Multi-Granularity Archaeological Dating of Chinese Bronze Dings Based on a Knowledge-Guided Relation Graph
Supplementary Materials

Rixin Zhou[1] Jiafu Wei[1] Qian Zhang[3] Ruihua Qi[3] Xi Yang[1,2,*] Chuntao Li[3,*]
[1]School of Artificial Intelligence, Jilin University
[2]Engineering Research Center of Knowledge-Driven Human-Machine Intelligence, MoE, China
[3]School of Archaeology, Jilin University

1. Overview

The supplementary materials we provide are organized as follows:

- Supplementary Materials
  - * Code and Data
  - * Typical Data Examples
  - * Annotation Lists
  - * Supplementary Materials.pdf (this file)

In the following, we organize each section as follows:

- In Section 2, we introduce more archaeological knowledge that guides our network design.
- In Section 3, we provide more details about data collection and annotation.
- In Section 4, we describe the implementation details of our network.
- In Section 5, we describe the detailed computation of our loss function.
- In Section 6, we perform additional ablation experiments on the hyper-parameter settings in the network.
- In Section 7, we compare the experimental results of our proposed loss function with two other commonly used loss functions.
- In Section 8, we provide more comparison results between SOTA methods and our network.

2. Archaeological Knowledge

Meaning. The Shang, Western Zhou, and Warring States were the most prosperous times for bronze wares in Chinese history. Therefore, the study of bronzes is of great significance in archaeology, history, linguistics, and other aspects. And bronze dating is an important part of bronze research. Only on the basis of dating can bronze wares become effective historical materials.

Roadmap. Bronze wares have been unearthed in the Han dynasty, but the number is too small to study in depth. In the Song dynasty, epigraphy became popular, and the study of bronzes also began. Until the Qing dynasty, people paid attention to the age of the bronzes, but the research was still scattered and non-systematic. The real start of scientific research on bronze dating was at the beginning of the 20th century. The "Standard Utensil Dating Method" was proposed, that is, to first determine the era of a certain utensil according to the content of the inscription, use it as a standard utensil, and then infer the era of the utensils with related or similar shapes and characteristics [3]. The proposal of this method has laid a scientific foundation for the dating of bronzes, and it has also been widely accepted by scholars. Then, the shape and characteristics can be summarized in a certain type of bronze ware, combined with the unearthed pottery, and correspond to the underlying relationship in the archaeological unit, the eras of new bronze wares can be determined.

Methodology. The development of bronze wares has a general process, and its shape and characteristics have a law of evolution.

In terms of shape, it was mainly influenced by social productivity, handicraft technology, and social thoughts at that time. For example, a general development process of the bronze ding foot is cone foot, column foot, and hoof foot. The bronze ding with cone feet is a very typical feature of the early Shang dynasty. The column foot was popular from the late Shang dynasty to the mid and late Western Zhou dynasty, while the hoof foot first appeared in the mid Western Zhou dynasty, and became the main foot type during the Spring and Autumn and Warring States.

In terms of characteristics, they also have certain era and region features. For example, animal face patterns were most popular in the late Shang dynasty and early Western Zhou dynasty; the phoenix and bird pattern was more com-

*Corresponding authors

mon in the mid Western Zhou dynasty; the swallowtail pattern was only seen in the Xingan Oceania region in the late Shang dynasty.

3. Data Collection and Annotation

We have collected all the available bronze ding data as much as possible.

3.1. Statistics

Table 1 shows the number of images in each era in our dataset and the number of images split into training, validation, and test sets. The imbalance is because bronze ware was more popular as utensils and ritual vessels during the Shang and Western Zhou dynasties. However, during the Spring and Autumn and Warring States dynasties, bronze wares were mostly used as utensils, and their role as ritual vessels was greatly weakened. Funeral bronzes are no longer popular as symbols of status in burials. In addition, with the development of smelting technology and handicraft technology, utensils made of iron, lacquered wood, and other materials gradually replaced bronze. Therefore, the number of bronze dings in the Spring and Autumn and Warring States dynasties was much smaller than that in the Shang and Western Zhou dynasties.

