FFD Augmentor: Towards Few-Shot Oracle Character Recognition from Scratch[⋆]

Xinyi Zhao[0000−0003−0555−243X], Siyuan Liu[0000−0002−6002−3779],
Yikai Wang[0000−0001−6107−5063], and Yanwei Fu[⋆⋆][0000−0002−6595−6893]

School of Data Science, Fudan Univerisity
{xyzhao19, liusiyuan19, yikaiwang19, yanweifu}@fudan.edu.cn

Abstract. Recognizing oracle characters, the earliest hieroglyph discovered in China, is recently addressed with more and more attention. Due to the difficulty of collecting labeled data, recognizing oracle characters is naturally a Few-Shot Learning (FSL) problem, which aims to tackle the learning problem with only one or a few training data. Most current FSL methods assume a disjoint but related big dataset can be utilized such that one can transfer the related knowledge to the few-shot case. However, unlike common phonetic words like English letters, oracle bone inscriptions are composed of radicals representing graphic symbols. Furthermore, as time goes, the graphic symbols to represent specific objects were significantly changed. Hence we can hardly find plenty of prior knowledge to learn without negative transfer. Another perspective to solve this problem is to use data augmentation algorithms to directly enlarge the size of training data to help the training of deep models. But popular augment strategies, such as dividing the characters into stroke sequences, break the orthographic units of Chinese characters and destroy the semantic information. Thus simply adding noise to strokes perform weakly in enhancing the learning capacity.

To solve these issues, we in this paper propose a new data augmentation algorithm for oracle characters such that (1) it will introduce informative diversity for the training data while alleviating the loss of semantics; (2) with this data augmentation algorithm, we can train the few-shot model from scratch without pre-training and still get a powerful recognition model with superior performance to models pre-trained with a large dataset. Specifically, our data augmentation algorithm includes a B-spline free form deformation method to randomly distort the strokes of characters but maintain the overall structures. We generate 20 - 40 augmented images for each training data and use this augmented training set to train a deep neural network model in a standard pipeline. Extensive experiments on several benchmark datasets demonstrate the effectiveness of our augmentor. Code and models are released in https://github.com/Hide-A-Pumpkin/FFDAugmentor.

---

1   Introduction

Oracle inscriptions are one of the earliest pictographs in the world. Dating back to Shang Dynasty [1], ancient Chinese carved oracle characters on animal bones or tortoise shells to record the present and divine the future. Due to the gradual abandonment of oracle characters and the loss of oracle samples in the long history, we now can only discover very limited oracle bone inscriptions, and most of the characters are incompletely preserved. What is worse, as oracle bone inscriptions were created by people of different ethnic groups in different regions and were written on nail plates of various shapes and curvatures, the oracle characters are hard to recognize and distinguish from each other.

In the early time, archaeologists [1] could identify some widely used and easily identifiable characters. Then, with recognition models, researchers [2–4] can identify some new characters after training on annotated oracle characters. However, the ultimate goal of fully understanding the oracle inscription system is far from attainable since there remain a lot of undecipherable oracle pieces. For those deciphered words, mostly we only collect very limited characters [5] that is far from enough to train a powerful recognition model. To go further, we can formulate the oracle recognition task as a Few-Shot Learning (FSL) [6, 7] problem targeted at recognizing oracle characters in the limited data case.

FSL is a popular machine learning topic that aims to train a learning model with only a few data available. Based on the motivation to solve the limited training data issue, we can roughly classify FSL algorithms into the following categories: (1) Learning to learn, or meta-learning algorithms [8–14], aims to train the few-shot model to learn the capacity of learning from a few examples by simulating the few-shot scenario in the training stage; (2) Data-augmentation algorithms [5, 15, 16] directly expand the training set by generating new data based on the few training samples; (3) Semi-supervised algorithms [17–23] have the access to the additional unlabeled dataset and try to utilize this unlabeled knowledge to help train the few-shot model.

However, most FSL algorithms assume the existence of related large labeled datasets to pre-train the few-shot model. But as we do not have such datasets, these algorithms are not suitable for the oracle character recognition problem. Current algorithms to tackle the oracle character recognition problem in the FSL pipeline mainly focus on data-augmentation, including using hierarchical representation [3], dividing characters into structure and texture components [24], converting characters to stroke vectors [5].Nevertheless, they still fail to solve the problem of limited training data for their under-utilization of structured information and the mining of stroke information is limited and costly.

In this paper, we propose a new data augmentation approach by adopting Free Form Deformation (FFD) [25] which is initially used in the field of non-rigid registration [26]. FFD deforms the image by manipulating an underlying
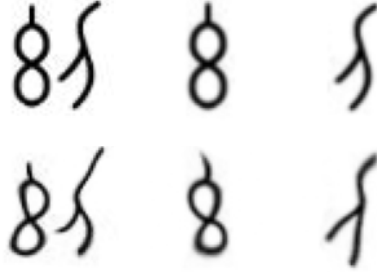
Fig. 1. The oracle character 'Ci' (in the left column) is made up of two parts, an eight-like character (in the middle column) and a knife-like radical (in the right column). After the FFD transformation (from top to bottom), though the strokes are distorted, the overall structure remains unchanged. We can still re-recognize the character 'Ci' based on them.

mesh of control points [25] and calculates the displacement rule for each pixel in the image individually. When it is applied to oracle character images, each pixel that makes up a stroke moves by their each displacement rule, which leads to the distortion of strokes and the corruption of local information. Meanwhile, the two adjacent pixels' displacement rules are similar in general, which maintains the consistency of radicals and stability of the global information (See Fig. 1 as an example). By corrupting local information while maintaining global information, FFD well preserves the radical features of oracle bone inscriptions and randomly distorts stokes, making the augmented image more representative.

