Ink Enhancement for Ancient Bamboo Manuscripts Using Iterative Restoration-Degradation Adversarial Learning

Chongsheng Zhang[1,2]
cszhang@henu.edu.cn
Junchao Ma[1]
jc8023cb@henu.edu.cn
Wenhao Zhang[1]
xzxs77889@henu.edu.cn
Kamel Aouaidjia[1]
kamel@henu.edu.cn
Qilong Li[1]
qilonghenu@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 Slips are crucial documentary evidences for investigating ancient Chinese history and civilization. However, due to the lengthy deterioration process for about 2000 years, inscriptions on Bamboo Slips typically suffer from ink fading problem, which seriously impacts their clarity and legibility. This paper proposes BambInk, which is among the first technical initiatives that explore AI-enabled ink enhancement for Bamboo Slips with faded inks. BambInk is a new self-supervised learning method that devises an iterative restoration-degradation adversarial learning mechanism to progressively enhance the faded inks on the Bamboo Slips, yet without requiring any annotated data. In specific, BambInk designs two generators which are ink enhancer and ink degrader, and two discriminators for evaluating the effects of the enhanced and degraded images, respectively. Moreover, in the generators, it introduces a dynamic convolution mechanism that integrates features captured via three heterogeneous types of attentions; it also adopts a simple yet effective high-low frequency information differentiation mechanism for extracting the fine-details of the ink traces. Experiments conducted on the real-world Bamboo Slips dataset demonstrate the effectiveness of our method. Code, data and the enhanced results are available at: https://github.com/cszhangLMU/BambInk.

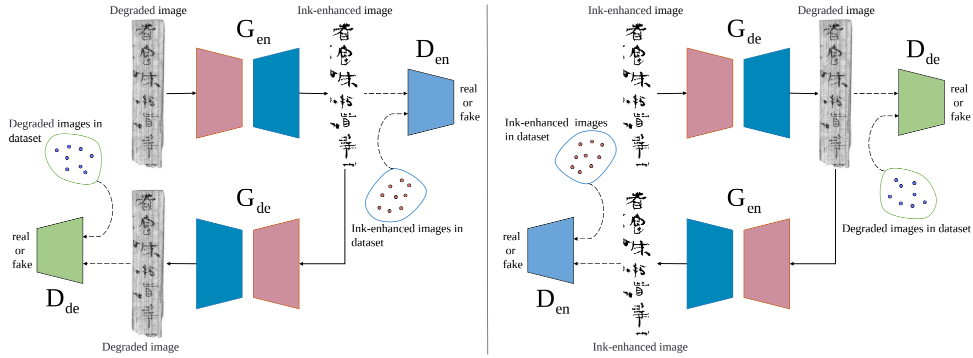

Figure 1: The proposed self-supervised ink enhancement method BambInk using iterative enhance-degrade adversarial learning. Left: the degraded image is processed by the ink enhancer G_{en} to produce the ink-enhanced image, which is then restored to the degraded image by the ink degrader G_{de}. Right: the ink-enhanced image is processed by the ink degrader G_{de} to produce the degraded image, which is then restored to the ink-enhanced image by the ink enhancer G_{en}.

1 Introduction

Xuanquan Bamboo Slips of the Han dynasty [13] excavated in northwestern China are the primary historical materials that offer unique documentary evidence and value for investigating the culture, economy, politics and military affairs during that period. However, due to the text degradation over millennia, bamboo slip images exhibit five challenges: (1) ink fading, (2) discontinuous strokes, (3) background soiling, (4) textural fissures, and (5) significant noise. Furthermore, these images demonstrate variations in character clarity and background complexity. To enhance both legibility (perceptual quality) and readability (semantic comprehension) of the inscriptions on Bamboo slip images, it becomes imperative to design specific ink enhancement technique for such degraded Bamboo Slips.