Furthermore, Four characteristics appear in 1000+ images. Standing ear, cylindrical foot, and hoof-shaped foot are very common as the key components, and the string pattern is relatively simple, so there are many.

3.2. Data Annotation Tool

In addition to labeling the era and shape for each image, we modify LabelImg [4] to annotate characteristics, as shown in Figure 1. Before starting to label, it is necessary to pre-define the categories according to the type of characteristics in our dataset, with 96 types in total. We provide categories lists for era, shape, and characteristics in the annotation list files.

3.3. Calculation of Information Gain

To quantify the information provided by the additional annotations, we calculate the information gain of the shape and characteristic annotations for era judgement, as shown in Equation (1).

$$g(D, A) = H(D) - H(D \mid A) \tag{1}$$

where $\mathrm{H(D)} = -\sum_{\mathrm{i}=1}^{\mathrm{C}} \mathrm{e_i} \log(\mathrm{e_i})$ is the entropy of dataset D, $\mathrm{H(D \mid A)} = \sum_{\mathrm{j}=1}^{\mathrm{K}} \mathrm{e_j H(D \mid A = A_j)}$ is the conditional entropy of annotation A on dataset D, and C and K are the number of categories in D and categories of the additional attribute annotation A. We assume that the data in D corresponding to each category is uniformly distributed, thus the entropy of dataset D is $\mathrm{H(D)} = \log(C)$. The increased information gain indicates that the annotation provides more significant.

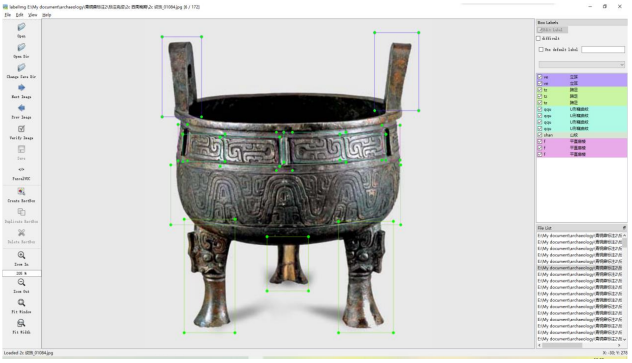

Figure 1. Annotation tool we use to label characteristics.

4. Network Details

Table 2 shows the structure of each head in our network. The size of the input image is 400×400×3. The dynasty head in MGM outputs a 4-dim vector corresponding to 4 dynasties and then applies $sigmoid$ projection to form the dynasty nodes of the AKG. The period head outputs an 11-dim vector corresponding to 11 periods and then applies $sigmoid$ and $softmax$ projection to form the period nodes of the AKG and computes the cross-entropy loss to enhance the exclusive relation between the period nodes, respectively. In KEM, the shape head outputs a 29-dim vector corresponding to 29 shapes and then applies $sigmoid$ and $softmax$ projection to form the shape nodes of the AKG and compute the focal loss. The characteristic head outputs a 96-dim vector corresponding to 96 characteristics and then applies $sigmoid$ projection to form the characteristic nodes of the AKG and computes the multi-label focal loss.

We conducted experiments and show that adding softmax projection to the dynasty head and constraining it with cross-entropy loss will lead to decreases in the dating performance (2.04% OA on dynasty and 2.91% OA on period). Therefore, we did not implement it in our proposed network.

5. Focal-type Probabilistic Classification Loss

Follow the probabilistic classification loss in [1], we describe the details of our extended calculation process. Given an input image $\mathbf{x}$, the unnormalized era joint probability of all era nodes concerning the era label assignment $\mathbf{y_e}$ is computed as:

$$\tilde{P}_e(\mathbf{y_e}|\mathbf{x}) = \prod_{i=1}^{n} \phi_{e_i}(\bar{x}_{e_i}, y_{e_i}) \prod_{i,j \in \{1,\ldots,n\}} \psi_{e_{i,j}}(y_{e_i}, y_{e_j}) \tag{2}$$

Table 1. Statistics for our dataset. We counted the specific number of samples in each era, as well as the number of samples in the training, validation, and test set.

