Binarizing Severely Degraded Ancient Bamboo Slips: Dataset and Baseline

Chongsheng Zhang[1,2]
cszhang@henu.edu.cn
Wanwan Fu[1]
fww@henu.edu.cn
Qilong Li[1]
qilonghenu@henu.edu.cn
Samra Zafar[1]
samra.z@henu.edu.cn
Zhanshuo Zhang[1]
zzs@henu.edu.cn
Qiyan Li[1]
qiyanli@henu.edu.cn
Gaojuan Fan[1]
fangaojuan@henu.edu.cn
Christian Heumann[2]
christian.heumann@stat.uni-muenchen.de

[1] Henan University
MOE Key Research Institute of Yellow River Civilization and Sustainable Development,
Henan Provincial Key Laboratory of Big Data Analytics and Processing,
School of Computer and Information Engineering,
Kaifeng, China

[2] LMU Munich
Department of Statistics
Munich, Germany

Abstract

Bamboo and wooden slips were the primary writing materials in China for more than 800 years, carrying valuable historical records. However, due to the lengthy corrosion and/or weathering effects over a period of two millennia, texts on the ancient bamboo (and wooden) slips typically suffer from severe degradation problems, such as ink deterioration and text blur, which render the binarization of severely degraded Bamboo Slips manuscripts a very challenging task. Due to the scarcity of benchmark dataset in this direction, in this work we aim to build *HanBamboo*, a specialized Bamboo Slips dataset with pixel-level annotations for research on binarization of severely degraded ancient manuscripts. HanBamboo comprises 1,000 infrared Bamboo Slips images displaying varying levels of ink degradation and text blur. Among them, Bamboo Slips exhibiting low-contrast ink traces characterized by significant fading and diminished visibility requires remarkably greater time in pixel-level annotation, indicating the inherent difficulty in binarizing these Bamboo Slips. As a technical baseline, we also propose a baseline approach MRA-Net, which is an Mamba-based Encoder-Decoder framework for degraded Bamboo Slips binarization that devises additional multi-scale discrete wavelet transform and adaptive reweighting attention mechanisms to capture and enhance the stroke details of the texts and suppress background noise during binarization. Comprehensive experiments on both HanBamboo and public document binarization benchmark datasets DIBCO 2017 and 2018 demonstrate the effectiveness of our baseline. Our dataset and code are available at: https://github.com/cszhangLMU/MRA-Net.

1 Introduction

As the primary writing materials in ancient China from the 4th century BC to the 4th century AD, before paper became widely adopted, bamboo (and wooden) slips preserve crucial documentary evidence for ancient Chinese history and civilization. To date, over 300,000 pieces of Bamboo Slips have been unearthed. However, after two millennia of exposure to corrosion, and oxidation or weathering, the inscriptions on the Bamboo Slips often suffer from severe degradation problems, as demonstrated in Figure 1, which significantly compromise the readability of the ancient texts, creating substantial obstacles for Bamboo Slips studies, especially for the new Bamboo Slips that are being continuously excavated, for which domain experts have to manually check, understand, and transcribe the texts from scratch.

As an important document preprocessing technique, binarization is essential for separating inscriptions from the background, removing noise interference and enhancing the clarity and legibility of the texts. The binarization of Bamboo Slips will also be beneficial for potential tasks such as ink enhancement, text inpainting and recognition, writing style verification, and fragments rejoining. However, little research has been conducted in this direction, primarily due to the scarcity of a specialized pixel-level annotated Bamboo Slips dataset—a prerequisite for this task, which is prohibitively expensive and time-consuming to curate.

On the other hand, although existing document image binarization methods such as Encoder-Decoder based approaches/architectures [2, 28] and Generative approaches [21, 27] have made substantial advances on common/standard document binarization benchmark datasets such as DIBCO [14], they encounter new challenges when applied to Bamboo Slips images that exhibit unique degradation pattern, including text degradation, blurred character strokes, and low contrast between text and background, as shown in Figure 1.

To address the above problems, in this work we aim to curate a specialized Bamboo Slips dataset with pixel-level annotations for severely degraded ancient manuscript binarization research, referred to as *HanBamboo*, based on a large collection of *Xuanquan* Bamboo Slips [20] written by the border guards stationed in the northwestern regions of China during the *Han* Dynasty, approximately two millennia ago. Our dataset consists of 1000 infrared images of Bamboo Slips, with varying degrees of ink degradation and text blur. For each infrared Bamboo Slips image, we create pixel-level annotations for the characters and their strokes in the Bamboo Slips, during which we observed that Bamboo Slips exhibiting low-contrast ink traces characterized by significant fading and diminished visibility demands substantially greater manual annotation time. This motivates us to further divide the dataset into four subsets based on the degradation degree and estimated annotation difficulty level. We will also show that, the overall noise and entropy metric scores of the Bamboo Slips in different subsets exhibit a consistent positive correlation with the overall degradation degree of the subset.