Existing image processing techniques such as denoising, binarization [17], super-resolution [11], can be used for improving the quality of bamboo slip images. However, these methods primarily focus on global contrast adjustment and noise suppression, without fully considering the inherent characteristics and requirements of the inscriptions on the Bamboo Slips. For instance, losing certain strokes in one character can result in a different character, which cannot be overlooked during the enhancement process. Moreover, binarization methods may struggle to distinguish the subtle differences between certain ink traces and background textures, leading to the loss of structural information in the characters.

To overcome these challenges, we introduce BambInk, a specialized self-supervised learning framework designed for ink enhancement on ancient Bamboo Slips. As depicted in Figure 1, our model uses unpaired high-quality (ink-legible) and low-quality (ink-faded) bamboo slip images as dual style domains for unsupervised training. BambInk consists of four key components: (1) a generator as an ink enhancer, (2) a generator as an ink degrader, (3) an ink enhancement effect discriminator, and (4) a degradation effect discriminator.

Moreover, in the generators of BambInk, we devise a dynamic convolution block that integrates different types of features captured using three types of attention: spatial attention, channel attention, and filter attention to extract both global and local stroke features for better ink traces preservation. We also design a simple yet effective high- and low-frequency information differentiation mechanism for extracting the fine-details of the ink traces. Furthermore, self-attention mechanism is integrated into both discriminators to improve the discrimination capability between generated enhanced images and the originally ink-legible images. We conduct comprehensive experiments on a real-world Bamboo Slips dataset which demonstrate that BambInk significantly outperforms existing methods in ink enhancement.

Our main contributions are summarized as follows:

- As far as we know, this work is among the first technical initiatives that explore AI-empowered ink enhancement on deteriorated bamboo slip manuscripts, which can significantly improve the clarity and legibility of ancient inscriptions on the Bamboo Slips, thereby facilitating more accurate transcription and advancing research in this field, yet without requiring any data annotation.

- We design BambInk, which is a self-supervised learning scheme that adopts an iterative restoration/enhancement-degradation adversarial learning mechanism for progressive ink enhancement, yet without the need of annotated data.

- BambInk consists of two generators and two discriminators. In the generators, we devise a dynamic convolution block that fuses heterogeneous attentions to effectively preserve both global and local stroke features; we also adopt a high- and low-frequency information differentiating mechanism to better extract the fine-details of the ink traces.

- Comparisons with various existing methods on real-world Bamboo Slips dataset demonstrate that BambInk achieves state-of-the-art performance in ink enhancement.

2 Related Work

In recent years, AI-empowered interdisciplinary ancient manuscript research [26, 27] has emerged as a rapidly developing field, which also bridges the Bamboo Slip research community and the computer science/artificial intelligence community. There are a few preliminary efforts devoted to Bamboo Slips related ancient character detection and recognition [12, 13, 28], and ancient language understanding [3], etc. While documentary image and manuscript enhancement has received lots of attention [22], ink enhancement for low-quality Bamboo Slips remains under-explored, despite being both an important prerequisite and a significant real-world challenge for Bamboo Slip research.

Ancient manuscript enhancement using image processing. As one of the earliest handcrafted image processing approaches, image binarization methods employ thresholds and adaptive parameters to handle various manuscript conditions in complex backgrounds. These include thresholding techniques [17], fuzzy methods [19],etc. Later on, several algorithms that combine many image processing techniques, such as were proposed to handle complex documents and manuscripts [6, 10]. Although these methods have their strengths in global contrast enhancement, noise removal, and deblurring, they heavily rely on predefined filter kernels, prior parameters, making them unsuitable for the complex degradation scenarios commonly found in Bamboo slip images.

Ancient manuscript enhancement using deep learning. Deep learning approaches, particularly Convolutional Neural Networks (CNNs), have been applied for enhancement tasks including binarization, deblurring, denoising, and artifact removal [21]. Generative Adversarial Networks (GANs) [5, 29] are key in ancient manuscript enhancement for generating unrestricted enhanced images. Souibgui and Kessentini [18] proposed the first GAN-based enhancement called Document Enhancement Generative Adversarial Networks (DE-GAN) that uses a conditional GAN to address document enhancement. Jemni et al. [8] incorporated text information via Handwritten Text Recognition (HTR) model. Yang et al. [24] proposed DocDiff, a diffusion-based document enhancer that combines coarse prediction and high-frequency refinement. However, these methods only consider overall image enhancement, but ignore the fidelity requirement on inscriptions during enhancement.

