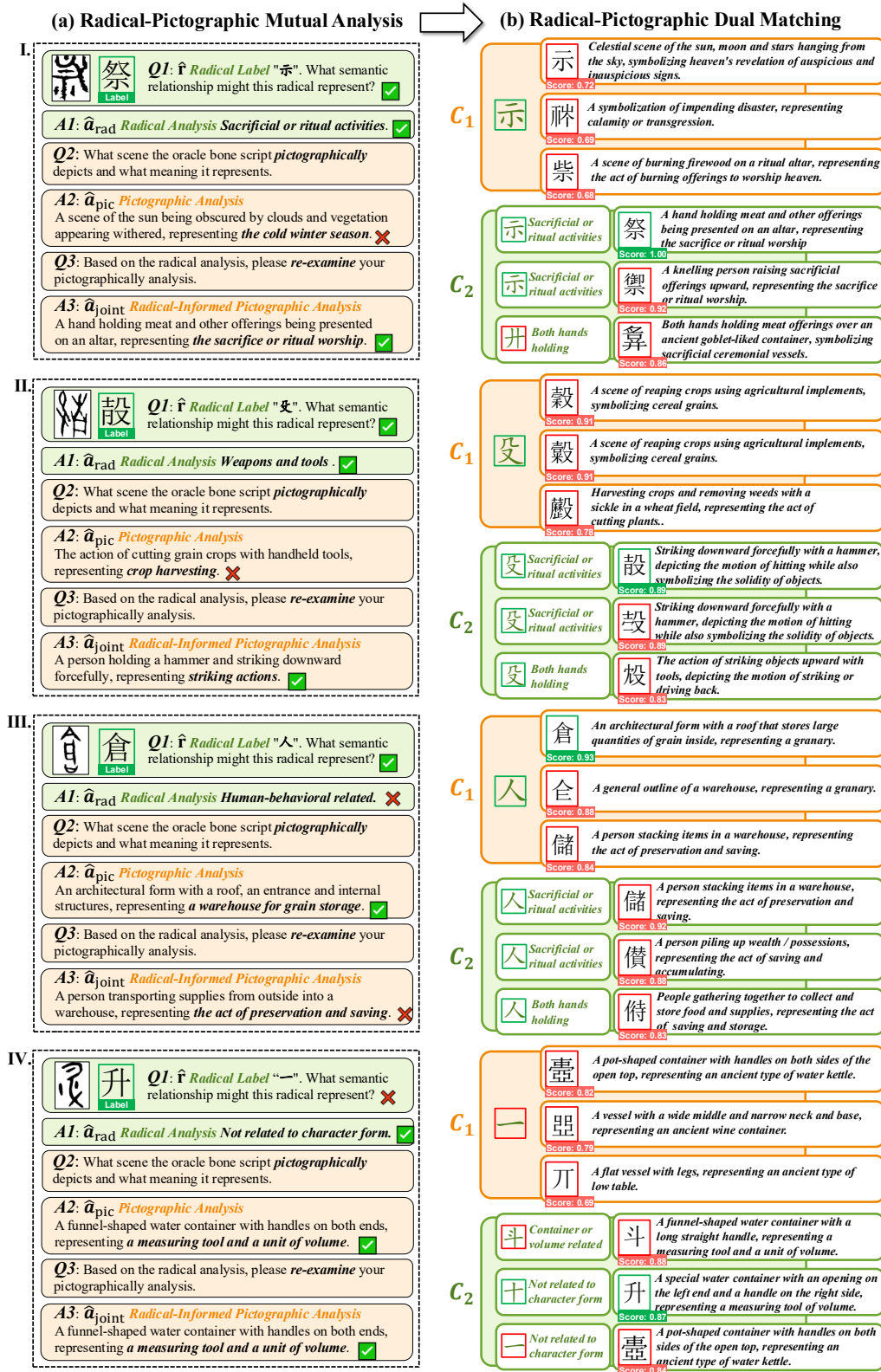


Figure 7: Visualization of Radical-Pictographic Mutual Analysis and candidate sets C_1 & C_2 in Dual Matching. Green rectangles and checkmarks indicate correct contents and results.