With this FFD Augmentor, we can now tackle the few-shot oracle characters recognition by utilizing the online Free Form Deformation algorithm as our augmentor and generating a bunch of augmented data from each annotated training instance for each category. As the generated training data is of high quality and diversity, we now make it possible to train the few-shot model in a standard supervised manner with our generated data *from scratch*. To better show the effectiveness of our FFD Augmentor, we select a powerful few-shot training method called EASY [27], which is composed of widely used training modules without specifically designed algorithms, and train the model with the generated data in a standard pipeline to ensure that our proposed algorithm can be utilized as a general module for the oracle recognition task. Extensive experiments on benchmark datasets, Oracle-FS [5] and HWOBC [28], verify the effectiveness of our proposed algorithm. We further conduct experiments on a sketch dataset [29] to show that the effectiveness of our augmentor is not limited to the oracle characters.

The main contributions of our work are as follows:

(1) To the best of our knowledge, we are the first to apply the non-rigid transformation, namely FFD, to the field of data augmentation.

(2) Our generated training data are diverse and informative such that the deep model can be trained with generated data from scratch without the help of an additional large unlabeled or labeled dataset.

(3) We demonstrate the effectiveness of our approach through comprehensive experiments on Oracle-FS and HWOBC datasets, reaching a significant accuracy improvement over competitors.

2   Related Works

2.1   Oracle Character Recognition

Here we mainly survey machine learning algorithms for oracle character recognition. Conventional researches [30–32] regarded the characters as a non-directional graph and utilized its topological properties as features for graph isomorphism. Guo et al. [3] proposed a hierarchical representation and released a dataset of hand-printed oracle characters from 261 different categories.

In the deep learning era, neural networks gradually became the mainstream recognition method. Yang et al. [33] studied several different networks and suggested using the Gabor method for feature extraction. Zhang et al. [34] utilized the nearest-neighbor classifier to reject unseen categories and configure new categories. Other models like capsule network [4] and generative model [24] are also proposed to solve this task.

However, few researchers focus on few-shot learning of oracle character recognition. Orc-Bert [5] converts character images to stroke vectors and learns the stoke features from large unlabeled source data. The method requires large unlabeled source data and the generated augmented stroke data performs average compared with the benchmark. Instead, we show that a single oracle character is informative enough to train the recognition model via our proposed data augmentation algorithm without pretraining on related large datasets.

2.2   Few-Shot Learning

Few-Shot Learning (FSL) [35, 36] aims to train a machine learning model with only a few training data. Gidaris et al. [6] created an attention-based few-shot classification weight generator with a small number of gradient steps. MAML [8] searched for the best initial weights to accelerate the learning process reducing the risk of over-fitting. Chen et al. [37] applied self-supervised learning in a generalized embedding network to provide a robust representation for downstream tasks. LEO [38] reduced the complexity by learning a low-dimension model embedding and used the nearest neighbor criterion to classify. Chen et al. [39] generated multiple features at different scales and selected the most important local representations among the entire task under the complex background.

In this paper, we used the EASY [27] framework, which combined several simple approaches like backbone training, and featured vectors projection in the above literature, and reached the state-of-art performance in training without huge computation cost.

2.3   Data Augmentation Approaches

Traditional data augmentation methods are mainly focused on rigid transformation such as flipping [40], rotation [9], shearing [41] and cropping [40]. Blending or occluding[42] parts of images is also widely used, but they require expert domain knowledge to prevent critical information from being disrupted. In the context of few-shot learning, some efficient data augmentation algorithms like AutoAugment [43] are not applicable for the lack of large-scale labeled data.

Distinguished from the traditional image, the character image is regarded as an intermediate between text and image. As an image, it retains textual features, such as the ability to be reconstructed into a sequence of strokes. Han et al. [5] captured the stroke order information of Chinese characters and utilized the pre-trained Sketch-Bert model to augment few-shot labeled oracle characters. Yue et al.[44] designed a dynamic dataset augmentation method using a Generative Adversarial Network to solve the data imbalance problem. However, none of the state-of-art data augmentation approaches addressed the problem of oracle bone characters recognition from the overall structure.

As the ancestor of Chines characters, the oracle character is at least a logographic language and the upper-level orthographic units, like radicals, should contain richer information. Simply introducing noise to strokes cannot enhance or may even weaken the model's ability to recognize radical components.

2.4   Non-Rigid Transformation

Non-Rigid transformation is a critical technique in the field of image registration, which focuses on finding the optimal transformation from one input image to another [26]. In the medical field, extensive research has been conducted on non-Rigid transformation for its essential role in analyzing medical effects over time. Existing non-rigid transformation techniques include AIR [45], Diffeomorphic Demons [46], and FFD [47].

FFD is a commonly used algorithm for image registration [25]. The registration from moving image to fixed image is modeled by the combined motion of a global affine transformation and a free-form deformation based on B-splines. Compared with rigid transformation, FFD has a higher degree of freedom which can better model the motion between two images. When applied to data augmentation, this can also bring greater flexibility to enlarge the training dataset.

