Article

Image Translation for Oracle Bone Character Interpretation

Feng Gao [1,2], Jingping Zhang [3,4], Yongge Liu [1,2] and Yahong Han [5,*]

1 School of Computer and Information Engineering, Anyang Normal University, Anyang 455000, China; gaof@aynu.edu.cn (F.G.); liuyongge@aynu.edu.cn (Y.L.)
2 Key Laboratory of Oracle Bone Inscriptions Information Processing, Ministry of Education, Anyang 455000, China
3 Digital Media Art Department, Shanghai Theatre Academy, Shanghai 200040, China; zhangjingping@sta.edu.cn
4 Key Laboratory of Integrated Innovation of Digital Performing Arts, Ministry of Culture and Tourism, Shanghai 200040, China
5 College of Intelligence and Computing, Tianjin University, Tianjin 300072, China
* Correspondence: yahong@tju.edu.cn

Abstract: The Oracle Bone Characters are the earliest known ancient Chinese characters and are an important record of the civilization of ancient China.The interpretation of the Oracle Bone Characters is challenging and requires professional knowledge from ancient Chinese language experts. Although some works have utilized deep learning to perform image detection and recognition using the Oracle Bone Characters, these methods have proven difficult to use for the interpretation of uninterpreted Oracle Bone Character images. Inspired by the prior knowledge that there exists a relation between glyphs fromOracle Bone Character images and images of modern Chinese characters, we proposed a method of image translation from Oracle Bone Characters to modern Chinese characters based on the use of a generative adversarial network to capture the implicit relationship between glyphs from Oracle Bone Characters and modern Chinese characters. The image translation process between Oracle Bone Characters and the modern Chinese characters forms a symmetrical structure, comprising an encoder and decoder. To our knowledge, our symmetrical image translation method is the first of its kind used for the task of interpreting Oracle Bone Characters. Our experiments indicated that our image translation method can provide glyph information to aid in the interpretation of Oracle Bone Characters.

Keywords: Oracle Bone Character interpretation; generative adversarial network; symmetrical image translation

1. Introduction

The Oracle Bone Characters (OBCs), which were recorded on animal bones and turtle shells, are defined as the earliest known form of Chinese writing and the most famous writing system in the world [1–4]. The OBCs on animal bones and turtle shells record the development of civilization in the Shang dynasty of ancient China. Therefore, ancient Chinese language experts have tried to interpret the connotations of the OBCs to obtain information about the Shang dynasty. After the unearthing of the oracle bones, research on the OBCs has been carried out since the late 1890s, although the interpretation of OBCs is difficult. Firstly, the amount of unearthed oracle bones and turtle shells is small. There are only approximately 4500 different OBCs, which leads to a lack of sufficient OBCs to conduct research. Secondly, detecting and organizing the raw OBCs among the unearthed animal bones and turtle shells requires substantial labor costs. Above all, the great gap between the OBCs and modern Chinese characters leads to difficulties in OBC interpretation, so that only 2200 OBCs have been deciphered to date. To this day, many OBCs still cannot be interpreted fully.

Due to the success of artificial intelligence and deep learning in real-world applications [5,6], some researchers have utilize computers to aid research on the OBCs. Computer-aided OBC studies related with OBC interpretation mainly include OBC detection tasks [7–10] and OBC recognition tasks [11–14].

The task of OBC detection regards OBCs as a special kind of data object in order to perform object detection using rubbing images. These methods detect the locations of OBCs on rubbing images and then classify the OBC. Due to the presence of noise caused by corrosion and excavation, the directly use of object detection methods will lead to poor performance. Therefore, researchers [9] applied the YOLOv4 model to alleviate the problem of noise present on rubbing images. In another study [8] the authors proposed a kind of oracle bone inscription detector based on a single-shot multibox detector to improve the performance of small-object detection in OBC detection tasks. The researchers in another study [7] proposed a simpler but more effective detector for OBC detection based on a kind of anchor-free scheme. These OBC detection methods can reduce the labor costs involved in detecting OBCs in rubbing images.

The task of OBC recognition aims to classify OBCs based on a large-scale labeled OBC dataset [2,3]. Due to the imbalanced distribution of classes in the OBC dataset, the direct use of deep learning methods cannot obtain good performance. To solve the problem of the unbalanced distribution of classes, some researchers [14] proposed a kind of deep metric learning method to conduct classifications based on the nearest-neighbor rule. The authors of another study [15] proposed a mix-up strategy, leveraging majority and minority classes to augment samples and using the triplet loss function to overcome the problem of the imbalanced distribution of classes. Another study [12] integrates self-supervised learning and data augmentation in order to solve the problem of data limitation and imbalance.