Figure 8: Visualization of interpretable outputs from LVLm-based methods. Green rectangles and texts indicate correct contents and results. The two BERT-Score values correspond to: the similarity score between the model outputs and the PD-OBS label / the similarity score between the model outputs and the authoritative dictionary definition.

method used for comparison. The visualization results further demonstrate the effectiveness and robustness of the proposed method on multiple validation set samples and zero-shot samples. We also demonstrated a failed zero-shot case due to an error in pictographic analysis.

A.6.4 Results on Undeciphered Characters

As shown in Figure 11, we represent decipherment outputs for previously undeciphered characters. For each character, we show the Top-4 decipherment results and the related analysis contents. We will subsequently publish all prediction results of the model for undeciphered characters, and open-source both the model and dataset to support research in this field.

A.7 Extended Generalization Results

A.7.1 Cross-dataset Results

To validate the cross-dataset generalization of the proposed model, we also conducted experiments on the Oracle-50K dataset (Han et al., 2020) as shown in Table 11. The proposed method achieved highly

competitive Top-1 accuracy and markedly superior Top-10 accuracy. These experimental results collectively validate the effectiveness and generalization of the proposed approach.

A.7.2 Rubbing-form Results

To validate the cross-form generalization capability and practical application value of the proposed model, we further conducted verification on the Rubbing form OBI-125 dataset (Yue et al., 2022) as shown in Table 11. Experiments demonstrate that even without training on rubbing data, the proposed model maintains competitive Top-1 accuracy and significantly outperforms method OBSD in Top-10 accuracy when confronted with rubbing data containing substantial noise. In addition, the visualization results are shown in Figure 9 to demonstrate the deciphering process of our method for rubbing data.

A.8 The use of large language models

Large Language Models (LLMs) were only used to assist with grammar checking of the manuscript. In

OBS Image	GT	Top-3 Result (OBS)	Top-3 Results (Ours)	Radical Analysis (Ours)	Pictographic Character Form Analysis (Ours)
	貞	貞 [貞] [頁] [貞]	[貞] [貞] [禎]	Radical “貞” is related to currency, transactions, or shellfish creatures; in the character form it symbolizes shells	Using shells for divination activities, representing the meaning of divination and observation
	未	未 [未] [未] [未]	[未] [杕] [木]	Radical “木” is related to trees or plants, in the current character form it symbolizes a big tree	A tree with flourishing branches and foliage, representing the state of a tall, vigorously growing tree.
	雨	𩇔 [局] [易] [居]	[雨] [宋] [雨]	Radical “雨” indicates rain or related weather phenomena, directly relevant to the character form.	A scene of rainwater falling from the sky, representing rainfall
	王	𣎵 [𣎵] [沐] [似]	[立] [𣎵] [𣎵]	Radical “立” represents a standing posture or an upright state, directly relevant to the character form.	A person standing on the ground with feet apart, expressing the meaning of standing upright or being established.

Figure 9: Visualization of extended decipherment results for rubbing form data from OBI-125 dataset.

some cases, minor wording suggestions from the LLM were adopted to improve clarity and readability. No part of the research ideation, experimental design, data analysis, or results interpretation relied on LLMs.

Data Type	OBS Image	GT	Top-3 Result (OBS)	Top-3 Results (Ours)	Radical Analysis (Ours)	Pictographic Character Form Analysis (Ours)
# validation		監	監	[監] [鑒] [鑑]	Radical "皿" is related to basin-type vessels, in the current character form it appears as the image of a water basin	A person kneeling beside a basin, observing their own reflection in the water, representing the meaning of observation and examination
		俛	俛	[俛] [規] [俛]	Radical "人" is related to human behaviors, in the current character form it symbolizes a person	A figure with focused gaze upon an object, representing the meaning of observation and metaphorical reasoning
		兮	兮	[兮] [兌] [允]	Radical "八" is related to paths or flow; in the character form it symbolizes air current	A stream of air flowing downward from above, representing the meaning of the transmission of breath or sound
		刈	刈	[刈] [刊] [刈]	Radical "丩" is related to blade cutting, in the current character form it represents a farm sickle	A farmer harvesting wheat with a sickle, representing the meaning of crop harvesting
		鑄	鑄	[鑄] [厝] [厝]	Radical "金" is related to metallic objects, in the current character form it symbolizes metallurgy engineering	A person holding an ancient tripod vessel (li) pouring molten metal into a mold, representing the meaning of the process of casting and metallurgy
# zero-shot		汜	汜	[汜] [已] [汜]	Radical "水" is related to water flow, in the current character form it symbolizes branches of water flow	Rivers flowing into a lake water system, indicating river inlet to the lake
		昔	昔	[昔] [巷] [肯]	Radical "日" indicates the sun, in the current character form it represents sunshine	A scene of meat and food under the scorching sun, representing the meaning of dried meat
		盜	盜	[盜] [盜] [盜]	Radical "皿" is related to containers, in the current character form it represents a water jar	A hand holding a brush to clean vessels, expressing the meaning of emptiness within containers
		割	割	[割] [割] [割]	Radical "糸" is related to silk threads, in the current character form it represents a bundle of ropes	A scene of cutting hemp rope with a knife, indicating a clean cut, symbolizing the meaning of termination and severance
		剛	剛	[剛] [剛] [剛]	Radical "丩" is related to blade cutting, in the current character form it represents a sharp knife	A scene of cutting through the interwoven network of rope fibers with a blade, representing the meaning of the qualities of rigidity and resilience

