FaRE: A Feature-aware Radical Encoding Strategy for Zero-shot Chinese Character Recognition

Hongjian Zhan[1,2], Yangfu Li[1], Yu-jie Xiong[†,3], and Yue Lu[1]

[1] Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University Shanghai 200241, China
[2] Chongqing Key Laboratory of Precision Optics, Chongqing Institute of East China Normal University, Chongqing 401120, China
[3] School of Electronic and Electrical Engineering, Shanghai University of Engineering Science Shanghai 201620, China

Abstract. Due to the complexity of glyphs and the vast vocabulary, zero-shot Chinese character recognition (ZSCCR) remains a prominent research topic. A mainstream approach involves radical-based character decomposition. However, existing methods typically employ random encoding for each radical post-decomposition, leading to potential topology distortions in the radical encoding and glyph spaces. To address these issues, we propose a novel Feature-aware Radical Encoding (FaRE) strategy that incorporates visual feature clues into radical encodings to generate feature-aware representations. Initially, we create radical images by rendering TTF files and then apply a pre-trained feature extractor to obtain the feature representation of each radical. Finally, projection and binarization operations are performed to produce compact and efficient radical encodings. Extensive experiments on the public benchmark ICDAR2013 demonstrate that the proposed FaRE significantly enhances the state-of-the-art ZSCCR performance. Additionally, abundant ablation studies are conducted to validate the effectiveness of the proposed FaRE.

Keywords: Chinese Character Recognition · Feature-aware Radical Encoding Strategy · Zero-shot

1 Introduction

Character recognition as a foundation task of pattern recognition has been studied for decades, and extends many research areas [12, 16–18, 20, 23–27, 33, 34]. Among them, because of complex glyphs and massive vocabulary, Chinese character recognition has been one of the most difficult directions, especially zero-shot Chinese character recognition (ZSCCR).

Existing ZSCCR methods can be approximately divided into three categories: stroke-based methods [3, 9, 19], radical-based methods [2, 8, 11, 13, 21, 30–32], and

† Corresponding Author

Glyph-based methods [1,7]. Specifically, stroke-based methods decompose Chinese characters into numerous individual strokes and use the sequence of encoded stroke information to represent each character. In contrast, glyph-based methods treat Chinese characters as indivisible units and achieve recognition through similarity-based matching between character instances. Different from the former two methods, radical-based methods fully consider the characteristics of Chinese characters, decompose Chinese characters into a series of special components, i.e., radicals, through specific rules, and employ the sequence of encoded radicals to represent different characters, the whole pipeline is shown in Fig. 1. Due to their flexibility and scalability, radical-based methods have gradually become the most promising method for zero-shot CCR.

In early radical-based methods, Chinese characters are represented by counts or distributions of radicals [21]. These methods ignore the structural information of characters, resulting in unsatisfactory recognition performance. In order to solve this problem, numerous radical-based works have been proposed in recent years, which decompose Chinese characters into Ideographic Description Sequences (IDS) composed of structures and radicals and utilize the encoded IDS to represent characters [2,30–32]. In these methods, the recognizer is employed to predict the code of the IDS and obtain recognition results by matching the predicted code with the codebook.

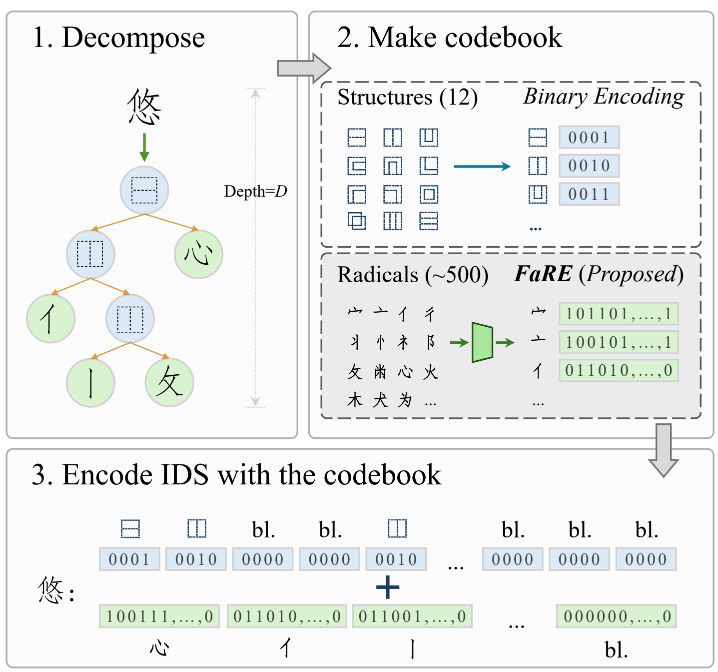

Fig. 1: Illustration of the decomposition and representations with FaRE in radical-based methods, where blue and green denote the structure and radicals.