Era	Shang		Western Zhou			Spring and Autumn			Warring States			Total
	Early	Late	Early	Mid	Late	Early	Mid	Late	Early	Mid	Late	
Number	93	1012	852	428	279	302	193	217	87	78	149	3690
Split	37/9/47	404/101/507	340/85/427	171/42/215	111/27/141	120/30/152	77/19/97	86/21/110	34/8/45	31/7/40	59/14/76	1470/363/1857

Table 2. The structure of each head in our network. The size of the input image is 400×400×3.

Part	Output Size	Layer Information
backbone	13 × 13 × 2048	ResNet50
conv1	13 × 13 × 1024	CONV-(N1024, K1, S1, P0), BN, ReLU
conv2	13 × 13 × 2048	CONV-(N2048, K3, S1, P1), BN, ReLU
	1 × 1 × 2048	average poolling
fc1	1024	BN, MLP-(N1024)
fc2	512	BN, ReLU, MLP-(N512)
fc3	4	ReLU, MLP-(N4)
projection	4	Sigmoid

(a) Network architecture of dynasty head in MGM.

Part	Output Size	Layer Information
backbone	13 × 13 × 2048	ResNet50
conv1	13 × 13 × 1024	CONV-(N1024, K1, S1, P0), BN, ReLU
conv2	13 × 13 × 2048	CONV-(N2048, K3, S1, P1), BN, ReLU
	1 × 1 × 2048	average poolling
fc1	1024	BN, MLP-(N1024)
fc2	512	BN, ReLU, MLP-(N512)
fc3	11	ReLU, MLP-(N11)
projection	11	Sigmoid and Softmax

(b) Network architecture of period head in MGM.

Part	Output Size	Layer Information
backbone	13 × 13 × 2048	ResNet50
conv1	13 × 13 × 1024	CONV-(N1024, K1, S1, P0), BN, ReLU
conv2	13 × 13 × 2048	CONV-(N2048, K3, S1, P1), BN, ReLU
	1 × 1 × 2048	average poolling
fc1	1024	BN, MLP-(N1024)
fc2	512	BN, ReLU, MLP-(N512)
fc3	29	ReLU, MLP-(N29)
projection	29	Sigmoid and Softmax

(c) Network architecture of shape head in KEM.

Part	Output Size	Layer Information
backbone	13 × 13 × 2048	ResNet50
conv1	13 × 13 × 1024	CONV-(N1024, K1, S1, P0), BN, ReLU
conv2	13 × 13 × 2048	CONV-(N2048, K3, S1, P1), BN, ReLU
	1 × 1 × 2048	average poolling
fc1	1024	BN, MLP-(N1024)
fc2	512	BN, ReLU, MLP-(N512)
fc3	96	ReLU, MLP-(N96)
projection	96	Sigmoid

(d) Network architecture of characteristic head in KEM.

And, the unnormalized era-shape joint probability concerning the era-shape label assignment $\mathbf{y_{es}}$ and the unnormalized era-characteristic joint probability concerning the era-characteristic label assignment $\mathbf{y_{ec}}$ are computed as:

$$\tilde{P}_{es}(\mathbf{y_{es}}|\mathbf{x})=\prod_{i=1}^{n+m}\phi_{es_i}\left(\bar{x}_{es_i},y_{es_i}\right)\prod_{i,j\in\{1,\ldots,n+m\}}\psi_{es_{i,j}}\left(y_{es_i},y_{es_j}\right) \tag{3}$$

$$\tilde{P}_{ec}(\mathbf{y_{ec}}|\mathbf{x})=\prod_{i=1}^{n+k}\phi_{ec_i}\left(\bar{x}_{ec_i},y_{ec_i}\right)\prod_{i,j\in\{1,\ldots,n+k\}}\psi_{ec_{i,j}}\left(y_{ec_i},y_{ec_j}\right) \tag{4}$$