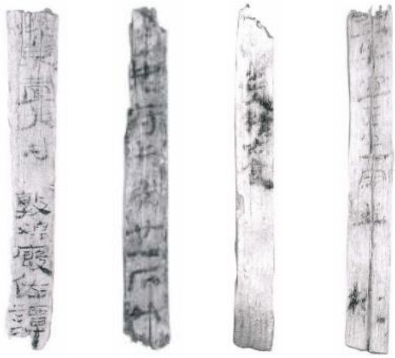

Figure 1: Four severely degraded ancient Bamboo Slips, with challenges such as uneven ink distribution, corrosion of the slip surface, ink bleeding, and fading of the inscriptions.

As a technical baseline, we also propose a baseline approach MRA-Net, which is an end-to-end document binarization architecture for Bamboo Slips. It incorporates a Visual

State Space (VSS) block to leverage Mamba's global modeling capabilities [6, 11] to enhance the feature extraction of ink traces. Moreover, it devises a Multi-scale Wavelet Feature Fusion block (MSWF) which employs frequency-domain decoupling mechanism to further improve the localization and representation of the stroke details. It also contains an Adaptive Reweighting Attention block (ARA) to suppress background noise and enhance text details.

To validate the effectiveness of our baseline, we conduct comprehensive experiments on both HanBamboo and document binarization benchmark datasets DIBCO 2017 and 2018, which demonstrate the outstanding document binarization performance of our baseline, particularly on our HanBamboo dataset, where MRA-Net outperforms existing method across all the evaluation metrics, which are F-Measure (FM), pseudo F-Measure (pFM), Peak Signal-to-Noise Ratio (PSNR), and Distance Reciprocal Distortion metric (DRD).

In summary, the main contributions of this paper are as follows:

- We build a pixel-level annotated Bamboo Slips dataset with 1000 infrared images for promoting research in severely degraded ancient manuscript binarization.
- We propose MRA-Net, a baseline method which can preserve text details and reduce complex background interference during degraded Bamboo Slips binarization.
- On our dataset, the proposed baseline outperforms latest document binarization approaches by 0.74% and 0.68% in terms of FM and pFM, and 0.17 and 0.29 in terms of PSNR and DRD, confirming its effectiveness in degraded documents binarization.

2 Related Work

In recent years, AI-empowered interdisciplinary ancient manuscript research [29, 30] started attracting attention from both the Bamboo Slips research community and the Computer Science/Artificial Intelligence community. Several preliminary efforts have been devoted to Bamboo Slips-related tasks, including ancient character detection and recognition [9, 10, 31], and ancient language understanding [3], etc. While manuscript and documentary image binarization has been a long-standing research focus [24], the binarization of low-quality ancient Bamboo Slips has received limited attention, despite being a real-world problem with significant challenges.

2.1 Document Binarization Datasets

In the field of document image binarization, several publicly available datasets have been released. Among them, the most representative ones are the DIBCO benchmark series [5, 14], each edition of DIBCO has provided a new small-scale but high-quality dataset which typically containing 10 to 20 images with various forms of degradation. The ICFHR 2016 competition on the analysis of handwritten text in images of Balinese Palm Leaf manuscripts [1] addresses the unique challenges of these ancient documents. A key sub-task is binarization, aimed at enhancing the readability of the manuscripts. In addition, specialized datasets such as LS-HDIB [16] focus on different cultural backgrounds and document formats, offering rich resources for expanding document binarization algorithms across diverse scenarios.

In latest years, preliminary explorations into AI-enabled Bamboo Slips research have emerged [3, 9, 10], with a primary focus on text spotting tasks. Nonetheless, to the best of our knowledge, little attention has been paid to the binarization of severely degraded Bamboo

Slips, primarily due to the lack of a pixel-level annotated Bamboo Slips dataset, which is essential for this task but is prohibitively expensive and time-demanding to construct. To bridge this gap, in this work we present the *HanBamboo* dataset, aiming to establish it as a benchmark dataset for advancing research in binarizing severely degraded Bamboo Slips.