However, existing radical-based methods rarely consider the visual information of radicals during encoding. They commonly apply a random multi-hot encoding strategy to encode radicals and then derive character-level representations. As a result, the similar radicals may have distinct representations in the encoding space, which further leads to distortion of the character-level representations composed of these radicals, as shown in Fig. 2. To address this issue, we propose a Feature-aware Radical Encoding strategy termed FaRE. Specifically, it first renders the associated radical images with True Type Fonts, then employs a pre-trained convolutional neural network to extract the visual features, and finally performs a dimension reduction together with a binarization to derive the radical codes containing the visual clues. FaRE can effectively maintain the topology of the encoding space, thereby constructing a more reasonable character-level representation and improving the performance of recognizers. Solid experimental results on public handwriting benchmark ICDAR2013 show that the proposed FaRE achieves noticeable improvements then the random multi-hot encoding strategy [32]. The main contributions of this paper are summarized as follows:

- We propose a noval radical encoding strategy, i.e., *FaRE*, for zero-shot Chinese character recognition methods. Different from existing random multi-hot strategies, FaRE introduces the visual features of the radicals into the character encodings, thereby improving the recognition performances.
- Based on FaRE, we present a zero-shot Chinese character recognition method that achieves state-of-the-art results on wellknown Chinese handwritten benchmark ICDAR2013.
- Extensive experimental results demonstrate the effectiveness of the proposed FaRE. Besides, a series of detailed ablation studies are provided to explore the impact of key components in FaRE, which provides some valuable insights for the design of zero-shot Chinese character recognition methods in the future.

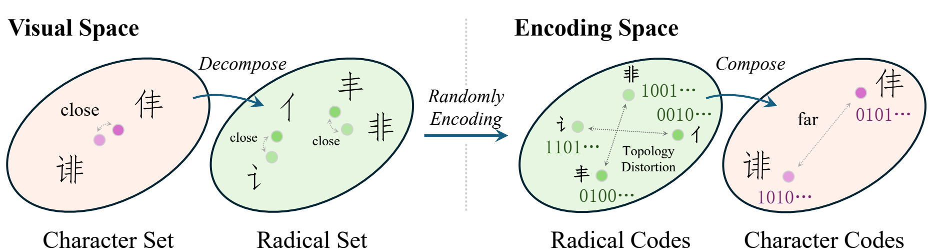

Fig. 2: Diagram illustrating the topology distortion caused by the random encoding strategy, where the light purple denotes the set of Chinese characters and the light green denotes the set of radicals.

2 Proposed Method

2.1 Motivation

Most of the prevailing radical-based methods for zero-shot Chinese character recognition [2,8,11,13,21,30–32] do not pay enough attention to the visual features in radicals. A natural solution is to represent radicals with random sparse multi-hot codes [32], which guarantees differentiation among distinct radicals. However, this encoding strategy overlooks the inherent visual clues of radicals, consequently resulting in a topology distortion of the encoding space, as illustrated in Fig. 2. More specifically, the features extracted from the characters composed of similar radicals are commonly relevant. However, the representations of these similar characters derived by the random multi-hot encoding strategy are likely orthogonal. As a result, a prediction head is forced to predict almost orthogonal representations through the relevant features, which increases the difficulty of recognition and leads to suboptimal results. To address this issue, we propose *FaRE*, a feature-aware encoding strategy that takes into account both the differentiation and the visual features among radicals in the encoding space, which reduces the difficulty for the prediction head to make correct decisions, thereby improving the recognition performance.