where $\bar{x}_{e_i}$, $\bar{x}_{es_i}$, and $\bar{x}_{ec_i}$ are the sigmoid output of the corresponding i-th era nodes, era-shape nodes, and era-characteristic nodes, respectively; n, m, and k denote the number of nodes for the eras, shapes, and characteristics. Then, $\phi_{e_i}\left(\bar{x}_{e_i},y_{e_i}\right)=e^{\bar{x}_{e_i}\left[y_{e_i}=1\right]}$, and $\psi_{e_{i,j}}\left(y_{e_i},y_{e_j}\right)$ is the constraint defined in the relation graph between any two era labels in $\mathbf{y_e}$:

$$\psi_{e_{i,j}}(y_{e_i},y_{e_j})=\begin{cases}0, & \text{if violates era nodes assignment}\\ 1, & \text{otherwise}\end{cases} \tag{5}$$

, $\phi_{es_i}\left(\bar{x}_{es_i},y_{es_i}\right)$, $\phi_{ec_i}\left(\bar{x}_{ec_i},y_{ec_i}\right)$, $\psi_{es_{i,j}}\left(y_{es_i},y_{es_j}\right)$, and $\psi_{ec_{i,j}}\left(y_{ec_i},y_{ec_j}\right)$ are calculated in the same way.

The era joint probability is then normalized by $\mathrm{Pr_e}(\mathbf{y_e}\mid\mathbf{x})=\frac{\tilde{P}_e(\mathbf{y_e}|\mathbf{x})}{Z_e(\mathbf{x})}$, where $Z_e(x)$ is the era partition function that sums over all legal era assignments $\overline{\mathbf{y}}_e\in S_{G_e}$:

$$Z_e(\mathbf{x})=\sum_{\overline{\mathbf{y}}_e\in\{0,1\}^n}\prod_{i=1}^{n}\phi_{e_i}(\bar{x}_{e_i},\bar{y}_{e_i})\prod_{i,j\in\{1,\ldots,n\}}\psi_{e_{i,j}}\left(\bar{y}_{e_i},\bar{y}_{e_j}\right) \tag{6}$$

And the normalized era-shape partition function and era-characteristic partition function are computed as:

$$Z_{es}(\mathbf{x})=\sum_{\overline{\mathbf{y}}_{es}\in\{0,1\}^{n+m}}\prod_{i=1}^{n+m}\phi_{es_i}(\bar{x}_{es_i},\bar{y}_{es_i})\prod_{i,j\in\{1,\ldots,n+m\}}\psi_{es_{i,j}}\left(\bar{y}_{es_i},\bar{y}_{es_j}\right) \tag{7}$$

$$Z_{ec}(\mathbf{x})=\sum_{\overline{\mathbf{y}}_{ec}\in\{0,1\}^{n+k}}\prod_{i=1}^{n+k}\phi_{ec_i}(\bar{x}_{ec_i},\bar{y}_{ec_i})\prod_{i,j\in\{1,\ldots,n+k\}}\psi_{ec_{i,j}}\left(\bar{y}_{ec_i},\bar{y}_{ec_j}\right) \tag{8}$$

If input image $\mathbf{x}$ has the i-th era label, we obtain the era marginal probability $\mathrm{Pr_e}(y_{e_i}=1\mid\mathbf{x})$ of era label i by summing over all legal ear assignments $\overline{\mathbf{y}}_e$ that include $\overline{y}_{e_i}$ = 1.

Table 3. Evaluation of the recall performance of all 4 coarse-grained dynasties and 11 fine-grained periods for each independent era classification.