2.2 Document Binarization Techniques

Document binarization methods can be divided into traditional thresholding methods [13] and deep learning-based approaches [8], the latter can be further categorized into Encoder-Decoder based architectures/approaches [2, 28] and Generative approaches [21, 23, 27]. Methods in the former subcategory formulate document binarization as a supervised semantic segmentation task, but require pixel-level annotations, which are thus often challenged by limited amount of annotated documents. The latter subcategory treats document binarization as a generative task, consisting of both supervised GAN and unsupervised image-to-image translation approaches.

Recently, researchers have explored using Vision Transformers (ViT) for document binarization task [22, 26]. There are also attempts that also apply diffusion models to document image binarization [17], yielding promising preliminary results.

3 Dataset Construction

3.1 Data Collection

We obtain the original Bamboo Slips from domain experts in Bamboo Slips studies, which come from *Xuanquan Han Slips* I∼III [20], unearthed at the *Xuanquan* guard post of *Gansu* province in northwestern China, close to 7,000 Bamboo Slips were discovered in total. Both color and infrared images were provided, but given that infrared Bamboo Slips images can better reflect ink strokes and text details, as illustrated in Figure 2, they are thus adopted as the main research materials by domain experts in their manual transcription process.

Out of these 7,000 Bamboo Slips, we select 1,000 infrared images, based on different levels of ink/text degradation, reflecting varying degrees of legibility (from easy to hard) of the texts present in the Bamboo Slips. Upon this, we divide our dataset into 4 subsets, shown in Table 1, which are *minor*, *moderate*, *severe* and *extreme* degradation groups. For each subgroup of Bamboo Slips, we equally (and randomly) split it into the training and test sets.

Figure 2: Comparison between color and infrared Bamboo Slips images in text legibility.

Figure 3: Comparison between automated annotation results and manually labeling results.

3.2 Dataset Annotation

In order to preserve the original text details in the Bamboo Slips, our annotation process is primarily manual, with the support of annotation tool. In this process, we adopt a differentiated annotation strategy for speeding up the labeling process. In specific, for infrared Bamboo Slips in the *minor degradation* group (A), we first utilize automated annotation software to produce the initial labeling results, then we check and manually correct a small portion of inaccurate annotations. But for other groups, existing automated software/model yields unsatisfactory annotation results. As shown in Figure 3, challenges such as strong background noise, faint ink, low-contrast between text and background in low-quality Bamboo Slips images cannot be effectively handled by automated annotation software. So we have to label the images from scratch for the remaining groups. In total, one of our team members has spent 6 months in the pixel-level annotation for all the infrared images in our dataset.

3.3 Dataset Statistics and Characteristics

Bamboo Slips present extremely low aspect ratios, since they are characterized by a very narrow (e.g., only 1 cm) but relatively long shape (e.g., 23 cm), as shown in Figure 1. In Table 1, we present the number of images in the four subgroups. In Figure 4, we calculate the averaged noise levels and entropy values for Bamboo Slips in these subgroups, in which the last two subgroups C and D have higher noise interference and complexity, indicating a gradual increase in both noise and entropy with respect to the degradation degrees.

Category	Train	Test	Total
A:Minor Degradation	269	269	538
B:Moderate Degradation	145	145	290
C:Severe Degradation	59	59	118
D:Extreme Degradation	27	27	54

Table 1: HanBamboo is divided into four subgroups based on degradation degree.

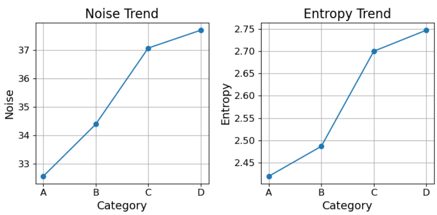

Figure 4: Overall noise and entropy levels for Bamboo Slips images in different subgroups.

Our dataset presents distinctive characteristics, including: (i) low image quality; (ii) low inscription quality, with severely degraded inks and blur texts; (iii) low contrast between the texts and background; (iv) varying degrees of noise and complexity in the images. Such characteristics pose unique challenges, making our dataset a valuable benchmark for advancing research on the binarization of severely degraded, low-quality ancient manuscripts.