3   Methodology

3.1   Problem Formulation

Here we define the few-shot oracle character learning problem without related large datasets for pre-training, thus degenerating to the standard supervised learning manner. We are provided with labeled dataset, $\mathcal{D}$, which comprises the category set $\mathcal{C}$, $|\mathcal{C}| = n$. For a certain k-shot learning task, our augmentor and classifier would only have access to k annotated training instances for each
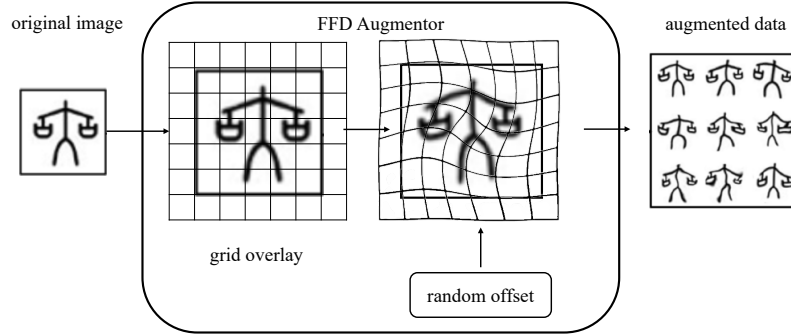
Fig. 2. Illustration of our FFD Augmentor. We generate random offsets for each control point and construct a deformed image based on the recalculation of the world coordinates of each vertex of the image.

category. We randomly sample k and q images for each category $C_i \in \mathcal{C}, i = 1, ...n$ to construct the training set $\mathcal{S}$ and the evaluation set $\mathcal{Q}$ respectively. We aim to train on $\mathcal{S}$ and generalize to $\mathcal{Q}$. We take accuracy on $\mathcal{Q}$ as the evaluation metric.

3.2 Overview of Framework

As shown in Fig. 2, our data augmentation method, FFD Augmentor, consists of several parts. For each image, we first create a local coordinate system by splitting the whole image into local patches using grids. The grid vertexes are used as the control points to define the local position of the neighboring pixels. Then random offsets are generated to shift the control points, thus shifting the neighboring pixels. Due to random offsets being used to shift control points, the whole image is modified with destroyed local information.

Using our proposed FFD Augmentor, we generate several augmented data for each training image and then store them to expand the training set. With this expanded training set, we now can train the few-shot model from scratch.

3.3 FFD Augmentor

Though there are few studies on the composition of oracle characters, as the ancestor of Chinese characters [48], oracle characters intuitively retain similar characteristics. For example, compared with strokes, radicals contain richer information. This motivates us to perform a data augmentation algorithm to generate local diversity while preserving global structures.

In non-rigid registration, researchers [25, 49, 50] applies FFD to achieve this goal by calculates displacement rule for each pixel in the image individually. When it is applied to oracle character images, each pixel that makes up a stroke moves according to their each displacement rules. Thus it leads to the distortion of strokes and the corruption of local information. On the other hand, thanks to the continuity of these features and physical rules, the displacement rules for two adjacent pixels are similar in general. Globally, the relative positions of radicals remain consistent and the character's structure is well preserved. See Fig. 1 for an illustration. Hence in this paper, we adopt FFD to generate new training data for few-shot oracle character recognition problem.

Free Form Deformation. As the oracle characters can be represented as a grayscale image, we implement the 2D version of Free Form Deformation based on B-splines [25]. Specifically, for the oracle character grayscale image $\boldsymbol{x} \in \mathbb{R}^{h \times w}$, we design a two-dimensional mapping

$$\mathrm{T} : (x_1, x_2) \quad \rightarrow \quad (x_1^{'}, x_2^{'}), \tag{1}$$

to simulate the non-rigid transformation. We decouple T by a global deformation mapping and a local deformation mapping as:

$$\mathrm{T}(x_1, x_2) = \mathrm{T_{global}}(x_1, x_2) + \mathrm{T_{local}}(x_1, x_2). \tag{2}$$

The global deformation mapping is a simple affine transformation defined as:

$$\mathrm{T_{global}}\,(x_1, x_2) = \begin{pmatrix} \theta_{11}\ \theta_{12} \\ \theta_{21}\ \theta_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} \theta_{13} \\ \theta_{23} \end{pmatrix}, \tag{3}$$

while the local deformation mapping is the major concern in our algorithm. Specifically, we first distribute a series of grid points over the image at a certain spacing based on the predetermined patch number. Denote the area of the oracle character grayscale image as $\Omega = \{(x_1, x_2) \mid 0 \le x_1 \le X_1, 0 \le x_2 \le X_2\}$, we split it by control points $\Phi = \{\phi_{i,j}\}$ into several patches of size $n_1 \times n_2$, where n_i is the distance between adjacent control points in the i-th dimension. Then we can define the local deformation mapping as the product of B-splines functions:

$$\mathrm{T_{local}}(x_1, x_2) = \sum_{l=0}^{3} \sum_{m=0}^{3} B_t(u) B_m(v) \phi_{i+l, j+m}, \tag{4}$$

where $i = \lfloor x_1/n_1 \rfloor - 1, j = \lfloor x_2/n_2 \rfloor - 1, u = x_1/n_1 - \lfloor x_1/n_1 \rfloor, v = x_2/n_2 - \lfloor x_2/n_2 \rfloor$, and the B-splines functions are defined as:

$$
\begin{aligned}
B_0(u) &= \frac{(1-u)^3}{6}, B_1(u) = \frac{3u^3 - 6u^2 + 4}{6}, \\
B_2(u) &= \frac{-3u^3 + 3u^2 + 3u + 1}{6}, B_3(u) = \frac{u^3}{6}.
\end{aligned}
\tag{5}
$$

Then the augmentation comes when we randomly apply offsets with a predefined range $O = [O_{\min}, O_{\max}]$ to shift the control points. For a specific grid