Ink enhancement. In the ancient manuscript enhancement literature, ink enhancement is not sufficiently addressed. To the best of our knowledge, the work of Kaneko et al. [9] is the only work specifically proposed to enhance the degraded (faded ink) text of Kuzushi-ji characters in ancient Japanese books using CycleGAN. They alleviate this by adopting the domain discriminator that enforces clearer transformations by penalizing insufficient changes, and an augmented identity loss to better preserve character structure by encouraging identity mapping when no restoration is needed. However, their ink enhancement method was specifically designed for historical texts written on papers with a specific writing style, while little effort has been paid to ink enhancement on Bamboo Slips.

3 Proposed Method

3.1 Overall Framework

We propose the BambInk framework for Bamboo Slips ink enhancement, which is depicted in Figure 1, more details about the architecture are given in Figure 2, in particular the generator. Inspired by CycleGAN [29] and DualGAN [25], BambInk consists of an ink enhancer ($\mathrm{G_{en}}$) and an ink degrader ($\mathrm{G_{de}}$) as generators; and an enhancement effect discriminator ($\mathrm{D_{en}}$) and a degradation effect discriminator ($\mathrm{D_{de}}$). It is noteworthy that identical network architectures are used for $\mathrm{G_{en}}$ and $\mathrm{G_{de}}$, so is the case with $\mathrm{D_{en}}$ and $\mathrm{D_{de}}$.

In the training, a degraded bamboo slip image denoted as X, is split into multiple patches. Each patch is processed by a series of convolutions in the encoder to extract high-level features. These features are then fed into a sequence of blocks, where each block is an attention-based module (AttEnc) that accurately locates and enhances textual ink traces and stroke edge information. The resulting features are upsampled to their original resolution in the decoder through multiple upsampling stages to get a coarse ink-enhanced image $G_{\mathrm{en}}(X)$. During the adversarial training, the real ink-enhanced bamboo slip image patches and the generated $G_{\mathrm{en}}(X)$ are input into the enhancement effect discriminator $\mathrm{D_{en}}$ to distinguish between real ink-legible images and the generated ones (Figure 1(Left)). Simultaneously, to enforce cycle consistency, $G_{\mathrm{en}}(X)$ is also processed by an ink degrader $\mathrm{G_{de}}$, which mirrors the architecture of $\mathrm{G_{en}}$. This process produces a coarse degraded image $G_{\mathrm{de}}(G_{\mathrm{en}}(X))$. The cycle-consistency loss L_{cyc} is applied to enforce it can reconstruct the original input. The same process is then repeated to generate a degraded image from the ink-enhanced image in an inverse way (Figure 1(Right)).

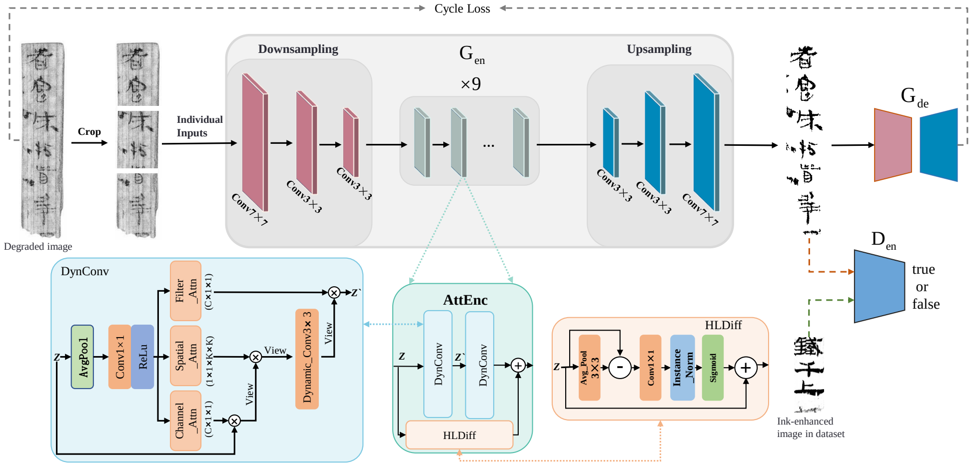

Figure 2: Detailed architecture of the generator of BambInk.