4 Baseline Method

4.1 Overall Framework

In Figure 5, we present our **M**amba based bina**R**ization **A**rchitecture MRA-Net, which adopts an Encoder-Decoder framework. In the encoder, it first extracts the initial convolutional features, then adopts hierarchical residual layers that progressively capture semantic context, in which it incorporates Visual State Space (VSS) and Adaptive Reweighting At-

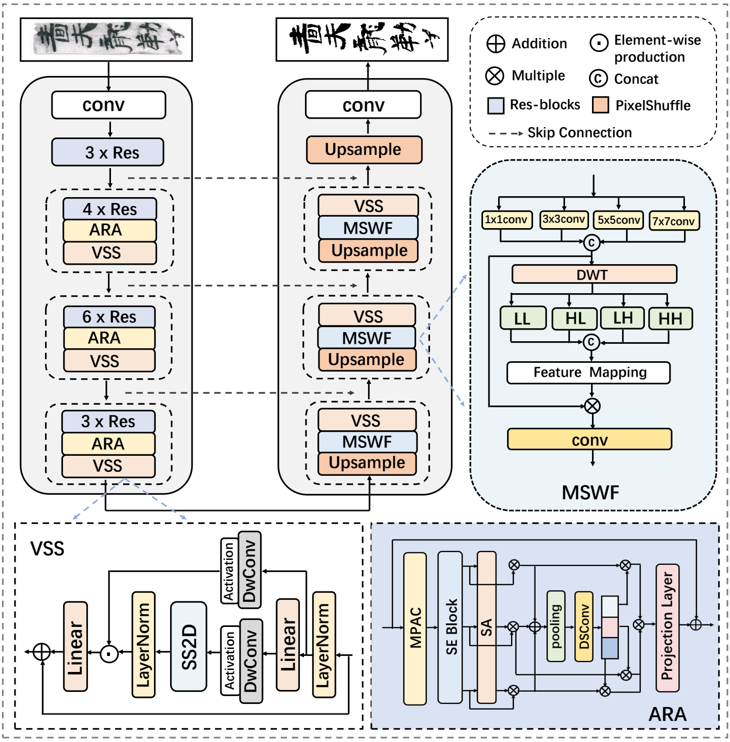

Figure 5: Overall architecture of MRA-Net.

tention (ARA) blocks to enhance feature representation of ink traces. In the decoder, it restores the spatial resolution via pixel shuffling [19] based up-sampling; more importantly, it combines Multi-Scale Wavelet Feature Fusion (MSWF) and VSS to separate texts from background noises and improve the recovery of fine-details of texts in degraded texts.

4.2 Visual State Space Block

In each VSS block, input features first pass through a linear embedding layer and then split into two branches. The first branch is essentially a fully connected layer with activation; the second branch sequentially applies linear transformation, depthwise separable convolution, and 2D Selective Scanning (SS2D) operations. The outputs from both branches are then integrated as the final output of the VSS block.

4.3 Adaptive Reweighting Attention Block

To address the challenges of complex background noise and severe ink degradation in Bamboo Slips images, we introduce the ARA block to capture and enhance the text details. It involves a MPAC module which utilizes directional perception feature modeling mechanism to capture spatial variations along both height and width dimensions. Additionally, Squeeze-and-Excitation (SE) and cross-dimensional attention gating mechanisms are introduced to adaptively model the importance of features across the channel, height, and width dimensions, enhancing the model's focus on the stroke details of degraded texts.

4.4 Multi-scale Wavelet Feature Fusion Block

The MSWF block first applies convolutional kernels of varying scales to extract spatial features, then apply Discrete Wavelet Transform (DWT) [12] to decompose the features into one low-frequency and three high-frequency components to capture the structural information and directional details of the texts. Overall, by integrating spatial-domain multi-scale convolution with frequency-domain wavelet-based modulation, the MSWF module significantly enhances the model's ability in capturing character/stroke details.

4.5 Loss Function

To preserve stroke details while suppressing background noise interference, we adopt a comprehensive loss function that integrates Binary Cross Entropy, Soft Dice loss, and Mean Absolute Error, to balance between pixel-level classification accuracy and structural restoration.

5 Experiments

We conduct comprehensive experiments on both HanBamboo and public document binarization benchmark datasets DIBCO 2017 and H-DIBCO 2018 (DIBCO 2017/2018 in short) to compare the performance of different algorithms. We adopt four standard metrics: (1) F-Measure (FM); (2) pseudo F-Measure (pFM); (3) Peak Signal-to-Noise Ratio (PSNR); and (4) Distance Reciprocal Distortion Metric (DRD). We consistently use the same official protocol when calculating the metrics in all the evaluations.

5.1 Comparative Experiments

We first compare our proposed baseline with both traditional and deep learning-based document binarization techniques on the Hanbamboo dataset.

As shown in Table 2, the proposed baseline achieves the best performance across all four evaluation metrics, demonstrating its superiority in degraded Bamboo Slips image binarization. Specifically, our proposed baseline outperforms latest document binarization methods FourBi [15] by 0.74% and 0.68% in terms of FM and pFM, respectively, and achieves improvements of 0.17 and 0.29 in PSNR and DRD. In particular, the DRD value is further reduced to 4.33 by our baseline, indicating a notable advantage in text structure/strokes preservation.