---

Algorithm 1 FFD Augmentor

Require: Training Image $\boldsymbol{x}$, patch size of $n_1 \times n_2$, offset range O.
Ensure: Augmented Image.
 1: Initialize the new image with empty pixel value: $\tilde{\boldsymbol{x}}$;
 2: For each control point C, random initialize its degree of shift based on Eq. (6);
 3: **for** all pixel $\boldsymbol{x}_i = (x_{i,1}, x_{i,2})$ in $\boldsymbol{x}$ **do**
 4:      Generate the new coordinates $\tilde{\boldsymbol{x}}_i = (\tilde{x}_{i1}, \tilde{x}_{i2})$ based on Eq. (4);
 5: Use bi-linear interpolation algorithm to complete unvisited pixels in $\tilde{\boldsymbol{x}}$.
 6: **return** $\tilde{\boldsymbol{x}}$

---

point $\phi_{i,j} = (x_{\phi_i}, x_{\phi_j})$, we randomly initialize its degree of shift within the offset range.

$$\mathrm{T} : (x_{\phi_i}, x_{\phi_j}) \quad \rightarrow \quad (x_{\phi_i} + \Delta x_{\phi_i}, x_{\phi_j} + \Delta x_{\phi_j}), \quad \Delta x_{\phi_i}, \Delta x_{\phi_j} \in O. \quad (6)$$

Then for each pixel (x_1, x_2) within the image, we calculate the deformed location based on Eq. (4) with the shifted control points deformed in Eq. (6). After the displacement transformation rules for all pixels are determined, we finally re-sample the image by pixels according to their rules to achieve non-rigid deformation. If the transformed pixel coordinates exceed the image size, the grayscale value will be replaced by 255. Finally, for the unvisited pixels in the generated image, we use a bi-linear interpolation algorithm [51] to fill these empty holes.

Augmentor. As mentioned in Section 1, non-rigid transformation can destroy local information while maintaining global information. By generating multiple FFD-augmented training samples, the model extracts and learns the structured information of the oracle characters, rather than relying on some particular strokes to classify the character. This is critical in the task of few-shot oracle character recognition since it will alleviate the problem of bias and overfitting caused by the limited training samples.

Algorithm 1 illustrates the detailed pseudo-code for our FFD Augmentor. An ablation study about the number of FFD-augmented training samples and the selection of FFD hyperparameters will be further discussed in Section 4.3.

3.4 Training with FFD Augmentor

To show the effectiveness of the FFD Augmentor, we adopt a popular training algorithm, the Ensemble Augmented-Shot Learning method (EASY) [27] to combine with our proposed FFD Augmentor.

Specifically, we test our algorithm on several widely used CNN architectures, including ResNet-12 [52], ResNet-18, ResNet-20, and WideResNet [53], respectively. For the FFD augmented training set, we also apply standard data augmentation strategies, including cropping, flipping, and color jittering. When training, each mini-batch is divided into two parts: the first part is input to the

Table 1. Accuracy (%) of oracle characters recognition on **Oracle-FS** under all three few-shot settings with classifiers ResNet18 including Orc-Bert, EASY with and without FFD augmentor. Because we only share the ResNet18 classifier in common with Orc-Bert, we compare with their best-performance method, i.e. Orc-Bert Augmentor with point-wise displacement on ResNet18.

k-shot	Orc-Bert	EASY	FFD + EASY
1	31.9	55.17	**78.90**
3	57.2	79.45	**93.36**
5	68.2	90.34	**96.47**

Table 2. Accuracy (%) of oracle characters recognition on **Oracle-FS** under all three few-shot settings with different architectures. The Basic model is the pure model without any augmentation method involved.

k-shot	model	Basic	FFD+Basic	EASY	FFD + EASY
1	ResNet12	14.95	33.95	58.46	**76.79**
	ResNet18	21.22	26.47	55.17	**78.90**
	ResNet20	8.61	30.28	54.67	**77.09**
	WideResNet	15.35	33.85	53.77	**77.75**
3	ResNet12	33.06	68.83	84.91	**92.89**
	ResNet18	35.90	57.25	79.45	**93.36**
	ResNet20	32.28	66.08	82.53	**92.57**
	WideResNet	40.23	71.12	84.81	**93.42**
5	ResNet12	54.53	77.92	91.30	**95.38**
	ResNet18	46.70	70.00	90.34	**96.47**
	ResNet20	55.82	77.60	91.55	**95.26**
	WideResNet	60.62	78.38	92.43	**97.59**

standard classifier with the feature augmentation strategy Manifold-MixUp [54]; the second part is with the rotation transformation and input to both heads. For details on training the EASY model, we suggest to read the original paper [27].

4   Experiments

We conduct extensive experiments to validate the effectiveness of our FFD augmentor and provide ablation studies to analysis each part of our algorithm.

4.1   Experimental Settings

Datasets. We demonstrate the effectiveness of our FFD Augmentor on Oracle-FS [5] and HWOBC [28]. Oracle-FS contains 200 oracle character categories. We run experiments on 3 different few-shot settings, including k-shot for $k = 1, 3, 5$ where for each category we only have access to k labeled training data.

Table 3. Accuracy (%) of oracle characters recognition on **HWOBC** under 1-shot setting with different architectures.

Model	Basic	FFD+Basic	EASY	FFD + EASY
ResNet12	14.92	55.72	65.24	**98.53**
ResNet18	15.73	38.43	41.85	**98.97**
ResNet20	7.65	54.63	52.32	**98.60**
WideResNet	10.2	63.37	62.16	**99.52**

To evaluate the performance, we randomly select 20 instances to construct the testing dataset for each category. HWOBC consists of 3881 oracle character categories, each containing 19 to 25 image samples. We randomly selected 200 categories, and each category is divided into 1-shot training sets and 15-sample test sets. Because the accuracy of our model in the 1-shot setting is high enough, we did not test on more k-shot settings.