The abovementioned studies on OBC detection and OBC recognition require a large-scale labeled dataset to train a model to perform generalizations effectively on unlabeled data. However, there are still nearly 2300 OBCs that need to be interpreted by ancient Chinese language experts. Detection and recognition models can not perform well on these unknown classes. The question of how to utilize deep learning methods to aid in OBC interpretation is challenging and has been the subject of few studies. Due to the fact that modern Chinese characters evolved from ancient Chinese characters, e.g., OBCs, there exists a relation between modern Chinese characters and OBCs in terms of glyphs.Inspired by the application of generative adversarial networks (GAN) in image translation, we aimed to utilize the generative model to capture the relationship between modern Chinese characters and OBCs. In particular, we trained the GAN model to translate the OBCs to modern Chinese characters, aiming to capture the implicit relations between OBCs and modern Chinese characters. Then, we used the GAN model to translate the unknown OBCs into modern Chinese characters in order to provide glyph information for OBC interpretation. Our method utilizes the symmetrical structure of encoding and decoding to conduct image translation. Experiments performed on the OBC dataset indicated the effectiveness of our symmetrical image translation method.

In summary, the contributions of our method are as follows:

(1) In contrast with the existing research, we focus on the more challenging task of OBC interpretation for the under-interpreted OBCs.

(2) We propose a symmetrical method of image translation from OBCs to modern Chinese characters to provide glyph information for OBC interpretation. To our knowledge, this is the first work using the generative model to solve the task of OBC interpretation.

(3) The experimental results on the OBC dataset show that our method can capture the glyph relations between OBCs and modern characters and provide information for OBC interpretation.

2. Materials and Methods

2.1. Related Work

There are some image generation methods [16–18] related with our method, including the generative adversarial network (GAN) [18] and conditional generative adversarial network (cGAN) [19,20].

2.1.1. Generative Adversarial Network

The generative adversarial network (GAN) is a generative model which captures the data distribution by means of a minimax two-player game. A general framework of GAN includes a generative model G and a discriminative model D. During training, the generative model G tries to generate fake data to maximize the probability of D making a mistake in distinguishing between real data and fake data. This procedure increases the generative ability of the generative model G. On the contrary, the discriminative model D tries to distinguish the real data and fake data to maximize the probability of G generating fake data which is different from real data. By alternativelytraining the generative model G and the discriminative model D on the two-player game, the generative model G can capture the training data distribution and generate fake data which has a similar distribution compared with the real data. The overall objective function of GAN can be shown as in Equation (1):

$$\min_{G} \max_{D} V(D,G) = \mathbb{E}_{x \sim p_{data}(x)}[\log D(x)] + \mathbb{E}_{z \sim p_z(z)}[\log(1 - D(G(z)))], \qquad (1)$$

where the p_z is a prior oninput noise variables z and $G(z)$ is the mapping from z to the data space. The GAN, which generates fake data similar to real data, has been widely used in various unsupervised image generation applications [21–23].

2.1.2. Conditional Generative Adversarial Network

The conditional generative adversarial network (cGAN) improves the generative adversarial network (GAN) by generating images conditioned on class labels. Similarly to the framework of GAN, cGAN is also composed of a generative model G and a discriminative model D. During training, the generative model G and the discriminative model D alternatively confuse each other in a minimax two-player game in order to capture the conditional distribution of training data.

Differently from GAN, the generative model G and discriminative model D are conditioned on class labels by feeding class labels y into the generative model G and the discriminative model D. The inputs of generative model G include a noise variable $p_z(z)$ and a class label y. For the discriminative model D, x and the class label y are presented as inputs. The objective function of cGAN can be shown as in Equation (2):

$$\min_{G} \max_{D} V(D,G) = \mathbb{E}_{x \sim p_{data}(x)}[\log D(x|y)] + \mathbb{E}_{z \sim p_z(z)}[\log(1 - D(G(z|y)))], \qquad (2)$$

The cGAN can generate images conditioned with class labels compared with GAN. The cGAN has been widely used in supervised image translation applications [24–26].

2.2. Proposed Method

2.2.1. Problem Description

Character style conversion is a deep learning model designed to transform characters from one style to another style. Convolutional neural networks are very suitable for all kinds of image processing tasks, from feature mapping close to the input layer to feature mapping close to the output layer, which can focus on describing the texture information and content information of the image, respectively. The font style translation task is a special type of image transfer task,performed automatically by the convolution neural networkbased on the depth of high-level features the font style for migration. This will greatly improve the efficiency of font style conversion tasks. Compared with manual

geometric modeling and the non-automatic font design method, the design of an end-to-end network framework without human intervention to handle font style transfer tasks has very important guidance and research significance for a variety of Chinese font design tasks.