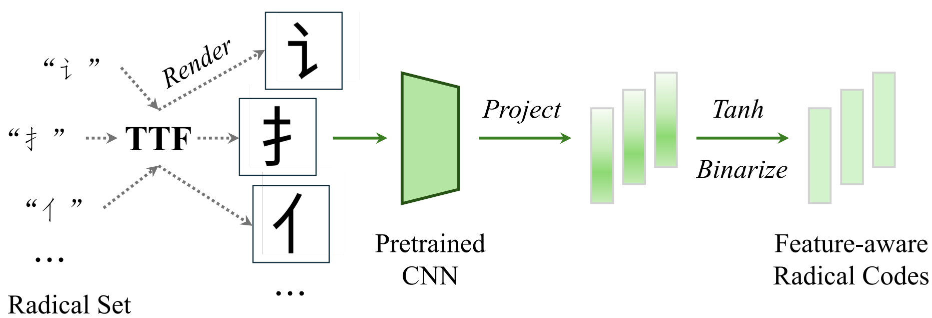

Fig. 3: Illustration of the main pipeline of feature-aware radical encoding.

2.2 Feature-aware Radical Embedding

The main pipeline of the proposed FaRE is shown in Fig. 3. Specifically, the first step in FaRE is to render the radical images with TTF file. Based on the rendered radical images, FaRE can easily gain the associated visual features through a pretrained neural network. Subsequently, taking efficiency into account, the visual feature embeddings are projected into a more compact space with a dimension reduction aglorithm. Finally, the feature-aware radical codes are derived from a binarization of the projections. To represent sufficient numbers of Chinese characters, it is inevitable to deal with some alien-shape radicals (about 18%

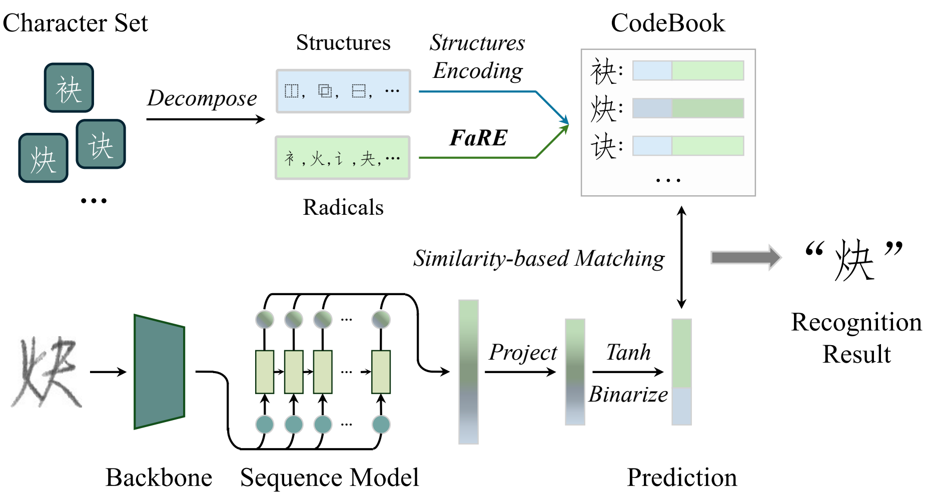

Fig. 4: The overall architecture for character recognition using FaRE.

of the total) that are not described in any type of TTF. As for these radicals, we follow the random encoding strategy from existing work [32]. The details are explianed in Algorithm 1. Besides, to avoid the collision of radical codes, we perform a check and replacements after the binarization.

2.3 Character Recognition with FaRE

Fig. 4 presents the pipeline for character recognition with FaRE. Similar to previous radical-based zero-shot Chinese character recognition methods [32], after decomposition, structure encoding, and radical encoding, i.e., FaRE, we concatenate the structure codes with the feature-aware radical codes for character-level representation and build a codebook $\mathbf{E} = [\mathbf{e}_1, ..., \mathbf{e}_N,] \in \mathbb{R}^{N \times C}$. The codebook is employed with a recognition model to achieve zero-shot recognition of Chinese characters. Specifically, given a character image $\mathbf{X}$ as input, the recognizer outputs the character representation $\mathbf{e}$:

$$\mathbf{e} = \mathrm{Binarize}(\mathrm{Tanh}(\mathrm{model}))(\mathbf{X}) \in \mathbb{R}^{C}. \tag{1}$$

Subsequently, a similarity-based matching is performed with the codebook for recognition:

$$\hat{y}_i = \arg\max_i(\mathbf{e} \cdot \mathbf{E}_i^T), \tag{2}$$

where $\hat{y}_i$ is a recognition result. Noteworthily, to approximate the gradients, a straightforward estimator [4] is applied during the training phase. Besides, the cross-entropy is applied as the loss function, which can be formulated as follows:

$$\mathcal{L} = -\sum_i \log p(y \mid \hat{y}_i), \tag{3}$$

where y denotes the ground-truth label.