Method		w/ Attributes	Shang		Western Zhou			Spring and Autumn			Warring States		
			Early	Late	Early	Mid	Late	Early	Mid	Late	Early	Mid	Late
Single-Granularity	ConvNeXt		**89.36**	79.39	74.24	86.98	80.85	75.66	61.86	*75.45*	36.67	26.25	*83.55*
	Part-based R-CNN	Bounding box	53.19	68.64	**86.42**	59.53	74.47	76.32	55.67	*75.45*	11.11	27.50	60.00
	MCL		74.47	76.23	72.48	79.77	68.44	72.04	56.19	61.82	18.89	36.25	69.74
	CrossX		59.57	*84.12*	62.53	84.42	65.25	*81.25*	44.85	68.64	18.89	25.00	70.39
	BCNN		61.70	77.91	63.00	80.93	71.63	75.66	51.55	66.36	0.00	0.00	67.11
	NTS-Net		78.72	71.99	77.52	79.53	*83.69*	75.66	54.64	68.18	28.89	*42.50*	81.33
	A^3M	✓	82.98	82.25	73.07	86.51	77.30	71.71	54.64	74.55	26.67	37.50	80.26
	SPS		*87.23*	80.67	*77.99*	86.51	80.14	79.61	**64.95**	66.36	35.56	30.00	81.33
	P2PNet		**89.36**	**86.59**	67.21	**91.63**	80.85	78.95	61.86	70.91	35.56	**47.50**	**84.00**
Multi-Granularity	YourFL		80.87			*87.10*			85.79			*85.53*	
			89.36	77.51	75.88	79.53	75.89	75.66	54.64	70.91	31.11	37.50	80.60
	C-HMCNN		*81.77*			84.93			88.30			82.50	
			82.98	79.29	74.47	80.00	77.30	75.00	57.73	73.64	*40.00*	**47.50**	73.33
	HRN		**86.64**			82.38			*90.53*			82.89	
			87.23	83.43	70.02	83.26	80.14	**82.24**	62.89	67.27	**53.33**	32.50	73.13
	Ours	✓	**86.64**			**89.53**			**90.81**			**87.58**	
			87.23	82.84	77.75	*88.84*	**86.52**	78.95	60.82	**76.36**	37.78	*42.50*	78.95

Given m training samples, $\mathcal{D} = \{x^l, y_e^l, y_{es}^l, y_{ec}^l, g_e^l, g_{es}^l, g_{ec}^l\}$, $l = 1, \ldots, m$, where y_e^l, y_{es}^l and y_{ec}^l are the ground-truth label vector of the era, era-shape combination, and era-characteristic combination, respectively. And $g_e^l \in \{1, \ldots, n\}$, $g_{es}^l \in \{1, \ldots, n+m\}$, $g_{ec}^l \in \{1, \ldots, n+k\}$ are the indices of the observed era, era-shape combination, and era-characteristic combination labels, respectively. Subsequently, the era probabilistic classification loss $\mathcal{L}_{\text{e}}(\mathcal{D})$ is defined as follows:

$$-\frac{1}{m}\sum_{l=1}^{m}\ln(\Pr_{\text{e}}(y_{e_{g_e^l}}^l = 1 \mid \mathbf{x}^l)) \tag{9}$$

Then, because of the different importance of attributes, we define the focal-type era-shape probabilistic classification loss $\mathcal{L}_{\text{es}}(\mathcal{D})$ and the era-shape-characteristic probabilistic classification loss $\mathcal{L}_{\text{esc}}(\mathcal{D})$ as follows:

$$-\frac{1}{m}\sum_{l=1}^{m}\Big((1-\Pr_{\text{e}}(y_{e_{g_e^l}}^l=1|\mathbf{x}^l))^{\alpha_1}\ln(\Pr_{\text{es}}(y_{es_{g_{es}^l}}^l=1|\mathbf{x}^l))\Big) \tag{10}$$

$$-\frac{1}{m}\sum_{l=1}^{m}\Big((1-\Pr_{\text{es}}(y_{es_{g_{es}^l}}^l=1|\mathbf{x}^l))^{\alpha_2}\ln(\Pr_{\text{ec}}(y_{ec_{g_{ec}^l}}^l=1|\mathbf{x}^l))\Big) \tag{11}$$

6. Analysis of Hyper-parameters

We conduct more experiments with different hyper-parameters to further evaluate the impact of the decay factors α_1, α_2 and trade-off parameters β, λ employed in our loss function, as shown in Table 4. The parameter settings α_1=2, α_2=3, β=0.001, λ=0.1 make the network achieve the best period OA of 78.83%.