3.2 Ink Enhancer and Ink Degrader

BambInk consists of dual generator networks with identical model structures designated as the ink enhancer G_{en} and the ink degrader G_{de}, where both have an encoder-decoder structure. As shown in Figure 2, for the G_{en}, each single degraded patch is resized at 256×256 with borders extended via mirror padding. This preprocessed input image X is fed into a convolutional layer with filter sizes 7×7, followed by another two 3×3 convolutions, as downsampling steps in the encoder. The resulting features are passed into a sequence of 9 AttEnc modules. The structure of AttEnc is shown in Figure 2, which is specifically designed to pay attention to the ink strokes in the image. After that, the obtained features undergo three upsampling stages with transpose convolution to restore their size to the original dimensions of the input. The ink degrader G_{de} mirrors this architecture in the opposite direction, converting ink-enhanced images into degraded images.

3.3 Enhancement and Degradation Effect Discriminators

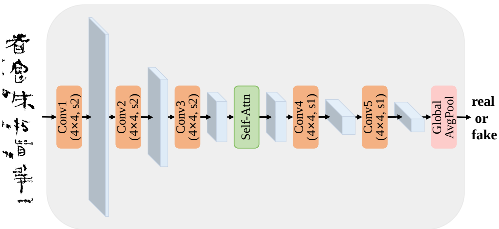

Figure 3: Architecture of the discriminator.

The network architecture of the enhancement/degradation discriminator is illustrated in Figure 3. The input image passes through three layers of 4×4 convolutions with a stride of 2, each followed by spectral normalization. Two additional 4×4 convolutional layers with a stride of 1 are employed. Finally, the number of channels is reduced to 1, and a global average pooling operation is applied to obtain the score for binary classification. We improve the discriminator architecture with a self-attention module after the third convolutional layer to enhance its capability to model global features that are captured using the attention mechanism over the image.

3.4 Dynamic Convolution and High-Low Frequency Differentiating

The Attention Encoder (AttEnc) module in the generator is composed of two consecutive Dynamic Convolutions (DynConv) with a residual connection that passes through a high- and low-frequency differentiating module (HLDiff) (Figure 2).

Dynamic Convolutions Block (DynConv): The DynConv module dynamically captures channel, spatial, and filter attentions to adaptively enable the model to preserve textual ink traces and stroke integrity. As shown in Figure 2, the input feature Z is fed into a global average pooling to derive the global context vector processed through a 1×1 convolution, followed by batch normalization and ReLU activation to generate the intermediate features, which are used to compute the channel, spatial, and filter attention weights through three separate 1×1 convolutions. The intermediate features are multiplied by the channel attention features, then by the spatial attention features to get channel-spatial attention features. The filter attention features are also multiplied by the channel-spatial attention features to get the final Z' output of the DynConv module.

High- and Low-frequency Differentiating Block (HLDiff): The HLDiff module performs an amplification on the features to enhance textual ink traces and stroke edges. As shown in Figure 2, the HLDiff module receives an input feature Z, which is processed with 3×3 average pooling, resulting in a smoothed map that contains low-frequency background information. Then, the original high-frequency feature map subtracts the smoothed low-frequency feature map to extract the high-frequency stroke information. After that, the extracted features pass through a 1×1 convolution, batch normalization, and ReLU activation. Finally, the enhanced features are summed up with the original high-frequency features to generate the final feature map.