Figure 10: Visualization of extended decipherment results in validation and zero-shot settings.

Oracle Bone Character	OBS Result	Results (Ours)	Radical Analysis (Ours)	Pictographic Character Form Analysis (Ours)
	汜	[漁] [汜]	Radical "水" is related to water flow, in the current character form it represents a flowing stream	A scene of fishing with nets in a stream, representing the meaning of fishing
	舂	[舂] [舂]	Radical "臼" indicates tools used for pounding rice, directly related to the character's meaning	The process of holding a pestle and mortar while pounding rice, expressing the meaning of husking grain
	汜	[洗] [汜]	Radical "水" is related to water flow, in the current character form it symbolizes a water pool	A foot being washed in water, representing foot washing
	眊	[目] [眊]	Radical "目" is related to eyes, directly related to the character's meaning	Eyes facing each other, expressing the meaning of mutual gaze or eye contact
	𤇀	[災] [𤇀]	Radical "火" is related to fire and heat, in the current character form it represents burning flames	A scene of fire spreading through trees, expressing the meaning of forest fire or catastrophic blaze
	𧈧	[蠱] [𧈧]	Radical "虫" represents insect creatures, directly related to the character form	An image of a long-tailed venomous insect crawling, indicating a type of insect creature
	𧈧	[𧈧] [𧈧]	Radical "虎" is related to tigers and beasts, in the current character form it symbolizes a pouncing tiger	A scene of a fierce tiger pouncing from the grass, indicating hidden threats and imminent danger
	𧈧	[弓] [𧈧]	Radical "弓" is related to bows and arrows, in the current character form it symbolizes hunting with a bow	A figure grasping a bow with both hands, representing a ceremonial act before hunting
	𧈧	[𧈧] [𧈧]	Radical "比" is related to comparison or parallelism, unrelated to the current character's pictographic meaning	A fierce beast with bared fangs and extended claws, representing the meaning of ferocity and savagery
	龜	[龜] [龜]	Radical "龜" is related to turtles and reptilian animals, directly related to the character's meaning	Turtle shell patterns and turtle form, indicating a type of large turtle creature
	𧈧	[𧈧] [𧈧]	Radical "丩" is related to blade cutting, in the current character form it represents execution tools	A criminal's legs and feet being severed, representing the ancient cruel punishment of foot amputation
	𧈧	[𧈧] [𧈧]	Radical "耳" is related to sound and hearing, in the current character form it represents the image of an ear	A suspended bell chime, representing the resonating sound of bells
	𧈧	[𧈧] [𧈧]	Radical "火" is related to burning scenes, in the current character form it symbolizes fiercely burning flames	A scene of flames burning fiercely, representing vigorous fire
	𧈧	[𧈧] [𧈧]	Radical "田" is related to fields/farmland, unrelated to the current character's pictographic meaning	A complete human figure with both hands raised high, representing a bizarre form that inspires fear
	𧈧	[𧈧] [𧈧]	Radical "女" is related to feminine qualities, in the current character form it manifests as a kneeling woman	A woman arranging her appearance beside a bed, expressing the meaning of grooming and dressing up

Figure 11: Visualization of extended decipherment results of undeciphered characters.