Method	FM↑	pFM↑	PSNR↑	DRD↓
Otsu [13]	69.52	70.16	13.72	30.91
Sauvola [18]	68.77	69.63	13.65	24.96
Gatos [4]	73.12	73.60	14.49	16.79
Suh [23]	83.94	82.78	17.27	5.86
DP-LinkNet [25]	83.96	83.89	<u>17.84</u>	4.81
cGANs [32]	83.14	83.26	17.53	5.04
Zhao [33]	81.84	81.52	17.25	6.48
DocEnTr [22]	83.30	83.55	16.43	6.11
FourBi [15]	<u>84.13</u>	<u>84.48</u>	17.81	<u>4.62</u>
Proposed (ours)	**84.87**	**85.16**	**17.98**	**4.33**

Table 2: Comparison results of different methods on the HanBamboo dataset.

In Tables 3 and 4, we respectively report the specific binarization results on the severely and extremely degraded Bamboo Slips (categories C and D in Table 1). We can see that, there

Method	FM↑	pFM↑	PSNR↑	DRD↓
Otsu [13]	60.31	60.56	11.81	59.34
Sauvola [18]	64.98	65.14	12.90	29.25
Gatos [4]	67.67	67.52	13.54	22.48
Suh [23]	81.68	81.41	17.02	4.67
DP-LinkNet [25]	81.89	81.79	17.08	4.55
cGANs [32]	81.14	81.56	16.88	4.82
Zhao [33]	80.06	80.22	16.38	5.48
DocEnTr [22]	81.38	81.52	16.63	4.35
FourBi [15]	<u>82.03</u>	<u>81.98</u>	<u>16.98</u>	<u>4.32</u>
Proposed (ours)	**82.34**	**82.49**	**17.24**	**4.16**

Table 3: Specific binarization results on category C (Severe degradation).

Method	FM↑	pFM↑	PSNR↑	DRD↓
Otsu [13]	56.25	56.68	9.68	51.72
Sauvola [18]	63.83	63.82	11.35	24.77
Gatos [4]	65.80	65.45	11.80	20.57
Suh [23]	79.14	78.45	14.87	6.56
DP-LinkNet [25]	79.26	78.82	15.04	5.16
cGANs [32]	78.14	78.19	14.93	5.78
Zhao [33]	76.24	76.05	14.67	7.23
DocEnTr [22]	77.21	76.82	14.33	7.11
FourBi [15]	<u>79.65</u>	<u>79.27</u>	<u>15.08</u>	<u>5.12</u>
Proposed (ours)	**79.97**	**79.86**	**15.16**	**4.78**

Table 4: Specific binarization results on category D (Extreme degradation).

Method	DIBCO 2017				H-DIBCO 2018			
	FM↑	pFM↑	PSNR↑	DRD↓	FM↑	pFM↑	PSNR↑	DRD↓
Otsu [13]	77.73	77.89	13.85	15.54	51.45	53.05	9.74	59.07
Sauvola [18]	77.11	84.10	14.25	8.85	67.81	74.08	13.78	17.69
Howe [7]	90.10	90.95	18.52	5.12	80.84	82.85	16.67	11.96
cGANs [32]	90.73	92.58	17.83	3.58	87.73	90.60	18.37	4.58
Suh [23]	90.95	94.65	18.40	2.93	91.86	<u>96.25</u>	20.03	2.60
GDB [28]	<u>94.32</u>	<u>96.58</u>	<u>20.04</u>	**1.79**	91.09	94.57	19.92	3.07
DocEnTr [22]	92.53	95.15	19.11	2.37	90.59	93.97	19.46	3.35
FourBi [15]	93.81	96.57	19.66	2.03	93.16	**98.40**	**20.92**	**2.23**
D^2BFormer [26]	93.52	95.09	19.35	2.12	88.84	93.42	18.91	3.99
Proposed	**94.59**	**96.87**	**20.12**	<u>1.98</u>	**93.89**	95.87	<u>20.12</u>	<u>2.47</u>

Table 5: Comparison of MRA-Net on the DIBCO 2017 and H-DIBCO 2018 datasets.

is substantial performance drop in these two categories, indicating the increased challenge and difficulty in binarizing such severely degraded Bamboo Slips.

To further evaluate the performance of our proposed baseline method, we also conduct comparative experiments on the DIBCO 2017/2018 benchmark datasets. In Table 5, we see that it obtains superior or comparable performance than latest binarization algorithms, demonstrating its applicability across various document types and domains.