Algorithm 1 Workflow of feature-aware Radical Encoding

Input: Set of radicals $\mathcal{R}$, TrueType Font file `ttf`
Initialization: Set of feature-aware radical codes $\mathcal{E}_{sa}$, Set of random radical codes $\mathcal{E}_r$
Parse ttf to extract character metadata $\mathcal{T} \leftarrow f_{parse}(\texttt{ttf})$
for radical $r \in \mathcal{R}$ **do**
    if $r \in \mathcal{T}$ **then**
        Get feature-aware radical code $e_{sa} \leftarrow FaRE(r)$
        Update feature-aware code set $\mathcal{E}_{sa} \leftarrow \mathcal{E}_{sa} \cup \{e_{sa}\}$
    else
        # Random encode for alien-shape radicals
        Generate a random multi-hot code $e_r \leftarrow \{0,1\}^D$
        # Check & Random replacement
        while $e_r \in \mathcal{E}_{sa} \cup \mathcal{E}_r$ **do**
            Replace by a random code $e_r \leftarrow \{0,1\}^D$
        end while
        Update random code set $\mathcal{E}_r \leftarrow \mathcal{E}_r \cup \{e_r\}$
    end if
end for
Obtain the full set of radical codes $\mathcal{E} \leftarrow \mathcal{E}_{sa} \cup \mathcal{E}_r$
Output: Set of radical codes $\mathcal{E}$

3 Experimental Setup

3.1 Datasets

We separately utilize the isolated character data from two popular handwritten Chinese datasets, i.e., **CASIA-HWDB** [10] and **ICDAR2013** [28], for training and testing, which covers 3755 high-frequency characters of Chinese. Specifically, to evaluate the zero-shot recognition performance, we separately select the first k characters from HWDB as the training set, where k ranges in $\{500, 1000, 1500, 2000, 2755\}$, and the test set is composed of the last 1000 character samples from ICDAR2013 to avoid the overlap of the characters between training and testing samples. As for the full-supervised recognition performance, the models are trained on all the 3755 characters from HWDB and evaluated with the same testing set. The details of the datasets are shown in Table 1.

3.2 Network Architecture

In this work, we employ ResNet34 [5] as the backbone to extract the visual features and utilize two Bi-LSTM [6] layers for sequence modeling. The model is trained with the Adadelta optimizer [29] with an initial learning rate of 0.1, a rho of 0.95, and a weight decay of 5×10^{-4}. The batch size is set to 256. All experiments are conducted on one NVIDIA RTX4090 GPU.

Table 1: Detials of the training and testing set.

	Training sets						Test set
#of Characters	500	1000	1500	2000	2755	3755	1000
#of Samples ($\times 10^6$)	0.36	0.71	1.07	1.43	1.97	2.68	0.06
#of Writers	1019	1019	1020	1020	1020	1020	60

Table 2: Comparison of the zero-shot and full-supervised recognition performance between FaRE and state-of-the-art methods.

Methods	Zero-shot Accuracy/%					Full-supervised
	500	1000	1500	2000	2755	Accuracy/%
CRNN [18]		-	-	-	-	95.78
DenseRAN [22]	1.70	8.44	14.71	19.51	30.68	96.66
HDE [2]	4.90	12.77	19.25	25.13	33.49	<u>97.14</u>
SLD [3]	5.60	13.85	22.88	25.73	37.91	96.73
CUE [13]	**7.43**	15.75	24.01	27.04	40.55	96.96
SideNet [8]	5.10	16.20	33.80	44.10	50.30	-
HierCode† [32]	5.85	<u>20.14</u>	<u>35.25</u>	<u>45.84</u>	<u>55.85</u>	96.44
FaRE	<u>7.21</u>	**21.78**	**36.58**	**47.33**	**57.17**	**97.82**
Δ	-0.22	+1.64	+1.33	+1.49	+1.32	+0.68

† The results is reproduced by us.