3.5 Loss Function

During the training phase, the total loss of the model is composed of contributions from the ink enhancer G_{en}, the ink degrader G_{de}, the enhancement discriminator D_{en}, and the degradation discriminator D_{de}. The generators' losses are defined as follows:

$$L_{G_{\mathrm{en}}} = L_{\mathrm{adv}}\left(G_{\mathrm{en}}, D_{\mathrm{en}}\right) + \lambda_{\mathrm{cyc}} L_{\mathrm{cyc}}\left(G_{\mathrm{en}}, G_{\mathrm{de}}\right) + \lambda_{\mathrm{id}} L_{\mathrm{id}}\left(G_{\mathrm{en}}\right) \tag{1}$$

$$L_{G_{\mathrm{de}}} = L_{\mathrm{adv}}\left(G_{\mathrm{de}}, D_{\mathrm{de}}\right) + \lambda_{\mathrm{cyc}} L_{\mathrm{cyc}}\left(G_{\mathrm{en}}, G_{\mathrm{de}}\right) + \lambda_{\mathrm{id}} L_{\mathrm{id}}\left(G_{\mathrm{de}}\right) \tag{2}$$

where L_{adv} denotes the adversarial loss, and L_{cyc} represents the cycle-consistency loss. Meanwhile, $L_{\mathrm{id}}\left(G_{\mathrm{en}}\right)$ and $L_{\mathrm{id}}\left(G_{\mathrm{de}}\right)$ correspond to the identity losses on the G_{en} and G_{de}, respectively. The discriminators' losses are defined as follows:

$$L_{D_{\mathrm{en}}} = \frac{1}{2}\mathrm{E}_{y \sim Y}\left[\left(D_{\mathrm{en}}(y) - 1\right)^2\right] + \frac{1}{2}\mathrm{E}_{x \sim X}\left[D_{\mathrm{en}}\left(G_{\mathrm{en}}(x)\right)^2\right] \tag{3}$$

$$L_{D_{\mathrm{de}}} = \frac{1}{2}\mathrm{E}_{x \sim X}\left[\left(D_{\mathrm{de}}(x) - 1\right)^2\right] + \frac{1}{2}\mathrm{E}_{y \sim Y}\left[D_{\mathrm{de}}\left(G_{\mathrm{de}}(y)\right)^2\right] \tag{4}$$

4 Experiments

4.1 Experimental Settings

Dataset: We conducted experiments and evaluation on a subset of *Xuanquan* Bamboo Slips collection from the Han Dynasty. From 2019 to 2023, three volumes (I to III) of Xuanquan Bamboo Slips were released, belong to different periods. Bamboo Slips in later periods show improvements in texts clarity and suppression of background noise, as can be seen from the four Bamboo Slips in Figure 4.

Therefore, we select 350 images from the ink-legible Xuanquan II Bamboo Slips as samples of the target domain for training, and randomly choose 400 images from the degraded Xuanquan I Bamboo Slips as the degraded source domain. As a preprocessing step, data augmentation, including flipping, rotation, was applied.

Figure 4: Comparison of Xuanquan Bamboo Slips of the Han dynasty from different periods.

Evaluation Metrics: We adopt two types of evaluations: unsupervised and supervised assessments. In the unsupervised evaluation, we compare the enhanced images with the original input images using three metrics [30]: the Frechet Inception Distance (FID), Kernel Inception Distance (KID), and Structural Similarity Index Measure (SSIM) to assess the effectiveness of the enhanced images in terms of distribution matching and overall structural integrity. In the supervised evaluation, we assess the binarization results of the ink-enhanced Bamboo Slips, for which we make the pixel-level annotations for 100 degraded Bamboo Slips, which will be used as the ground-truth (GT) in binarization. We calculate the F-Measure (FM), Pseudo F-Measure (pFM), Peak Signal-to-Noise Ratio (PSNR), and Distortion Ratio Difference (DRD) scores to quantitatively measure the ink-enhancement effects.

4.2 Unsupervised Evaluation

As presented in Table 1, our proposed BambInk method achieves the best performance on FID, KID, and SSIM with scores of 110.2, 0.040, and 0.852, respectively, on the unsupervised evaluation. This represents an improvement of 24.7, 0.020, and 0.050 over the original CycleGAN model on each metric. It also indicates that the images enhanced by BambInk are closer to the original inputs in terms of distribution similarity and structural integrity.