The model of oracle bone character evolution can be regarded as a model of character style transformation. Since there is a special character evolution law between Oracle Bone Characters and modern Chinese characters, this paper uses a generative adversarial neural network (GAN) to complete the font style conversion from Oracle Bone Characters to Chinese characters. GAN can be trained end-to-end using backpropagation algorithms that do not require the use of traditional inefficient Markov chains or complex reasoning. In addition, the gradient update information of the generator comes directly from the discriminator, rather than from the data sample, providing a new way to generate models. This paper uses a generative adversarial neural network to fit the evolution law of Oracle Bone Characters to modern Chinese characters, simulates the evolution of Oracle Bone Characters, and then achieves the effect of assisting in interpretation.

After years of development and evolution, the number of modern Chinese characters has gradually increased with history to express new things. The Chinese Character Sea, published in 1994, contains 87,019 characters, whereas the Chinese character database developed by Beijing Guoan Consulting Equipment Company, which has been authenticated by experts, contains 91,251 characters with provenance. However, the number of Chinese characters corresponding to the oracle bones in terms of their examination and interpretation is small, and most of the oracle bones that have not been examined and interpreted correspond to other Chinese characters. As a result, if deep learning is used to solve the problem of the interpretation of oracle-bone characters, the sample deviation between the training set and the test set will be very large, and the Chinese character labels corresponding to the test set and the training set will also be very different. Because of data bias between training data and test data, models that learn from training data often do not effectively generalize sample data from different distributions. In this case, we need to improve the generator's ability to generalize data outside of the training set.

2.2.2. Symmetrical Image Translation Based on Knowledge Expansion

There are three challenges in using GAN networks for the interpretation of Oracle Bone Characters:

(1) Chinese characters have complex structures and different writing styles. It is difficult for the generator to generate a Chinese character image without noise from some skeleton features.

(2) Unlike English or Latin fonts, which contain only a small number of glyphs, even the most commonly used Chinese character set (GB2312) is made up of 6763 characters. The examined and interpreted Oracle Bone Characters correspond to only a small portion of Chinese characters, so many Chinese characters do not appear in the training set, so the generator needs to have some generalization ability.

(3) The gap between Oracle Bone Characters and modern Chinese characters is large, so the generator needs a lot of training to learn some of the rules. However, due to the lack of data on Oracle Bone Characters, too much training will lead to an over-fitting phenomenon and affect the generalization ability of the generator.

Therefore, we need to use some constraints to balance these two results.

In this paper, we propose a structure-guided Chinese character generation system, which combines the prior domain knowledge of Chinese characters with a deep neural network to synthesize correct Chinese character images, as shown in Figure 1. We split the font generation task into two independent processes, namely, font feature coding and font rendering. In the first stage, each Oracle Bone Character is represented as a series of font feature codes. We use a multi-stage CNN model to transform the character encoding of the oracle bone script into the character skeleton of modern Chinese characters. In the second stage, the synthesized character skeleton is rendered in a specific font style via a GAN model to restore shape details on the outline of the glyph. Specifically, output fonts can be

synthesized from a learned skeleton flow vector that represents how corresponding pixels in the input are mapped to corresponding pixels in the output. In this way, we not only make the learning problem more manageable and avoid learning to generate characters from noise but also provide a natural way to preserve the information and structure of the symbol.