On the other hand, when cross-comparing the evaluation scores between HanBamboo and DIBCO 2017/2018, we see that HanBamboo is significantly more difficult than DIBCO, and there remains a large room for further performance improvements on our dataset.

5.2 Ablation Study

In Table 6 and 7, we carry out comprehensive ablation studies on both HanBamboo and H-DIBCO 2018 to assess the contributions of the VSS, ARA, and MSWF blocks to the overall model performance. In Table 6, we see that on the HanBamboo dataset, the VSS block enhances global modeling capabilities. The introduction of ARA leads to a notable increase in PSNR, indicating its effectiveness in preservation of fine details of texts in the ancient document. When MSWF is further integrated, pFM has been significantly improved, highlighting its capability in fine-detail capturing and preservation of critical text structural information such as subtle stroke characteristics, which is also beneficial for texts with faded inks. Similar observations also hold on the H-DIBCO 2018 dataset, as can be seen from Table 7.

ARA	MSWF	VSS	FM↑	pFM↑	PSNR↑	DRD↓
✓			83.20	83.50	16.56	5.81
	✓		82.99	82.89	16.89	5.21
		✓	83.40	83.37	17.12	5.38
	✓	✓	<u>84.56</u>	84.53	17.45	4.75
✓		✓	84.42	84.37	<u>17.91</u>	<u>4.62</u>
✓	✓		84.15	<u>84.64</u>	17.55	4.64
✓	✓	✓	**84.87**	**85.16**	**17.98**	**4.33**

Table 6: Ablation results on the HanBamboo dataset.

ARA	MSWF	VSS	FM↑	pFM↑	PSNR↑	DRD↓
✓			92.62	94.35	17.31	4.21
	✓		92.46	93.97	17.09	4.34
		✓	92.48	94.57	17.54	4.38
	✓	✓	93.02	94.85	18.42	<u>2.89</u>
✓		✓	<u>93.23</u>	<u>95.11</u>	<u>19.52</u>	3.12
✓	✓		92.85	94.61	18.55	3.64
✓	✓	✓	**93.89**	**95.87**	**20.12**	**2.47**

Table 7: Ablation results on the H-DIBCO 2018 dataset.

5.3 Visual Comparisons

We provide visual analysis of the binarization results on the HanBamboo dataset to further demonstrate the effectiveness of the proposed MRA-Net baseline method for binarizing degraded Bamboo Slips images. In Figure 6, we provide two original Bamboo Slips images, their corresponding ground-truth annotations (GT), and the visualized document binarization results of representative methods. These two images are representative infrared Bamboo Slips images that exhibit typical degradation problems, including faded ink and background noise interference, and low distinction and contrast between the inscriptions and background.

From Figure 6, we observe that traditional methods such as Otsu struggle to suppress background noise, and often introduce false contours, while existing deep learning-based document binarization methods can retain main text regions but are less effective in distinguishing strokes from noises on Bamboo Slips with severely faded inks, as can be observed from the top figure. In contrast, MRA-Net can effectively preserves stroke edges and character structures for severely degraded Bamboo Slips, while suppressing background noise.

6 Conclusion

In this work, we present HanBamboo, a specialized Bamboo Slips dataset with high-quality pixel-level annotations, which enables research in severely degraded Bamboo Slips binarization and lays the foundation for relevant tasks in the domain such as ink enhancement, text restoration, writing style verification and fragments rejoining. As a technical baseline, we propose a baseline method MRA-Net, a binarization network architecture tailored for degraded Bamboo Slips, which can effectively separate texts from background noise, and preserve fine-details and critical text structural information during binarization. Experimental results on the constructed HanBamboo dataset as well as public DIBCO 2017/2018 benchmark datasets validate the effectiveness and applicability of MRA-Net.

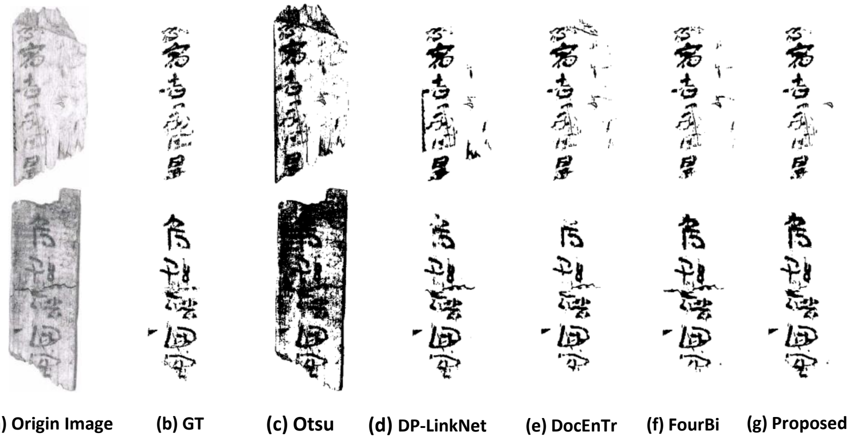

Figure 6: Visual comparisons between MRA-Net and existing methods on HanBamboo.