Furthermore, BambInk demonstrates a much higher performance over the best-performing traditional methods across all metrics. On the FID metric, BambInk is 71.0 lower than Niblack (181.2), which reflects stronger distribution consistency. On the KID metric, it is 0.055 lower than WAN (0.095), which shows more accurate feature mapping, and on the SSIM metric, it shows 0.251 higher than ZigZag (0.601), indicating superior structural detail preservation capabilities.

Methods	FID↓	KID↓	SSIM↑
Ostu [16]	244.3	0.170	0.561
NiBlack [15]	181.2	0.100	0.361
Sauvola [17]	235.4	0.166	0.596
Wolf [23]	189.7	0.106	0.568
Gatos [4]	347.4	0.327	0.575
Su [20]	350.8	0.333	0.589
ISauvola [7]	337.2	0.315	0.573
WAN [14]	185.8	0.095	0.519
AutoFoc [2]	324.5	0.278	0.565
ZigZag [1]	285.8	0.229	0.601
CycleGAN [29]	<u>134.9</u>	<u>0.060</u>	<u>0.802</u>
BambInk (Ours)	**110.2**	**0.040**	**0.852**

Table 1: Comparison with other methods on unsupervised evaluation metrics.

Methods	FM↑	pFM↑	PSNR↑	DRD↓
Ostu [16]	60.32	59.97	10.11	41.64
NiBlack [15]	32.96	32.94	4.62	125.3
Sauvola [17]	65.77	65.16	11.27	23.59
Wolf [23]	57.62	57.41	9.50	36.36
Gatos [4]	70.06	69.22	12.25	18.55
Su [20]	68.32	67.57	12.79	*15.39*
ISauvola [7]	68.20	67.45	11.81	20.52
WAN [14]	49.22	49.12	7.85	55.96
AutoFoc [2]	66.80	66.06	12.71	18.52
ZigZag [1]	*71.86*	*70.80*	*12.94*	15.78
CycleGAN [29]	<u>76.81</u>	<u>75.24</u>	<u>14.11</u>	<u>11.02</u>
BambInk(Ours)	**78.91**	**76.94**	**14.78**	**8.56**

Table 2: Comparison with other methods on supervised evaluation metrics.

4.3 Supervised Evaluation

We compared BambInk with ten traditional methods in Table 2. The results indicate significant differences among algorithms across various metrics. For example, due to the limitations of the thresholding strategy of Niblack, it shows the lowest performance across all metrics. The remaining traditional methods achieved relatively better scores. While ZigZag algorithm performs poorly on the DRD metric, it shows the best traditional method results on the FM, pFM, and PSNR metrics.

Compared to deep learning methods, traditional methods showed significant shortcomings across all metrics. The CycleGAN method surpassed the best traditional results by 4.95, 4.44, 1.17, and 4.37 on the four metrics. However, BambInk further improves these four metrics by 2.10, 1.70, 0.67, and 2.46 and achieves the best results on all metrics with scores of 78.91, 76.94, 14.78, and 8.56 on the FM, pFM, PSNR, and DRD metrics, respectively. This demonstrates the efficacy of BambInk in ink enhancement for degraded Bamboo Slips.

4.4 Ablation study

Table 3: Ablation study on the effects of different modules in BambInk.

Baseline	DynConv	HLDiff	Self-Attn	FM↑	pFM↑	PSNR↑	DRD↓
✓				76.81	75.24	14.11	11.02
✓	✓			77.52	75.56	14.35	9.25
✓	✓	✓		78.21	76.20	14.44	8.93
✓	✓	✓	✓	**78.91**	**76.94**	**14.78**	**8.56**

To assess the impact of the modules proposed in our method, we consider CycleGAN as the baseline and combine it with different modules. As shown in Table 3, adding DynConv, HLDiff, and Self-Attn to the baseline enhances the model's performance. When only adding AttEnc (DynConv + HLDiff) to the baseline, the model achieves improvements of 1.40, 0.96, 0.33, and 2.09 across the four evaluation metrics, which shows that this combination significantly enhances text feature extraction, ink enhancement of text, stroke edge reinforcement, and background noise suppression. Further incorporating the Self-Attn module into the discriminator shows the highest performance with an improvement of 2.10, 1.70, 0.67, and 2.46 over the baseline. This ablation study demonstrates the effectiveness of each proposed module in the proposed BambInk framework.