As shown in Figure 1, we use a generative model to simulate the evolution process from OBC to modern Chinese characters. The realization of symmetrical image translation based on GAN is more accurately based on cGAN (also called conditional GAN), because cGAN can guide image generation by adding conditional information, so in symmetrical image translation, the input image can be taken as a condition to learn the mapping between the input image and output image in order to obtain the specified output image.
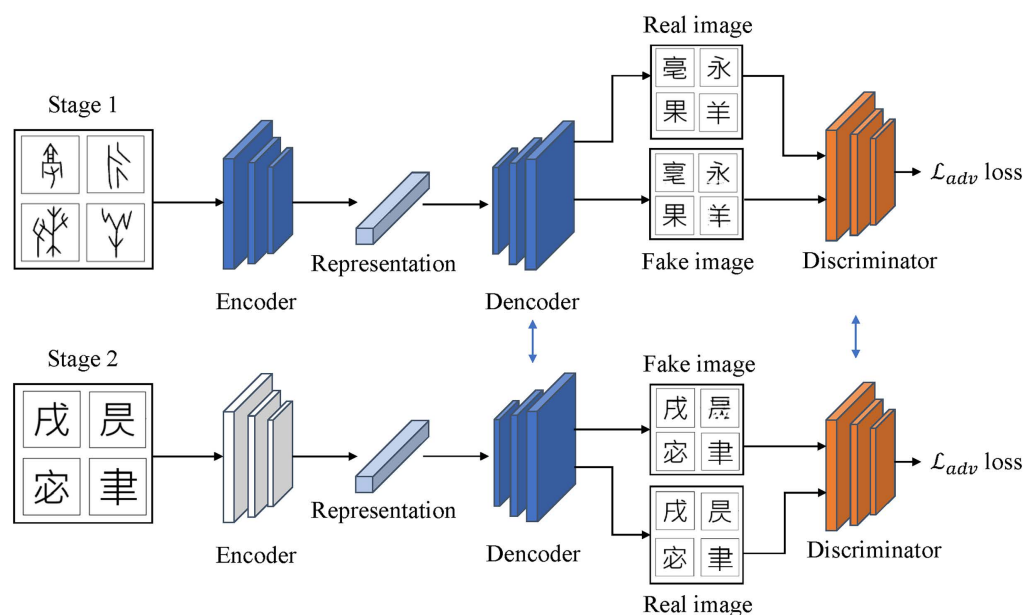

Figure 1. The process of two-stage training, utilizing the symmetrical structure of an encoder and a decoder. In the first stage, the pairs of Oracle Bone Characters and model Chinese characters are used to train the encoder and decoder via adversarial training to translate Oracle Bone Characters into modern Chinese characters. In the second stage, to generate unseen modern Chinese characters, we replace the encoder with a randomly initialized encoder and use the unseen modern Chinese character to fine-tune the decoder by means of adversarial training.

The model needs to use paired input images x and y during training. x is used as the input of the generator G to obtain the generated image $G(x)$, and then $G(x)$ and y are simultaneously input into discriminator D. Finally, as the input of discriminator D, the predicted probability value is obtained, which indicates whether the input is a pair of real images. Thus, the goal of discriminator D is to output a small probability when the input is not a pair of real images, and a large probability when the input is a pair of real images. The goal of generator G is to make the probability value of the output of discriminator D as large as possible when the generated $G(x)$ and y are used as inputs to the discriminator D, which is equivalent to successfully deceiving the discriminator D. Therefore, the adversarial loss can be defined as follows:

$$\mathcal{L}_{adv} = \mathbb{E}_x[\log D(x,y)] + \mathbb{E}_x[\log(1 - D(G(x),y))] \tag{3}$$

where x is the input Oracle Bone Character image and y is the is the corresponding image of modern Chinese characters.

Because the generator of the GAN algorithm is based on the use of random noise to generate images, it is difficult to control the output, so the image generation process is

guided by other constraints. In this paper, we use the distance L_1 to constrain the difference between the generated image and the real image. L_1 is used instead of L_2 to reduce the blur of the generated image.

$$\mathcal{L}_1(G) = \mathbb{E}_{x,y}[\|y - G(x)\|_1] \tag{4}$$

where x is the input Oracle Bone Character image and y is the is the corresponding image of modern Chinese characters.

Therefore, the loss function of generator G in the first stage can be expressed as follows:

$$\mathcal{L}_G = \arg \min_G \max_D \mathcal{L}_{adv} + \mathcal{L}_1(G) \tag{5}$$

The next step is the training setup for the second stage. Through the second stage of training, the decoder learns more about skeleton flow and modern Chinese character correspondence. In the first stage of training, not all modern Characters were able to take part. This is because the existing images of OBC cannot correspond to all modern Chinese character images. A large proportion of the oracle images have thus not been interpreted. Therefore, to ensure that the model is not limited to the modern Chinese character images in the training set, knowledge expansion is carried out on the generator in the second stage. We fine-tune the decoder by replacing the encoder in the generator G with another encoder with the same structure. As shown in Figure 1, we replace the encoder with another encoder.Then, the modern Chinese character images are used as the input for training. In this way, we fine-tune the decoder to expand the scope of knowledge of the decoder. The loss functions are shown as follows:

$$\mathcal{L}'_{adv} = \mathbb{E}_x[\log(1 - D(G(x'), x'))] \tag{6}$$

where x' is the input image of modern Chinese characters.

$$\mathcal{L}'_1(G) = \mathbb{E}_{x'}[\|x' - G(x')\|_1] \tag{7}$$

where x' is the input image of modern Chinese characters.

$$\mathcal{L}'_G = \arg \min_G \max_D \mathcal{L}'_{adv} + \mathcal{L}'_1(G) \tag{8}$$

3. Results

To demonstrate the effectiveness of our method, we split the OBC images data set into three parts: the training set, test set, and uninterpreted data set, to conduct the image translation task for OBC interpretation.