Competitors. We mainly compare our results with Orc-Bert [5], the SOTA algorithm for the few-shot oracle character recognition task. Orc-Bert masks some strokes, predicts them by a pre-trained model and make additional noise on each stroke to generate multiple augmented images. We also train EASY [27] for both with FFD augmentor and without FFD augmentor to compare the results.

Implementation Details. We implement FFD Augmentor training methods using PyTorch [55]. The number of training epochs is 100 with a batch size of 64. Unless otherwise specified, We follow the hyper-parameters for training EASY in their default settings. We conducted experiments with the FFD Augmentor of 5 patch num, 11 max offset, and 30 augmented samples. For 1-shot and 3-shot, we used a cosine scheduled learning rate beginning with 0.1, while we used a learning rate of 0.01 for 5-shot. The images are resized to 50×50.

4.2   Evaluation of FFD Augmented Training

It can be clearly noticed in Table 1 that our FFD-based data augmentor can defeat the state-of-the-art method by more than 30% under all few-shot settings. Also can be seen in Table 2, our data augmentation method plays a decisive role in improving accuracy.

On Oracle-FS, our 1-shot accuracy reaches 76.5% for all the classifiers, which outperforms EASY without being augmented by 20%. Our 3-shot accuracy achieved 93.42%, exceeding the accuracy of all the 5-shot models without FFD augmentation. For the 5-shot setting, our model's accuracy reaches 97.59% on WideResNet.

On the HWOBC dataset, the effect of our data augmentation tool is more prominent. We compare the result with EASY and Conventional Data Augmentation. As seen from the table 3, our FFD augmentor improves the accuracy from the original 65.24% to 99.52% for the 1 shot setting.
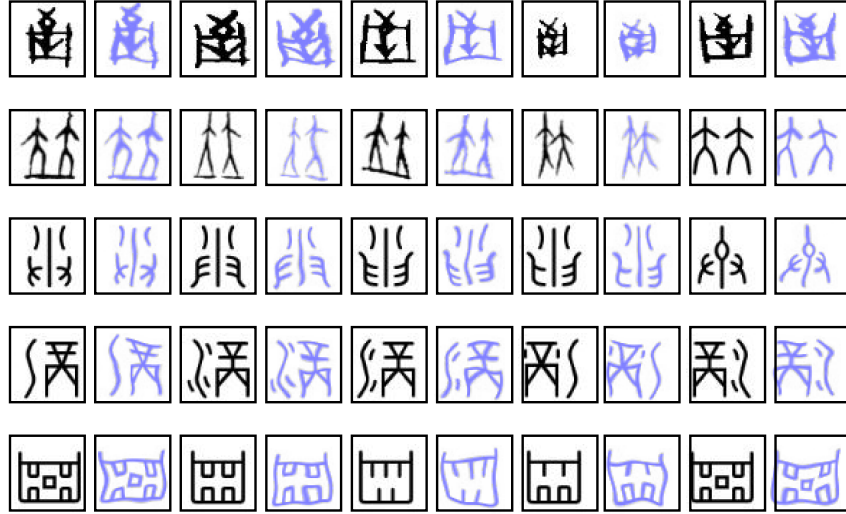
Fig. 3. Examples of oracle character images and the FFD-augmented samples.

4.3   Further Analysis of FFD Augmentor

Visualization. Here we provide more visualization of the oracle characters and FFD augmented images in Fig. 3. Clearly, for all kinds of oracle characters, the FFD Augmentor will consistently generate diverse and informative images. Hence we can provide many realistic augmented images to alleviate the lack of training data in the few-shot oracle character recognition task.

Ablation Study. In this part, we conduct more experiments to evaluate our FFD augmentor, including the min-max offset, the number of patches, and augmented samples. These experiments are running with ResNet18. To make our results more accurate, all experiments on hyperparameters are conducted twice and we average the results of the two experiments as the final accuracy.

(1)Max Offset Value: In our FFD augmentor, the random offset value was generated through a uniform distribution in the interval between the minimal and the maximal offset value. The maximal offset value is set as a hyperparameter while we negate it as the minimal offset. They together limit the movement range of the offset. The closer they are, the smaller the deformation of the image. In our experiment, we tested the maximum value from 0 to 15. Max offset value of 0 indicates that no free form transformation is performed on the original image, i.e., as an ablation experiment for our data augmentor.

(2)Num of Patches: The number of patches influences the number of control points of the grid. The more patches in the FFD transformation, the more transformation control points there are, and the more complex the deformation can be. In our realization, we test the number of patches from 3, 5 to 7. Considering
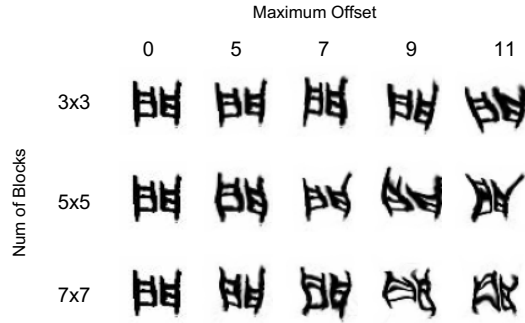
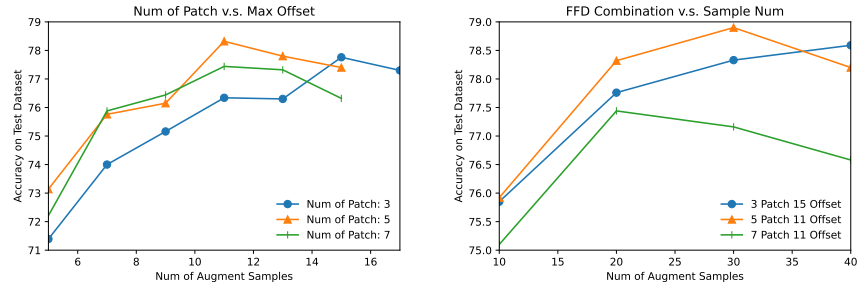**Fig. 4.** Illustration of num of patches and maximum offset difference.