3.3 Implementation Details

Inheriting from the previous work [32], we apply four bits to represent structures. The length of FaRE is defaultly set to 64, and the depth of the binary tree in decomposition is set to 6, finally resulting in a code with a length of 1212 for each character. Since the setting is slightly different from [32], for fair comparison, we produce 1212-dimensional character representations for HierCode in our experiments. All of the input character images are resized to 96×96. Furthermore, under the zero-shot evaluations, we also keep the same setting as [32]. Specifically, during the training phase, the similarity is calculated only with the representations of the characters appearing in the training set, while during the inference phase, the final classification results are obtained by matching the model predictions with the all characters comprising both the training and test sets.

4 Results

4.1 Comparison with State-of-the-Art Methods

To evaluate the effectiveness of FaRE, we conduct a comprehensive comparison between it and a classical recognition method, i.e., CRNN [18], together with five

Table 3: Alabtion study on the encoding dimension.

Encoding Dimension	Accuracy/%	
	Zero-Shot (2755)	Full-supervised
32	$56.42_{-0.75}$	$96.59_{-0.53}$
64	57.17	97.82
96	$56.91_{-0.26}$	$97.80_{-0.02}$
128	$56.35_{-0.82}$	$97.81_{-0.01}$

Table 4: Alabtion study on feature extractor in FaRE.

Feature Extractor	Accuracy/%	
	Zero-Shot (2755)	Full-supervised
ResNet18	$55.95_{-1.22}$	$97.23_{-0.59}$
ResNet34	57.17	97.82
ResNet50	$57.15_{-0.02}$	$97.78_{-0.04}$

advanced zero-shot Chinese text recognition methods, i.e., DenseRAN [22], HDE [2], SLD [3], CUE [13], SideNet [8], and even a state-of-the-art zero-shot Chinese text recognition method based on random sparse multi-hot encoding strategy, i.e., HierCode [32], in both zero-shot and full-supervised cases. Recognition performance is measured by character accuracy. As illustrated in Table 2, our method achieves a noticeable improvement when compared to the state-of-the-art zero-shot methods in most cases. Partically, under zero-shot scenarios, FaRE suprasses the SOTA method, i.e., HierCode, with absolute accuracy increments of 1.36%, 1.64%, 1.33%, 1.49%, and 1.32% at k in $\{500, 1000, 1500, 2000, 2755\}$, respectively, which demonstrates the effectiveness of the proposed feature-aware radical encoding strategy. As for the full-supervised case, FaRE can also boost the recognition performance by 2.04% and 0.68% when compared to the classic recognition method CRNN and the advanced recognition method HDE, respectively.

4.2 Ablation Study

This section conducts an ablation analysis to analyze the crucial components in FaRE, i.e., encoding dimension, feature extractor, and projection method. For efficiency, only the results in the zero-shot scenario with $k = 2755$ and the full supervised scenario are reported in Tables III–V, while the analogous tendencies are also observed across broader settings (details omitted). The gray denotes the default setting.

Encoding Dimension Impact As summarized in Table III, 64-dimensional encoding yields optimal performance in both zero-shot and fully supervised sce-

Table 5: Alabtion study on projection aglorithms.

Projection Aglorithm	Accuracy/%	
	Zero-Shot (2755)	Full-supervised
PCA	$56.82_{-0.35}$	$97.21_{-0.61}$
t-SNE	$56.61_{-0.56}$	$97.03_{-0.79}$
UMAP	57.17	97.82

narios. Specifically, the lower encoding dimension, like 32, inadequately captures radical feature complexity, compromising performance, while the higher encoding dimensions, e.g., 96 or 128, may lead to overfitting, detrimental in zero-shot recognition scenarios.

Feature Extractor Efficiency As evidenced in Table IV, ResNet34 emerges as the most efficacious backbone for extracting visual features from rendered radical images, striking an advantageous balance between computational efficiency and benefits for recognition performance.

Projection Algorithm Comparison We consider three prominent projection techniques, i.e., PCA, t-SNE [14], and UMAP [15], for dimension reduction of radical features. As shown in Table V, UMAP achieves the best result, which is attributed to its balanced preservation of local and global topological features in the visual space, thus generating both discriminative and feature-aware embeddings. Conversely, PCA's linearity limits topology preservation, and t-SNE pays too much attention to the local while sacrificing the global topology, resulting in ambiguous representation.

Fig. 5: Comparisions of several hard samples under both zero-shot and full-supervised scenarios. The correct and incorrect recognitions are separately marked in blur and red.