(1) Training set: The training set contains 19,526 OBC images. These OBC images have been interpreted as 1205 traditional Chinese characters by ancient Chinese language experts.

(2) Test set: The test set contains 1520 OBC images. These OBC images have been interpreted as 50 traditional Chinese characters by ancient Chinese language experts, and these 50 traditional Chinese characters do not appear in the training set.

(3) Uninterpreted data set: this data set contains the uninterpreted OBC images, a total of 20,927 OBC images. The uninterpreted OBC images have not been interpreted by ancient Chinese language experts.

We trained the image translation model on the training set and evaluated its performance on the test set to verify the effectiveness of our method. In the training set and test set, we used ground truth labels of the OBC images, enabling us to evaluate the effectiveness of image translation according to classification accuracy or the L_1 distance between the translated images and the ground truth traditional Chinese characters. As a practical problem of OBC interpretation, there were several uninterpreted OBC images, which lacked ground truth labels. We aim to assist the ancient Chinese language experts to interpret these uninterpreted OBC images by utilizing our image translation method. So an additional uninterpreted data set was used in our experiments. We translated the

uninterpreted OBC images to traditional Chinese characters, and the results were evaluated by ancient Chinese language experts, using their professional knowledge.

In our experiments, we used convolution layers to construct the backbone for the encoder and discriminator, as shown in Figure 1. Deconvolution layers were used to construct the backbone for the decoder in our method. The L_1 loss was used in our method for image generation, and the binary cross-entropy loss was used for the discriminator. For the practical computer vision problem, non-linear deep neural networks can achieve better performance compared with linear models.

To achieve a dynamic balance in the system between the encoder, decoder, and discriminator, we conducted a 'two-player' game during training. First, we fixed the discriminator and trained the encoder and decoder for better performance in image generation. Then, we fixed the encoder and decoder to train the discriminator for better discrimination performance using real images and the generated fake images. By training the generator and discriminator alternately in the 'two-player' game, the generator can achieve better generation performance, and the discriminator can achieve better discrimination performance in this dynamic system. Finally, the system achieved stability during training, and we used the encoder and decoder to translate the OBC images into modern Chinese character images. The discriminator was not used in the inference stage.

In our experiments, we first demonstrated the performance of our method on the training set. The results are shown in Figure 2. Our model was able to accurately generate the same modern Chinese characters for the OBC images. In Figure 2, the input of OBC images is shown on the left, and the output of modern Chinese characters is shown on the right. For the OBC images in the training set, our model was able to achieve good generalization ability, and each input OBC image could be translated to its corresponding modern Chinese character image.
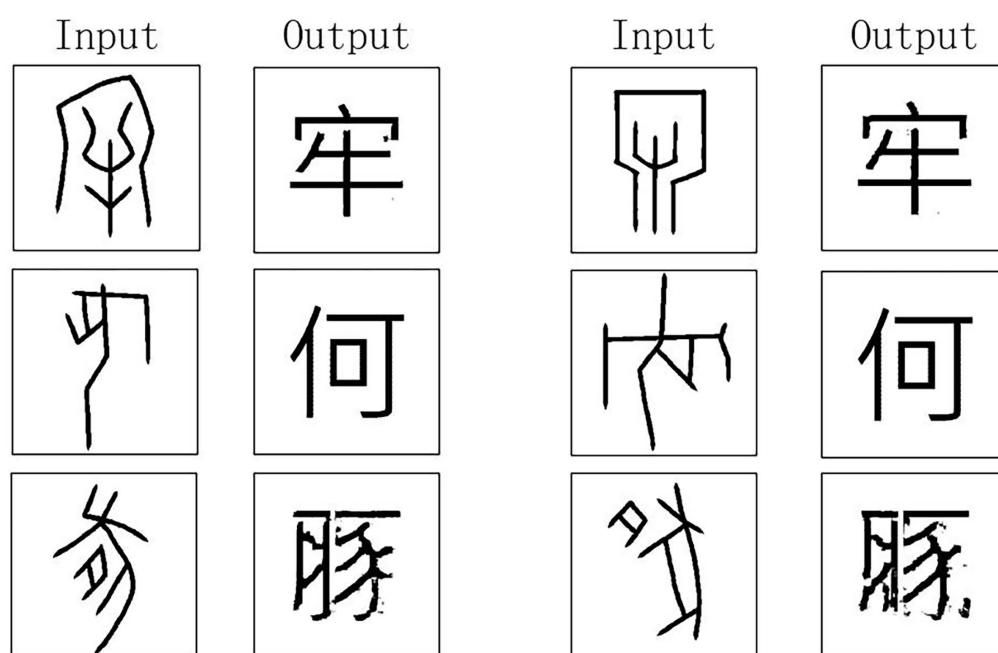