4.5 Visual Analysis

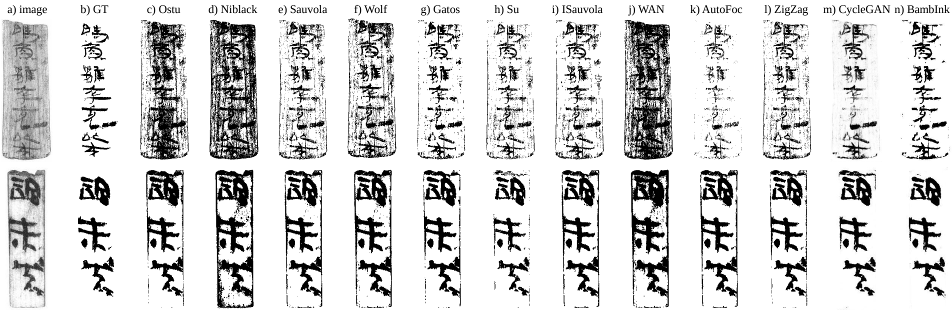

Figure 5: Visual comparisons with different methods.

Figure 5 presents the visual comparison results with different methods on two input Bamboo slip images. In the first row, it is clearly that BambInk ink-enhanced images can effectively distinguish texts/foreground from the background, while conventional methods and the original CycleGAN model without these improvements often misclassify a large number of background pixels as foreground, leading to residual noise and incomplete foreground-background separation results. The images in the second row encounter issues with unclear text boundaries. By comparing the enhancement effects of different methods, it can be observed that conventional methods generally struggle to accurately identify text boundaries. The original CycleGAN shows improvements over traditional methods in terms of background accuracy and weakening text boundaries, but there are still noticeable residual boundaries at the top of the image. In contrast, BambInk can accurately locate text features in the image and strengthen ink edges by removing boundaries.

Figure 6 presents the progressive ink enhancement results at different epochs. It can be observed that at epoch=80, the model reaches its optimal state as the majority of background noise is removed while the clarity of the ink is widely improved. At epoch=100, although the background noise is still well suppressed, the processing of text boundaries deteriorates, due to the instability of GAN. Overall, our model shows a clear enhancement in the inks of the Bamboo Slips.

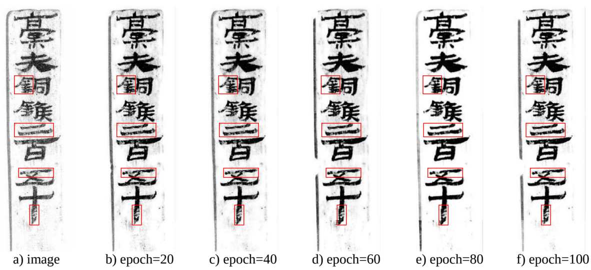

Figure 6: Model outputs at different epochs.

5 Conclusion

This paper explores improving the ancient Chinese manuscripts written on Bamboo Slips through ink enhancement. We proposed BambInk, a self-learning framework Using iterative enhancement/restoration-degradation adversarial learning. BambInk includes an ink enhancer and an ink degrader as generators and two associated discriminators to evaluate the generators' performance through end-to-end training with adversarial loss, cycle-consistency loss, and identity loss. We devise a dynamic convolution block that fuses channel, spatial, and filter attention features to preserve structural stroke integrity, and a high-low frequency differentiating block that actively enhances stroke clarity. Comprehensive experiments validate the superiority of our method over existing methods in ink enhancement.