Acknowledgments

This work was partially supported by the MOE Liberal Arts and Social Sciences Foundation (No.23YJAZH210), Major Program of National Social Science Foundation (No.23&ZD309), High Level Talent International Training Program of Henan Province (No. GCC2025010), Henan Provincial Center for Outstanding Overseas Scientists (No.GZS2025004), Chinese Scholarship Council (CSC, Grant No.202308410339), and the Department of Statistics, LMU Munich, Germany. We would like to express our sincere gratitude to Associate Professor Lei Yao of Xinyang Normal University, China, for providing the relevant data and materials, as well as offering professional guidance in domain-specific knowledge in Bamboo Slips research. Finally, we thank all the anonymous reviewers and area chair(s) for their constructive suggestions and kind help.

References

[1] Jean-Christophe Burie, Mickaël Coustaty, Setiawan Hadi, Made Windu Antara Kesiman, Jean-Marc Ogier, Erick Paulus, Kimheng Sok, I Made Gede Sunarya, and Dona Valy. Icfhr2016 competition on the analysis of handwritten text in images of balinese palm leaf manuscripts. In *ICFHR 2016)*, pages 596–601, 2016.

[2] Jorge Calvo-Zaragoza and Antonio-Javier Gallego. A selectional auto-encoder approach for document image binarization. *Pattern Recognition*, 86:37–47, 2019.

[3] Yingfa Chen, Chenlong Hu, Cong Feng, Chenyang Song, Shi Yu, Xu Han, Zhiyuan Liu, and Maosong Sun. Multi-modal multi-granularity tokenizer for chu bamboo slips. In *COLING 2025*, pages 6201–6211, 2025.

[4] Basilios Gatos, Ioannis Pratikakis, and Stavros J Perantonis. Adaptive degraded document image binarization. *Pattern recognition*, 39(3):317–327, 2006.

[5] Basilis Gatos, Konstantinos Ntirogiannis, and Ioannis Pratikakis. Icdar 2009 document image binarization contest (dibco 2009). In *ICDAR 2009*, pages 1375–1382, 2009.

[6] Albert Gu and Tri Dao. Mamba: Linear-time sequence modeling with selective state spaces. *arXiv preprint arXiv:2312.00752*, 2023.

[7] Nicholas R Howe. Document binarization with automatic parameter tuning. *International Journal on Document Analysis and Recognition*, 16(3):247–258, 2013.

[8] Seokjun Kang, Brian Kenji Iwana, and Seiichi Uchida. Complex image processing with less data—document image binarization by integrating multiple pre-trained u-net modules. *Pattern Recognition*, 109:107577, 2021.

[9] Shengnan Li, Chi Zhou, and Kaili Wang. Wa-net: Wavelet integrated attention network for silk and bamboo character recognition. *Engineering Applications of Artificial Intelligence*, 140:109674, 2025.

[10] Yiran Liu, Qiang Zhang, Ying Qi, Teng Wan, Defang Zhang, Yutong Li, Xin Zhang, Longbin Ma, Qiuyue Ruan, Huanting Guo, et al. Deepjiandu dataset for character detection and recognition on jiandu manuscript. *Scientific Data*, 12(1):398, 2025.

[11] Yue Liu, Yunjie Tian, Yuzhong Zhao, Hongtian Yu, Lingxi Xie, Yaowei Wang, Qixiang Ye, Jianbin Jiao, and Yunfan Liu. Vmamba: Visual state space model. In *NeurIPS 2024*, volume 37, pages 103031–103063, 2024.

[12] S. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693, 1989.

[13] Nobuyuki Otsu et al. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296):23–27, 1975.

[14] Ioannis Pratikakis, Konstantinos Zagoris, Xenofon Karagiannis, Lazaros Tsochatzidis, Tanmoy Mondal, and Isabelle Marthot-Santaniello. Icdar 2019 competition on document image binarization (dibco 2019). In *ICDAR 2019*, pages 1547–1556, 2019.