We adjust one parameter and keep other three parameters to perform the experiments. The decay factors control the degree of attribute learning while the model is learning the main dating task. When α_1 increases from 1 to 2, the OA of period dating increase from 77.32% to 78.83%. And when α_2 increases from 1 to 3, the OA of period dating increase from 78.02% to 78.83%. However, further increasing them makes the performance decrease. Such results show that appropriately reducing the learning of attributes can avoid interference to the main task, but reducing the learning of attributes too much will lead to a decline in the dating performance. The change of trade-off parameters β and λ from 0.0001 to 1 makes the performance of period dating fluctuate in a range. It is worth mentioning that, our network can robustly reach or outperform SOTA single-granularity method P2PNet [5] 77.32% on most parameter settings.

Table 4. The period overall accuracy w.r.t hyper-parameters used in our network. Bold indicates the best results, and underlined values are the second best result. When experimenting with different values for each parameter, we consider other parameters to be optimal.

α_1	1	2	3	4	5
period OA	*77.32*	**78.83**	76.72	77.05	75.48

α_2	1	2	3	4	5
period OA	78.02	*78.07*	**78.83**	77.21	74.57

β	0.0001	0.001	0.01	0.1	1
period OA	77.16	**78.83**	76.72	*78.56*	78.18

λ	0.0001	0.001	0.01	0.1	1
period OA	*78.07*	76.89	76.29	**78.83**	77.32

7. Comparison of Loss Function

Before proposing our new loss function, we implemented WCE loss and Focal loss on our network to learn the eras and attributes information directly. However, they got poor performance compared to our loss function (1.61% OA on dynasty & 2.27% OA on period decreases for WCE loss, and 3.61% OA on dynasty & 3.45% OA on period decreases for Focal loss).

8. Additional Comparison Results

Our proposed network archives SOTA results on coarse (dynasty) categories. Although few independent era classification results are worse than other approaches, we have superb performance (ranked in the top two) on many independent results. Other methods usually only perform well on a few independent categories but poorly on others. We also evaluate the recall performance of all 4 coarse-grained dynasties and 11 fine-grained periods for each independent era classification. And we achieve the best performance recall) for all 4 coarse-grained dynasties and 2 out of 11 fine-grained periods for each independent era classification, as shown in Table 3.

In Figure 2 and 3, we show 66 comparison results between our network and P2PNet [5] , C-HMCNN [2], and HRN [1]. All methods correctly identify most of the typical images of the Shang and Western Zhou dynasties, except for P2PNet [5] which gives a wrong prediction for the first row image of the late Shang dynasty. We believe that this is due to the relatively sufficient sample size.

The Spring and Autumn and Warring States dynasties have fewer samples, thus most of the wrong predictions are concentrated in these two periods. In particular, for the third row of images in the mid Spring and Autumn dynasty, the first row of images in the mid of the Warring States dynasty, and the fourth row of images in the mid Warring States dynasty, all methods give wrong predictions. However, our method makes the least false predictions for typical images of these two harder-to-recognize dynasties.

Furthermore, for failure cases, our network misclassifies them all to adjacent periods, in line with the development law of bronze ding; however, other methods misclassify some of them to relatively distant periods.

References

[1] Jingzhou Chen, Peng Wang, Jian Liu, and Yuntao Qian. Label relation graphs enhanced hierarchical residual network for hierarchical multi-granularity classification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4858–4867, 2022. 2, 5, 7

[2] Eleonora Giunchiglia and Thomas Lukasiewicz. Coherent hierarchical multi-label classification networks. *Advances in Neural Information Processing Systems*, 33:9662–9673, 2020. 5, 6

[3] Guo Moruo. The collection of jin wenci of the zhou dynasty. 1932. 1

[4] Maxim Tkachenko, Mikhail Malyuk, Andrey Holmanyuk, and Nikolai Liubimov. Label Studio: Data labeling software, 2020-2022. Open source software available from https://github.com/heartexlabs/label-studio. 2

[5] Xuhui Yang, Yaowei Wang, Ke Chen, Yong Xu, and Yonghong Tian. Fine-grained object classification via self-supervised pose alignment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7399–7408, 2022. 4, 5, 6