4.3 Case Study

There are numerous visually similar Chinese characters, particularly in handwritten styles. As illustrated in Fig. 5, we compared a series of challenging samples in both zero-shot and fully-supervised scenarios. Thanks to considering the visual features during radical encoding, our method produces more robust character-level representations, resulting in improved performance on these confounding samples.

5 Conclusion and Future Work

In this paper, we propose a feature-aware radical encoding strategy (FaRE) to enhance radical-based zero-shot Chinese character recognition methods. FaRE introduces visual clues into the encoding space, deriving more representative radical and character-level encodings for zero-shot recognition models. This prompts the prediction heads to make more reasonable decisions, thereby improving recognition performance. Experimental results demonstrate that FaRE achieves significant improvements compared to state-of-the-art zero-shot Chinese character recognition methods in both zero-shot and fully-supervised scenarios. Future work will focus on developing a unified encoding strategy to handle alien-shape radicals that cannot be rendered by TTF.

6 Acknowledgments

This work is jointly supported by the National Natural Science Foundation of China (No. 62176091, No. 62006150), Science and Technology Commission of Shanghai Municipality (No. 21DZ2203100), and Natural Science Foundation of Chongqing, China, under Grant CSTB2024NSCQ-MSX0877.

References

1. Ao, X., Zhang, X.Y., Liu, C.L.: Cross-modal prototype learning for zero-shot handwritten character recognition. Pattern Recognition **131**, 108859 (2022) 2
2. Cao, Z., Lu, J., Cui, S., Zhang, C.: Zero-shot handwritten chinese character recognition with hierarchical decomposition embedding. Pattern Recognition **107**, 107488 (2020) 1, 2, 4, 7, 8
3. Chen, J., Li, B., Xue, X.: Zero-shot chinese character recognition with stroke-level decomposition. In: Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21. pp. 615–621 (8 2021) 1, 7, 8
4. Courbariaux, M., Hubara, I., Soudry, D., El-Yaniv, R., Bengio, Y.: Binarized neural networks: Training deep neural networks with weights and activations constrained to+ 1 or-1. arXiv preprint arXiv:1602.02830 (2016) 5
5. 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. pp. 770–778 (2016) 6

6. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural computation **9**(8), 1735–1780 (1997) 6
7. Huang, G., Luo, X., Wang, S., Gu, T., Su, K.: Hippocampus-heuristic character recognition network for zero-shot learning in chinese character recognition. Pattern Recognition **130**, 108818 (2022) 2
8. Li, Z., Huang, Y., Peng, D., He, M., Jin, L.: Sidenet: Learning representations from interactive side information for zero-shot chinese character recognition. Pattern Recognition **148**, 110208 (2024) 1, 4, 7, 8
9. Liu, C.L., Kim, I.J., Kim, J.H.: Model-based stroke extraction and matching for handwritten chinese character recognition. Pattern Recognition **34**(12), 2339–2352 (2001) 1
10. Liu, C.L., Yin, F., Wang, D.H., Wang, Q.F.: Casia online and offline chinese handwriting databases. In: 2011 international conference on document analysis and recognition. pp. 37–41. IEEE (2011) 6
11. Liu, C.L., Yin, F., Wang, D.H., Wang, Q.F.: Online and offline handwritten chinese character recognition: benchmarking on new databases. Pattern Recognition **46**(1), 155–162 (2013) 1, 4
12. Liu, Y., Jin, L., Zhang, S., Luo, C., Zhang, S.: Curved scene text detection via transverse and longitudinal sequence connection. Pattern Recognition **90**, 337–345 (2019) 1
13. Luo, G.F., Wang, D.H., Du, X., Yin, H.Y., Zhang, X.Y., Zhu, S.: Self-information of radicals: A new clue for zero-shot chinese character recognition. Pattern Recognition **140**, 109598 (2023) 1, 4, 7, 8
14. Van der Maaten, L., Hinton, G.: Visualizing data using t-sne. Journal of machine learning research **9**(11) (2008) 9
15. McInnes, L., Healy, J., Melville, J.: Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426 (2018) 9
16. Peng, D., Jin, L., Liu, Y., Luo, C., Lai, S.: Pagenet: Towards end-to-end weakly supervised page-level handwritten chinese text recognition. International Journal of Computer Vision **130**(11), 2623–2645 (2022) 1
17. Peng, D., Jin, L., Ma, W., Xie, C., Zhang, H., Zhu, S., Li, J.: Recognition of handwritten chinese text by segmentation: a segment-annotation-free approach. IEEE Transactions on Multimedia (2022) 1
18. Shi, B., Bai, X., Yao, C.: An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. IEEE transactions on pattern analysis and machine intelligence **39**(11), 2298–2304 (2016) 1, 7
19. Su, Y.M., Wang, J.F.: A novel stroke extraction method for chinese characters using gabor filters. Pattern Recognition **36**(3), 635–647 (2003) 1
20. Wang, Q., Lu, Y.: A sequence labeling convolutional network and its application to handwritten string recognition. In: Proceedings of the 26th International Joint Conference on Artificial Intelligence. pp. 2950–2956 (2017) 1
21. Wang, T., Xie, Z., Li, Z., Jin, L., Chen, X.: Radical aggregation network for few-shot offline handwritten chinese character recognition. Pattern Recognition Letters **125**, 821–827 (2019) 1, 2, 4
22. Wang, W., Zhang, J., Du, J., Wang, Z.R., Zhu, Y.: Denseran for offline handwritten chinese character recognition. In: 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR). pp. 104–109. IEEE (2018) 7, 8
23. Wang, Z., Xie, H., Wang, Y., Xu, J., Zhang, B., Zhang, Y.: Symmetrical linguistic feature distillation with clip for scene text recognition. In: Proceedings of the 31st ACM International Conference on Multimedia. pp. 509–518 (2023) 1