Figure 2. Results of our method on the training set.

To indicate that our method has sufficient generalization ability on the training set, we conducted the second experiment on the training set. We have reported the retrieval results of the top N and calculated the retrieval accuracy of the top N. The experimental results are shown in Table 1.

Table 1. The retrieval accuracy(%) of our method on the training set.

Top N	10	20	30	50	100	200
Accuracy	90.8	93.5	96.3	96.7	97.9	99.4

Then, we tested our model on the test set. We selected some of the experimental results shown in Figure 3. The experimental results in Figure 3 show that the method was effective in analyzing uninterpreted OBC images. The ground truth refers to the modern Chinese character images corresponding to the inputs of the OBC images. We can observe that the output of this model showed consistency with the ground truth images.

Figure 3. Results of our method on test set.

To further illustrate the results of this experiment in a quantitative way, we report the L_1 distance between the output of the model and the ground truth images on the test set for every 50 epoch during training. The experimental results are shown in Table 2. The distance L_1 between the outputs of the model and the ground truth images decreased and then increased during training, which was caused by the overfitting problem after 100 epochs.

**Table 2.** L_1 distance of our method on test set.

epoch	0	50	100	150
L_1	244.2	180.7	173.6	178.3

To verify the reasoning ability of this method for uninterpreted OBC images, we used L_1 distance for the output results to retrieve the modern Chinese characters in the test set. We report the retrieval results of the top N and calculated the retrieval accuracy of the top N. The experimental results are shown in the Table 3.

Table 3. The retrieval accuracy (%) of our method on the test set.

Top N	10	20	30	50	100	200
Accuracy	10.4	15.5	23.1	27.5	35.6	49.0

Finally, we attempted to make inferences using uninterpreted data sets. We obtained some good results, which are presented in Figure 4. As we can see, the OBCs and modern Chinese characters translated using our method are similar in terms of their glyphs, which indicates that our method was able to capture the glyph information between OBC images and modern Chinese characters images. This glyph information can be used to assist ancient Chinese language experts to interpret these uninterpreted OBC images. The results of these translated images were used by ancient Chinese language experts for the interpretation of the uninterpreted OBC images.
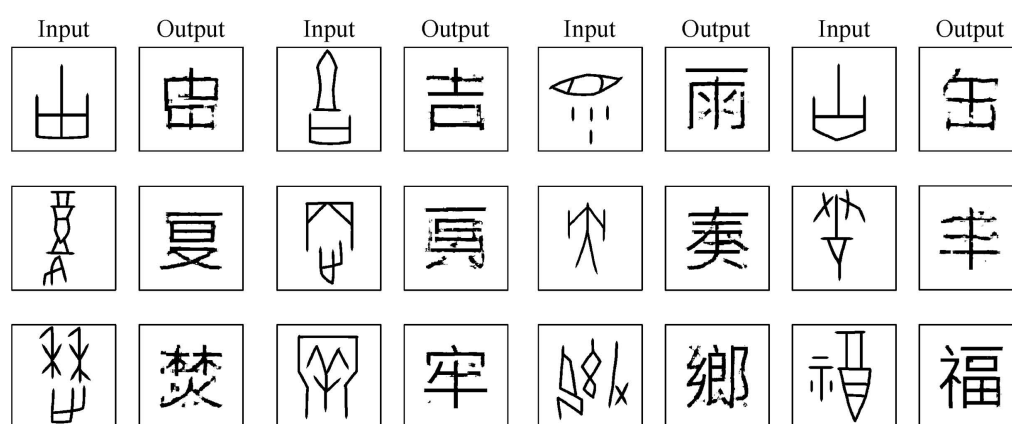
Figure 4. Results of our method on the uninterpreted dataset.