Fig. 5. Left: Different combination of Patch numbers and max offsets varies in accuracy. **Right:** Num of Augmented Sample influences the model accuracy.

the effectiveness of the transformation and the time overhead, we did not test on more patch nums. The time of generating the augmented dataset is proportional to the square of patch nums. When the number of patches increases to a large amount, the enhancement effect of the picture is not obvious but will take considerable time, much longer than the training cost.

Take Fig. 4 as an example, when the max offset value and num of patches are both limited, the deformation is closer to rigid transformation and the overall shape of the text remains unchanged. But when the num of patches and maximum offset value becomes too large, the deformation is so complex that the overall structure is severely damaged and the oracle character is hard to identify. As max offset value and num of patches influence the deformation process in different ways, we experimented on the trade-off between max offset value and num of patches. The results are shown in Fig. 5. With the rise of the max offset value, the accuracy of all patch nums increases. However, their accuracy begins to decrease when the offset exceeds a certain threshold. As can be seen, the top three accuracy combinations of patch Num and max offset is $(3, 11)$, $(5, 15)$ and $(7, 11)$.

Table 4. Top-1 Accuracy under different learning rates in the 5-shot task.

Learning rate	0.1	0.05	0.01	0.005	0.001
Accuracy	95.82	96.31	**96.47**	89.08	87.05

Table 5. Top-1 Accuracy under different support samples with no FFD augmentation.(k-shot=1)

Aug-sample	1	5	10	20	40
Accuracy	52.94	53.39	**53.52**	53.40	53.36

Num of Augmented Sample. We then experimented on the num of augmented samples to figure out whether the larger num of augmented samples contributes to the better performance of our model. Here, we mean generate num of augmented images for each training image. For example, under the 3-shot setting, if the num of Augmented Sample equals 30, we generate 90 augmented images.

Intuitively, with more augmented samples, the accuracy will be higher. However, the results of experiments (See Fig. 5) show that due to the limited number of samples, too many augmented images will lead to overfitting, i.e., the test accuracy will become lower with the decrease of training loss. Two FFD combinations show an increasing trend followed by a decreasing trend and the growth trend is also diminishing for the combination of 3 patches and 15 offset value. When the number of augmented samples equals 30, the FFD combination of 5 patch num and 11 max offset reaches the maximum accuracy.

Besides, the computation time of data augmentation is a crucial factor to be considered. FFD is a time-intensive transformation that increases with the size of the image and the number of patches. The flaw of FFD is less fatal in few-shot learning for a small number of images. Our FFD augmentor takes about 0.4 to 0.5s to generate each image of size 50×50 for 5 patch num. Due to the expensive time cost for data augmentation, we trade off both performance and computation time. Combining all the results above, we find that with 5 patch num, 11 max offset and 30 augment samples, our model achieves the best performance of 78.9% in 1-shot.

Learning Rate. Different learning rates affect the convergence speed and accuracy of the model. A low learning rate may cause the model to overfit the training dataset or converge too slowly. A high learning rate may prevent the model from convergence. We experimented on the influence of different learning rates under different k-Shot settings. As shown in Table 4, for k=5, learning rate of 0.01 reaches the highest accuracy.

Num of Augmented Samples Using Random Crop. Our experiments also test the number of image samples processed by random crop and flip before backbone training. The results in Table 5 show that the accuracy rate is highest when the size is 10, and there is a risk of overfitting when the size is larger.

4.4   Applicability To Other Problems

Though our paper is intended to tackle oracle character recognition, the innovative augmentor we proposed has much broader applications. To better demonstrate its versatility, we here report more experiments on the sketch recognition task. The task takes a sketch s as input and predicts its category label c. We did the toy experiment on the sketch dataset [29], which contains 20,000 unique sketches evenly distributed over 250 object categories, with the same dataset setting as before. The experiment results in Table 6 demonstrate superior performance after adopting the FFD Augmentor, which is 15% higher than the EASY without FFD augmentation. More applications of the augmentor will be explored in future works.

Table 6. Accuracy (%) of sketch recognition on the sketch dataset. The FFD Augmentor is with 5 patch num, 11 max offset, and 30 augmented samples.

k-shot	model	EASY	FFD + EASY
1	ResNet12	35.38	**49.75**
	ResNet18	38.51	**54.27**
	ResNet20	30.48	**53.91**
	WideResNet	40.27	**55.17**
3	ResNet12	61.72	**71.38**
	ResNet18	55.03	**73.38**
	ResNet20	62.21	**75.23**
	WideResNet	59.22	**76.04**
5	ResNet12	72.98	**80.76**
	ResNet18	67.00	**81.77**
	ResNet20	69.49	**80.87**
	WideResNet	70.87	**84.38**

5   Conclusion

We address the task for oracle character recognition with a few labeled training samples in this study. We propose a new data augmentation tool for few-shot oracle recognition problems, FFD Augmentor, which is based on the free form deformation method commonly used the registration. FFD Augmentor generates a series of augmented images by random FFD on the original images for the classifier for training. Numerous experiments in three few-shot scenarios support the efficacy of our FFD Augmentor. Our generated training data are so efficient and informative that the deep model can be trained with generated data from scratch, without any additional large unlabeled dataset for pretraining. Our model has broad prospects in the field of written character recognition field. A wider range of applications will be explored in future studies.