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-Luc Bloechle, Jean Hennebert, and Christophe Gisler. Zigzag: A robust adaptive approach to non-uniformly illuminated document image binarization. In *Proceedings of the ACM Symposium on Document Engineering 2024*, pages 1–10, 2024.

[2] Songxiao Cao, Zichao Shu, Zhipeng Xu, Dailiang Xie, and Ya Xu. Character segmentation and restoration of qin-han bamboo slips using local auto-focus thresholding method. *Multimedia Tools and Applications*, 81(6):8199–8213, 2022.

[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] Basilis Gatos, Konstantinos Ntirogiannis, and Ioannis Pratikakis. Icdar 2009 document image binarization contest (dibco 2009). In *2009 10th International conference on document analysis and recognition*, pages 1375–1382. IEEE, 2009.

[5] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. *Communications of the ACM*, 63(11):139–144, 2020.

[6] Avekash Gupta, Sunil Kumar, Ronak Gupta, Santanu Chaudhury, and S Joshi. Enhancement of old manuscript images. In *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, volume 2, pages 744–748, 2007.

[7] Zineb Hadjadj, Abdelkrim Meziane, Yazid Cherfa, Mohamed Cheriet, and Insaf Setitra. Isauvola: Improved sauvola's algorithm for document image binarization. In *Image Analysis and Recognition: 13th International Conference, ICIAR 2016, in Memory of Mohamed Kamel, Póvoa de Varzim, Portugal, July 13-15, 2016, Proceedings 13*, pages 737–745. Springer, 2016.

[8] Sana Khamekhem Jemni, Mohamed Ali Souibgui, Yousri Kessentini, and Alicia Fornés. Enhance to read better: a multi-task adversarial network for handwritten document image enhancement. *Pattern Recognition*, 123:108370, 2022.

[9] Hayata Kaneko, Ryuto Ishibashi, and Lin Meng. Deteriorated characters restoration for early japanese books using enhanced cyclegan. *Heritage*, 6(5):4345–4361, 2023.

[10] Seon Joo Kim, Fanbo Deng, and Michael S Brown. Visual enhancement of old documents with hyperspectral imaging. *Pattern Recognition*, 44(7):1461–1469, 2011.

[11] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690, 2017.

[12] 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.

[13] 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.

[14] Wan Azani Mustafa and Mohamed Mydin M Abdul Kader. Binarization of document image using optimum threshold modification. In *Journal of Physics: Conference Series*, page 012022, 2018.

[15] Wayne Niblack. *An introduction to digital image processing*. Strandberg Publishing Company, 1985.

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

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

[18] 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.

[19] João MC Sousa, JR Caldas Pinto, Cláudia S Ribeiro, and João M Gil. Ancient document recognition using fuzzy methods. In *The 14th IEEE International Conference on Fuzzy Systems*, pages 833–838, 2005.

[20] Bolan Su, Shijian Lu, and Chew Lim Tan. Robust document image binarization technique for degraded document images. *IEEE transactions on image processing*, 22(4): 1408–1417, 2012.

[21] 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.

[22] Ruilu Wang, Yang Xue, and Lianwen Jin. DocNLC: A document image enhancement framework with normalized and latent contrastive representation for multiple degradations. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI 2024)*, volume 38, pages 5563–5571, 2024.

[23] Christian Wolf and J-M Jolion. Extraction and recognition of artificial text in multimedia documents. *Formal Pattern Analysis & Applications*, 6:309–326, 2004.

[24] 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 *Proceedings of the 31st ACM international conference on multimedia*, pages 2795–2806, 2023.

[25] Zili Yi, Hao Zhang, Ping Tan, and Minglun Gong. DualGAN: Unsupervised dual learning for image-to-image translation. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV 2017)*, pages 2849–2857, 2017.

[26] 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.

[27] 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.

[28] 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.

[29] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.

[30] Yuanzhi Zhu, Zhaohai Li, Tianwei Wang, Mengchao He, and Cong Yao. Conditional text image generation with diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 14235–14245, 2023.