To demonstrate the effectiveness of our method in a practical scenario, we have selected a specific result as an example. As is shown in Figure 5, with the uninterpreted OBC image as the input, the model output a special character, which is similar to a modern Chinese character. This shows that the model exhibited an inference ability and was able to translate the OBC images to modern Chinese characters according to the learned glyph relation between the OBC images and modern Chinese character images. Due to huge gap in terms of glyphs between the OBC images and modern Chinese character images, the effectiveness of translating OBC images to modern Chinese character images still needs to be improved. Translating the unseen uninterpreted OBC images will lead to blurred outputs in some scenarios. However, these blurred outputs also provide important information for interpreting the OBC.

input   output
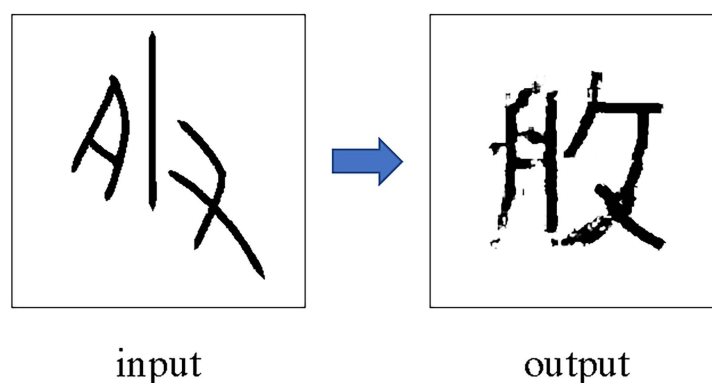
Figure 5. Specific example of results obtained using the uninterpreted dataset.

4. Discussion

We focused in this study on the under-explored task of OBC interpretation. As shown in the experiments, the interpretation of hitherto uninterpreted OBC images is difficult for ancient Chinese language experts. There are large differences in glyphs between OBCs and modern Chinese characters. To assist ancient Chinese language experts in the interpretation of the OBCs, we tried to use the artificial intelligence technology to interpret the uninterpreted OBC images. To capture the glyph relationshipsbetween the OBCs and modern Chinese characters, we have proposed a symmetrical image translation method based on a generative adversarial network. For the uninterpreted OBCs, our symmetrical image translation model was able to provide glyph information for the interpretation of the OBCs. As shown in our experiments, the uninterpreted OBC images can be translated to modern Chinese characters based on the presence of similar glyph information, which can guide ancient Chinese language experts to interpret the OBC images. The partial results of additional experiments on uninterpreted OBC images have provided important information on the relationships between the OBCs and modern Chinese characters for ancient Chinese language experts.

5. Conclusions

In this paper, we proposed a symmetrical image translation method for the OBC interpretation task. By utilizing a deep generative model in computer vision, we were able to capture the glyph relationships between OBCs and modern Chinese characters, which is important for interpreting OBC images. Our experiments demonstrated the ability of our method to capture glyph information. However, there is a great gap between OBCs and modern Chinese characters in terms of glyphs. In the future, we will translate the OBCs into other interpreted ancient Chinese characters, which have a smaller gap in terms of glyphs with OBCs, in order to obtain better performance.

References

1. Flad, R.K. Divination and power: A multiregional view of the development of oracle bone divination in early China. *Curr. Anthropol.* **2008**, *49*, 403–437. [CrossRef]
2. Huang, S.; Wang, H.; Liu, Y.; Shi, X.; Jin, L. Obc306: A large-scale oracle bone character recognition dataset. In Proceedings of the 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, Australia, 20–25 September 2019; pp. 681–688.
3. Li, B.; Dai, Q.; Gao, F.; Zhu, W.; Li, Q.; Liu, Y. Hwobc-a handwriting oracle bone character recognition database. *J. Phys. Conf. Ser.* **2020**, *1651*, 012050. [CrossRef]

4. Zhang, C.; Zong, R.; Cao, S.; Men, Y.; Mo, B. Ai-powered oracle bone inscriptions recognition and fragments rejoining. In Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence, Yokohama, Japan, 11–17 July 2021; pp. 5309–5311.

5. Madhu, P.; Kosti, R.; Mührenberg, L.; Bell, P.; Maier, A.; Christlein, V. Recognizing characters in art history using deep learning. In Proceedings of the 1st Workshop on Structuring and Understanding of Multimedia heritAge Contents, Nice, France, 21 October 2019; pp. 15–22.

6. Vaidya, R.; Trivedi, D.; Satra, S.; Pimpale, M. Handwritten character recognition using deep-learning. In Proceedings of the 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, India, 20–21 April 2018; pp. 772–775.