References

1. Keightley, D.N.: Graphs, words, and meanings: Three reference works for shang oracle-bone studies, with an excursus on the religious role of the day or sun. Journal of the American Oriental Society **117** (1997) 507–524 Full publication date: Jul. - Sep., 1997.
2. 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. AIIPCC '19, New York, NY, USA, Association for Computing Machinery (2019)
3. Guo, J., Wang, C., Roman-Rangel, E., Chao, H., Rui, Y.: Building hierarchical representations for oracle character and sketch recognition. IEEE Transactions on Image Processing **25** (2015) 104–118
4. Lu, X., Cai, H., Lin, L.: Recognition of oracle radical based on the capsule network. CAAI transactions on intelligent systems **15** (2020) 243–254
5. 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. (2020)
6. Vinyals, O., Blundell, C., Lillicrap, T., Wierstra, D., et al.: Matching networks for one shot learning. Advances in neural information processing systems **29** (2016)
7. Snell, J., Swersky, K., Zemel, R.: Prototypical networks for few-shot learning. (2017)
8. Finn, C., Abbeel, P., Levine, S.: Model-agnostic meta-learning for fast adaptation of deep networks. In: International conference on machine learning, PMLR (2017) 1126–1135
9. Santoro, A., Bartunov, S., Botvinick, M., Wierstra, D., Lillicrap, T.: Meta-learning with memory-augmented neural networks. In Balcan, M.F., Weinberger, K.Q., eds.: Proceedings of The 33rd International Conference on Machine Learning. Volume 48 of Proceedings of Machine Learning Research., New York, New York, USA, PMLR (2016) 1842–1850
10. Oreshkin, B., López, P.R., Lacoste, A.: Tadam: Task dependent adaptive metric for improved few-shot learning. (2018)
11. Sung, F., Yang, Y., Zhang, L., Xiang, T., Torr, P.H., Hospedales, T.M.: Learning to compare: Relation network for few-shot learning. (2018)
12. Nichol, A., Achiam, J., Schulman, J.: On first-order meta-learning algorithms. arXiv preprint arXiv:1803.02999 (2018)
13. Rusu, A.A., Rao, D., Sygnowski, J., Vinyals, O., Pascanu, R., Osindero, S., Hadsell, R.: Meta-learning with latent embedding optimization. (2019)
14. Mishra, N., Rohaninejad, M., Chen, X., Abbeel, P.: A simple neural attentive meta-learner. (2018)
15. Chen, Z., Fu, Y., Chen, K., Jiang, Y.G.: Image block augmentation for one-shot learning. In: Proceedings of the AAAI Conference on Artificial Intelligence. Volume 33. (2019) 3379–3386
16. Antoniou, A., Storkey, A., Edwards, H.: Augmenting image classifiers using data augmentation generative adversarial networks. In: International Conference on Artificial Neural Networks, Springer (2018) 594–603
17. Ren, M., Ravi, S., Triantafillou, E., Snell, J., Swersky, K., Tenenbaum, J.B., Larochelle, H., Zemel, R.S.: Meta-learning for semi-supervised few-shot classification. (2018)

18. Liu, Y., Lee, J., Park, M., Kim, S., Yang, E., Hwang, S., Yang, Y.: Learning to propagate labels: Transductive propagation network for few-shot learning. (2019)
19. Li, X., Sun, Q., Liu, Y., Zhou, Q., Zheng, S., Chua, T.S., Schiele, B.: Learning to self-train for semi-supervised few-shot classification. (2019)
20. Wang, Y., Xu, C., Liu, C., Zhang, L., Fu, Y.: Instance credibility inference for few-shot learning. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). (2020)
21. Wang, Y., Zhang, L., Yao, Y., Fu, Y.: How to trust unlabeled data? instance credibility inference for few-shot learning. IEEE Transactions on Pattern Analysis and Machine Intelligence (2021)
22. Yalniz, I.Z., Jégou, H., Chen, K., Paluri, M., Mahajan, D.: Billion-scale semi-supervised learning for image classification. arXiv preprint arXiv:1905.00546 (2019)
23. Springenberg, J.T.: Unsupervised and semi-supervised learning with categorical generative adversarial networks. arXiv preprint arXiv:1511.06390 (2015)
24. Wang, M., Deng, W., Liu, C.L.: Unsupervised structure-texture separation network for oracle character recognition. IEEE Transactions on Image Processing **31** (2022) 3137–3150
25. Rueckert, D., Sonoda, L., Hayes, C., Hill, D., Leach, M., Hawkes, D.: Nonrigid registration using free-form deformations: application to breast mr images. IEEE Transactions on Medical Imaging **18** (1999) 712–721
26. Keszei, A.P., Berkels, B., Deserno, T.M.: Survey of non-rigid registration tools in medicine. Journal of digital imaging **30** (2017) 102–116
27. Bendou, Y., Hu, Y., Lafargue, R., Lioi, G., Pasdeloup, B., Pateux, S., Gripon, V.: Easy: Ensemble augmented-shot y-shaped learning: State-of-the-art few-shot classification with simple ingredients. arXiv preprint arXiv:2201.09699 (2022)
28. Li, B., Dai, Q., Gao, F., Zhu, W., Li, Q., Liu, Y.: HWOBC-a handwriting oracle bone character recognition database. Journal of Physics: Conference Series **1651** (2020) 012050
29. Eitz, M., Hays, J., Alexa, M.: How do humans sketch objects? ACM Trans. Graph. **31** (2012)
30. Zhou, X.L., Hua, X.C., Li, F.: A method of jia gu wen recognition based on a two-level classification. In: Proceedings of 3rd International Conference on Document Analysis and Recognition. Volume 2., IEEE (1995) 833–836
31. Li, F., Woo, P.y.: The coding principle and method for automatic recognition of jia gu wen characters. International Journal of Human-Computer Studies **53** (2000) 289–299
32. Qingsheng, L.: Recognition of inscriptions on bones or tortoise shells based on graph isomorphism. Computer Engineering and Applications (2011)
33. Yang, Z., Wang, Q., He, X., Liu, Y., Yang, F., Yin, Z., Yao, C.: Accurate oracle classification based on deep convolutional neural network. In: 2018 IEEE 18th International Conference on Communication Technology (ICCT). (2018) 1188–1191
34. Zhang, Y.K., Zhang, H., Liu, Y.G., Yang, Q., Liu, C.L.: Oracle character recognition by nearest neighbor classification with deep metric learning. In: 2019 International Conference on Document Analysis and Recognition (ICDAR). (2019) 309–314
35. Sung, F., Yang, Y., Zhang, L., Xiang, T., Torr, P.H., Hospedales, T.M.: Learning to compare: Relation network for few-shot learning. In: Proceedings of the IEEE conference on computer vision and pattern recognition. (2018) 1199–1208