[15] Fabio Quattrini, Vittorio Pippi, Silvia Cascianelli, and Rita Cucchiara. Binarizing documents by leveraging both space and frequency. In *ICDAR 2024*, pages 3–22, 2024.

[16] Kaustubh Sadekar, Ashish Tiwari, Prajwal Singh, and Shanmuganathan Raman. LS-HDIB: a large scale handwritten document image binarization dataset. In *ICPR 2022*, pages 1678–1684, 2022.

[17] Saifullah Saifullah, Stefan Agne, Andreas Dengel, and Sheraz Ahmed. Coldbin: Cold diffusion for document image binarization. In *ICDAR 2023*, pages 207–226, 2023.

[18] Jaakko Sauvola and Matti Pietikäinen. Adaptive document image binarization. *Pattern recognition*, 33(2):225–236, 2000.

[19] Wenzhe Shi, Jose Caballero, Ferenc Huszár, Julian Totz, Andrew P Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *CVPR 2016*, pages 1874–1883, 2016.

[20] Silk Road Editorial Office. The Publication of Xuanquan Han Slips (Vol. III). *Silk Road*, (2), 2023. In Chinese.

[21] Mohamed Ali Souibgui and Yousri Kessentini. De-gan: A conditional generative adversarial network for document enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(3):1180–1191, 2020.

[22] Mohamed Ali Souibgui, Sanket Biswas, Sana Khamekhem Jemni, Yousri Kessentini, Alicia Fornés, Josep Lladós, and Umapada Pal. Docentr: An end-to-end document image enhancement transformer. In *ICPR 2022)*, pages 1699–1705, 2022.

[23] Sungho Suh, Jihun Kim, Paul Lukowicz, and Yong Oh Lee. Two-stage generative adversarial networks for binarization of color document images. *Pattern Recognition*, 130:108810, 2022.

[24] Richin Sukesh, Mathias Seuret, Anguelos Nicolaou, Martin Mayr, and Vincent Christlein. A fair evaluation of various deep learning-based document image binarization approaches. In *International Workshop on Document Analysis Systems (DAS 2022)*, pages 771–785, 2022.

[25] Wei Xiong, Xiuhong Jia, Dichun Yang, Meihui Ai, Lirong Li, and Song Wang. Dp-linknet: A convolutional network for historical document image binarization. *KSII Transactions on Internet and Information Systems (TIIS)*, 15(5):1778–1797, 2021.

[26] Mingming Yang and Songhua Xu. A novel degraded document binarization model through vision transformer network. *Information Fusion*, 93:159–173, 2023.

[27] Zongyuan Yang, Baolin Liu, Yongping Xxiong, Lan Yi, Guibin Wu, Xiaojun Tang, Ziqi Liu, Junjie Zhou, and Xing Zhang. Docdiff: Document enhancement via residual diffusion models. In *ACM Multimedia 2025*, pages 2795–2806, 2023.

[28] Zongyuan Yang, Baolin Liu, Yongping Xiong, and Guibin Wu. GDB: Gated convolutions-based document binarization. *Pattern Recognition*, 146:109989, 2024.

[29] Chongsheng Zhang, Ruixing Zong, Shuang Cao, Yi Men, and Bofeng Mo. Ai-powered oracle bone inscriptions recognition and fragments rejoining. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI 2020)*, pages 5309–5311, 2020.

[30] Chongsheng Zhang, Bin Wang, Ke Chen, Ruixing Zong, Bofeng Mo, Yi Men, George Almpanidis, Shanxiong Chen, and Xiangliang Zhang. Data-driven oracle bone rejoining: A dataset and practical self-supervised learning scheme. In *The 28th ACM SIGKDD Conference on Knowledge Discovery (KDD 2022)*, pages 4482–4492, 2022.

[31] Qiang Zhang, Yiran Liu, Ying Qi, Defang Zhang, Ruiqi Feng, and Shanxiong Chen. A feature-enhanced text detection model for jiandu manuscripts. *npj Heritage Science*, 13(1):355, 2025.

[32] Jinyuan Zhao, Cunzhao Shi, Fuxi Jia, Yanna Wang, and Baihua Xiao. Document image binarization with cascaded generators of conditional generative adversarial networks. *Pattern Recognition*, 96:106968, 2019.

[33] Penghai Zhao, Weilan Wang, Guowei Zhang, and Yuqi Lu. Alleviating pseudo-touching in attention u-net-based binarization approach for the historical tibetan document images. *Neural Computing and Applications*, 35(19):13791–13802, 2023.