24. Wei, X., Lu, S., Lu, Y.: Compact mqdf classifiers using sparse coding for handwritten chinese character recognition. Pattern Recognition **76**, 679–690 (2018) 1
25. Wu, X., Chen, Q., You, J., Xiao, Y.: Unconstrained offline handwritten word recognition by position embedding integrated resnets model. IEEE Signal Processing Letters **26**(4), 597–601 (2019) 1
26. Wu, Y.C., Yin, F., Chen, Z., Liu, C.L.: Handwritten chinese text recognition using separable multi-dimensional recurrent neural network. In: 2017 14th IAPR international conference on document analysis and recognition (ICDAR). vol. 1, pp. 79–84. IEEE (2017) 1
27. Xie, Z., Sun, Z., Jin, L., Ni, H., Lyons, T.: Learning spatial-semantic context with fully convolutional recurrent network for online handwritten chinese text recognition. IEEE transactions on pattern analysis and machine intelligence **40**(8), 1903–1917 (2017) 1
28. Yin, F., Wang, Q.F., Zhang, X.Y., Liu, C.L.: Icdar 2013 chinese handwriting recognition competition. In: 2013 12th international conference on document analysis and recognition. pp. 1464–1470. IEEE (2013) 6
29. Zeiler, M.D.: Adadelta: an adaptive learning rate method. arXiv preprint arXiv:1212.5701 (2012) 6
30. Zhang, J., Du, J., Dai, L.: Radical analysis network for learning hierarchies of chinese characters. Pattern Recognition **103**, 107305 (2020) 1, 2, 4
31. Zhang, J., Zhu, Y., Du, J., Dai, L.: Radical analysis network for zero-shot learning in printed chinese character recognition. In: 2018 IEEE International Conference on Multimedia and Expo (ICME). pp. 1–6. IEEE (2018) 1, 2, 4
32. Zhang, Y., Zhu, Y., Peng, D., Zhang, P., Yang, Z., Yang, Z., Yao, C., Jin, L.: Hiercode: A lightweight hierarchical codebook for zero-shot chinese text recognition. arXiv preprint arXiv:2403.13761 (2024) 1, 2, 3, 4, 5, 7, 8
33. Zhong, D., Lyu, S., Shivakumara, P., Pal, U., Lu, Y.: Text proposals with location-awareness-attention network for arbitrarily shaped scene text detection and recognition. Expert Systems with Applications **205**, 117564 (2022) 1
34. Zhong, D., Lyu, S., Shivakumara, P., Yin, B., Wu, J., Pal, U., Lu, Y.: Sgbanet: Semantic gan and balanced attention network for arbitrarily oriented scene text recognition. In: European conference on computer vision. pp. 464–480. Springer (2022) 1