36. Qiao, S., Liu, C., Shen, W., Yuille, A.L.: Few-shot image recognition by predicting parameters from activations. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2018) 7229–7238
37. Chen, D., Chen, Y., Li, Y., Mao, F., He, Y., Xue, H.: Self-supervised learning for few-shot image classification. In: ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE (2021) 1745–1749
38. Rusu, A.A., Rao, D., Sygnowski, J., Vinyals, O., Pascanu, R., Osindero, S., Hadsell, R.: Meta-learning with latent embedding optimization. arXiv preprint arXiv:1807.05960 (2018)
39. Chen, H., Li, H., Li, Y., Chen, C.: Multi-scale adaptive task attention network for few-shot learning. CoRR **abs/2011.14479** (2020)
40. Qi, H., Brown, M., Lowe, D.G.: Low-shot learning with imprinted weights. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2018)
41. Shyam, P., Gupta, S., Dukkipati, A.: Attentive recurrent comparators. In Precup, D., Teh, Y.W., eds.: Proceedings of the 34th International Conference on Machine Learning. Volume 70 of Proceedings of Machine Learning Research., PMLR (2017) 3173–3181
42. DeVries, T., Taylor, G.W.: Improved regularization of convolutional neural networks with cutout. arXiv preprint arXiv:1708.04552 (2017)
43. Cubuk, E.D., Zoph, B., Mane, D., Vasudevan, V., Le, Q.V.: Autoaugment: Learning augmentation strategies from data. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. (2019) 113–123
44. Yue, X., Li, H., Fujikawa, Y., Meng, L.: Dynamic dataset augmentation for deep learning-based oracle bone inscriptions recognition. J. Comput. Cult. Herit. (2022) Just Accepted.
45. Woods, R.P., Grafton, S.T., Holmes, C.J., Cherry, S.R., Mazziotta, J.C.: Automated image registration: I. general methods and intrasubject, intramodality validation. Journal of computer assisted tomography **22** (1998) 139–152
46. Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Diffeomorphic demons: Efficient non-parametric image registration. NeuroImage **45** (2009) S61–S72
47. Oliveira, F.P., Tavares, J.M.R.: Medical image registration: a review. Computer Methods in Biomechanics and Biomedical Engineering **17** (2014) 73–93 PMID: 22435355.
48. Ziyi, G., Zihan, Z., Bingshuai, L., Longquan, L., Qingju, J., Chenxi, H., Jianwei, Z.: An improved neural network model based on inception-v3 for oracle bone inscription character recognition. Scientific Programming (2022)
49. Jiang, J., Luk, W., Rueckert, D.: Fpga-based computation of free-form deformations in medical image registration. In: Proceedings. 2003 IEEE International Conference on Field-Programmable Technology (FPT)(IEEE Cat. No. 03EX798), IEEE (2003) 234–241
50. Rohlfing, T., Maurer, C.R.: Nonrigid image registration in shared-memory multiprocessor environments with application to brains, breasts, and bees. IEEE transactions on information technology in biomedicine **7** (2003) 16–25
51. Gribbon, K., Bailey, D.: A novel approach to real-time bilinear interpolation. In: Proceedings. DELTA 2004. Second IEEE International Workshop on Electronic Design, Test and Applications. (2004) 126–131
52. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. (2016) 770–778

53. Zagoruyko, S., Komodakis, N.: Wide residual networks. CoRR **abs/1605.07146** (2016)

54. Mangla, P., Kumari, N., Sinha, A., Singh, M., Krishnamurthy, B., Balasubramanian, V.N.: Charting the right manifold: Manifold mixup for few-shot learning. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision. (2020) 2218–2227

55. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S.: Pytorch: An imperative style, high-performance deep learning library. In: Advances in Neural Information Processing Systems 32. Curran Associates, Inc. (2019) 8024–8035