7. Liu, G.; Chen, S.; Xiong, J.; Jiao, Q. An oracle bone inscription detector based on multi-scale gaussian kernels. *Appl. Math.* **2021**, *12*, 224–239. [CrossRef]

8. Meng, L.; Lyu, B.; Zhang, Z.; Aravinda, C.; Kamitoku, N.; Yamazaki, K. Oracle bone inscription detector based on ssd. In Proceedings of the International Conference on Image Analysis and Processing, Trento, Italy, 9–13 September 2019; pp. 126–136.

9. Wang, N.; Sun, Q.; Jiao, Q.; Ma, J. Oracle bone inscriptions detection in rubbings based on deep learning. In Proceedings of the 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 11–13 December 2020; Volume 9, pp. 1671–1674.

10. Xing, J.; Liu, G.; Xiong, J. Oracle bone inscription detection: A survey of oracle bone inscription detection based on deep learning algorithm. In Proceedings of the International Conference on Artificial Intelligence, Information Processing and Cloud Computing, Sanya, China, 19–21 December 2019; pp. 1–8.

11. Guo, J.; Wang, C.; Roman-Rangel, E.; Chao, H.; Rui, Y. Building hierarchical representations for oracle character and sketch recognition. *IEEE Trans. Image Process.* **2015**, *25*, 104–118. [CrossRef] [PubMed]

12. Han, W.; Ren, X.; Lin, H.; Fu, Y.; Xue, X. Self-supervised learning of orc-bert augmentator for recognizing few-shot oracle characters. In Proceedings of the Asian Conference on Computer Vision, Kyoto, Japan, 30 November–4 December 2020.

13. Liu, M.; Liu, G.; Liu, Y.; Jiao, Q. Oracle-bone inscription recognition based on deep convolutional neural network. *J. Image Graph.* **2020**, *8*, 114–119. [CrossRef]

14. Zhang, Y.K.; Zhang, H.; Liu, Y.G.; Yang, Q.; Liu, C.L. Oracle character recognition by nearest neighbor classification with deep metric learning. In Proceedings of the 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, Australia, 20–25 September 2019; pp. 309–314.

15. Li, J.; Wang, Q.F.; Zhang, R.; Huang, K. Mix-up augmentation for oracle character recognition with imbalanced data distribution. In Proceedings of the International Conference on Document Analysis and Recognition, Lausanne, Switzerland, 5–10 September 2021; pp. 237–251.

16. Arjovsky, M.; Chintala, S.; Bottou, L. Wasserstein generative adversarial networks. In Proceedings of the 34th International Conference on Machine Learning (PMLR), Sydney, Australia, 6–11 August 2017; pp. 214–223.

17. Chen, X.; Duan, Y.; Houthooft, R.; Schulman, J.; Sutskever, I.; Abbeel, P. Infogan: Interpretable representation learning by information maximizing generative adversarial nets. In Proceedings of the Advances in Neural Information Processing Systems, Barcelona, Spain, 5–10 December 2016; Volume 29.

18. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 8–14 December 2014; Volume 27.

19. Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1125–1134.

20. Antipov, G.; Baccouche, M.; Dugelay, J.L. Face aging with conditional generative adversarial networks. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 2089–2093.

21. Esser, P.; Rombach, R.; Ommer, B. Taming transformers for high-resolution image synthesis. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 12873–12883.

22. Karras, T.; Aittala, M.; Hellsten, J.; Laine, S.; Lehtinen, J.; Aila, T. Training generative adversarial networks with limited data. In Proceedings of the Advances in Neural Information Processing Systems, Virtual, 6–12 December 2020; Volume 33, pp. 12104–12114.

23. Zhu, J.; Shen, Y.; Zhao, D.; Zhou, B. In-domain gan inversion for real image editing. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; pp. 592–608.

24. Rodríguez-de-la-Cruz, J.A.; Acosta-Mesa, H.G.; Mezura-Montes, E.; Cosío, F.A.; Escalante-Ramírez, B.; Montiel, J.O. Evolution of conditional-gans for the synthesis of chest X-ray images. In Proceedings of the 17th International Symposium on Medical Information Processing and Analysis, Campinas, Brazil, 17–19 November 2021; Volume 12088, pp. 85–94.

25. Deng, Y.; Yang, J.; Chen, D.; Wen, F.; Tong, X. Disentangled and controllable face image generation via 3d imitative-contrastive learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 5154–5163.

26. Odena, A.; Olah, C.; Shlens, J. Conditional image synthesis with auxiliary classifier gans. In Proceedings of the International Conference on Machine Learning (PMLR), Sydney, Australia, 6–11 August 2017; pp. 2642–2651.