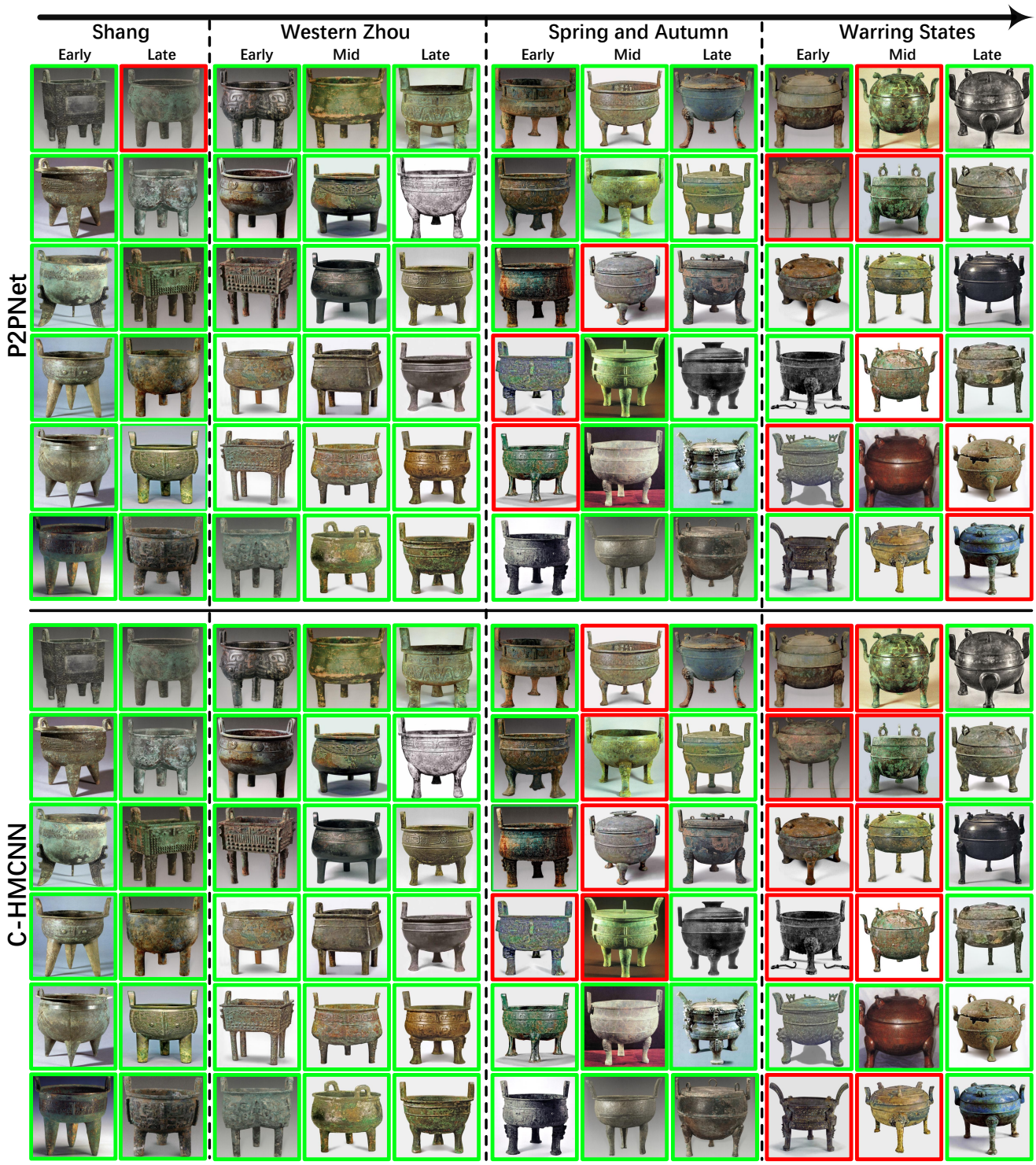

Figure 2. The comparison results of the P2PNet [5] and C-HMCNN [2] on 66 typical examples from 4 dynasties and 11 periods. The green box indicates success cases and the red box indicates failure cases.

