Seal-Robust KCR: A Robust Kuzushiji Character Recognition Framework under Seal Interference

Supplementary Material

A. Additional Dataset Details

The Japanese historical document images used in this work were sourced from the Center for Open Data in the Humanities (CODH) [1], whose original materials are preserved and maintained by the National Institute of Japanese Literature (NIJL). This resource is among the largest publicly available Kuzushiji datasets, containing over 1.08 million annotated character instances across 4,328 character categories from 6,151 page images of Japanese historical books.

In addition to document images, the dataset provides character-level annotations, including bounding box coordinates (i.e., the x- and y-coordinates, width, and height) and corresponding Unicode code points, obtained through expert-assisted transcription rather than automatic OCR. From this resource, we selected 13 representative books as data sources. From the selected books, page images that did not contain any Kuzushiji character instances were excluded, resulting in a final dataset of 1,000 document images containing 208,482 Kuzushiji characters, with an average of 208.5 characters per page. The bibliographic information and dataset statistics are summarized in Table 1.

B. Document Restoration Parameter Analysis

For document restoration, we employ Telea's fast marching-based image inpainting method [2] with a fixed inpainting radius of $\rho = 3$. By default, mask refinement is performed using a 3×3 dilation kernel with a single iteration. We further conduct a sensitivity analysis of the key hyperparameters τ_r and (τ_{rg}, τ_{rb}) for Kuzushiji document restoration, and the results are summarized in Table 2.

The first row reports a no-restoration baseline, where document restoration is disabled and the evaluation metrics are computed between the original document images and their corresponding synthetic counterparts. Overall, configurations with $\tau_r = 90$ generally achieve better restoration quality than those with $\tau_r = 80$. Furthermore, when $(\tau_{rg}, \tau_{rb}) = (1.3, 1.3)$, PSNR reaches its highest values of 34.09 dB and 34.13 dB on the validation and test sets, respectively. The corresponding SSIM scores of 0.9757 and 0.9750 are the second highest, only marginally lower than the best scores of 0.9763 and 0.9753.

Figure 1. Additional qualitative comparisons between the original historical document images (top row) and the corresponding restored results obtained from synthetic seal-interfered document images (bottom row) using $\tau_r = 90$ and $(\tau_{rg}, \tau_{rb}) = (1.3, 1.3)$.

Considering the overall performance across all metrics, we select $\tau_r = 90$ and $\tau_{rg} = \tau_{rb} = 1.3$ as the final hyperparameter configuration.

C. Character Ordering Parameter Analysis

To analyze the impact of key hyperparameters on character ordering, we conducted sensitivity experiments on the column grouping threshold λ and the column-center update strategy using character classification results produced by the proposed method on the synthetic test set. The experimental results are shown in Table 3.

Overall, the mean and median update strategies yielded highly similar performance, with CER differences of less than 0.2% across all parameter configurations, indicating that the proposed ordering algorithm is robust to the choice of column-center update strategy.

As λ increased from 0.5 to 0.8, the CER consistently decreased, indicating that a larger value of column grouping threshold facilitates more accurate aggregation of characters within the same text column, thereby improving ordering accuracy. Both update strategies achieved their best performance at $\lambda = 0.8$, yielding a CER of 13.67%. When

Table 1. Statistics of the dataset constructed from 13 selected Japanese historical books. Pages denotes the number of annotated page images, and Chars/Page denotes the average number of annotated Kuzushiji characters per page.

Index	NIJL ID	Book Title	Pages	Chars.	Chars/Page
1	100241706	Usonarubeshi (虚南留別志)	67	8,565	127.8
2	100249376	Gozenkashi Hiden-shou (御前菓子秘伝抄)	104	11,841	113.9
3	100249416	Mochigashi Sokuseki Teseishuu (餅菓子即席手製集)	58	7,967	137.4
4	100249476	Meshi Hyakuchin Den (飯百珍伝)	46	7,842	170.5
5	200006663	Diguchi (ぢぐち)	8	121	15.1
6	200015843	Nippon Eitaigura (日本永代蔵)	180	50,251	279.2
7	200017458	Soga Monogatari (曾我物語)	78	29,641	380.0
8	200020019	Chikusai (竹斎)	146	33,228	227.6
9	200021086	Isoho Monogatari (伊曾保物語)	60	15,410	256.8
10	200021763	Zenbu Ryouri-shou (膳部料理抄)	94	11,437	121.7
11	200021802	Ryouri Monogatari (料理物語)	105	19,609	186.8
12	200021869	Ryourikata Kokoroenokoto (料理方心得之事)	30	3,012	100.4
13	200022050	Ryouri Hiden-shou (料理秘伝抄)	24	9,558	398.3
Total	–	–	1,000	208,482	208.5

Table 2. Parameter sensitivity analysis of τ_r and (τ_{rg}, τ_{rb}) for Kuzushiji document restoration on the constructed dataset.

τ_r	(τ_{rg}, τ_{rb})	Validation Set		Test Set	
		PSNR	SSIM	PSNR	SSIM
–	–	29.15	0.9655	28.71	0.9639
80	(1.2, 1.2)	29.76	0.9470	29.61	0.9465
80	(1.3, 1.3)	33.64	0.9736	33.73	0.9731
80	(1.4, 1.4)	33.87	0.9756	33.77	0.9745
80	(1.5, 1.5)	31.97	0.9717	31.68	0.9706
90	(1.2, 1.2)	30.37	0.9522	30.19	0.9619
90	(1.3, 1.3)	**34.09**	0.9757	**34.13**	0.9750
90	(1.4, 1.4)	34.05	**0.9763**	33.94	**0.9753**
90	(1.5, 1.5)	32.03	0.9721	31.74	0.9710

Table 3. Sensitivity analysis of the column grouping threshold λ and column-center update strategy in the proposed method.

λ	Column-center update strategy	
	CER@Mean update	CER@Median update
0.5	13.87	14.02
0.6	13.76	13.82
0.7	13.73	13.73
0.8	**13.67**	**13.67**
0.9	13.74	13.69

λ was further increased to 0.9, the CER increased slightly, indicating that an excessively large threshold value may incorrectly merge characters from adjacent text columns and increase ordering errors.

Based on the overall performance across different parameter configurations, $\lambda = 0.8$ was selected as the default setting. Since both update strategies achieved the same CER under the optimal configuration, the median strategy was adopted as the default column-center update strategy.

D. Additional Qualitative Results

Figure 1 further presents qualitative comparisons between the input document images and their corresponding restoration results under the hyperparameter configuration of $\tau_r = 90$ and $(\tau_{rg}, \tau_{rb}) = (1.3, 1.3)$. The input Japanese historical document images exhibit severe seal interference, with many seals overlapping Kuzushiji characters and substantially hindering their readability. After restoration, most red seal interference is effectively removed, leaving only faint pink traces in a small number of cases.

However, in regions with dense seal coverage, the restoration process may still introduce locally overexposed areas or over-smoothed textures, leading to the loss of fine paper details. Although such artifacts remain visually acceptable, they will adversely affect pixel-level fidelity metrics, such as PSNR and SSIM, and partially compromise the integrity of local character stroke structures.

References

[1] National Institute of Japanese Literature. Japanese Classical Book Kuzushiji Dataset（日本古典籍くずし字データセット）, 2016. 1

[2] Alexandru Telea. An Image Inpainting Technique Based on the Fast Marching Method. *Journal of Graphics Tools*, 9(1): 23–34, 2004. 1