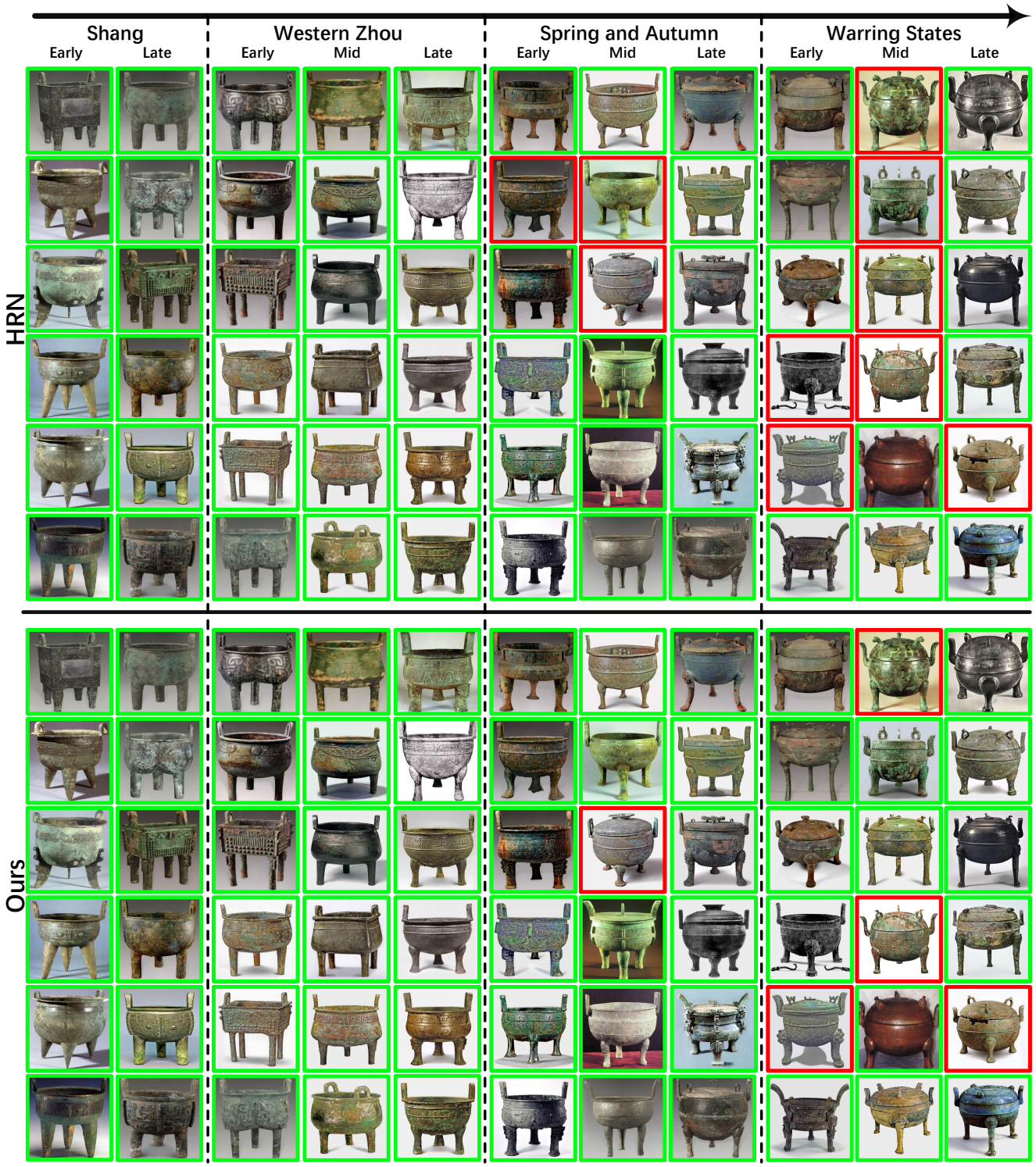

Figure 3. The comparison results of HRN [1] and our method on 66 typical examples from 4 dynasties and 11 periods. The green box indicates success cases and the red box indicates failure cases.