Binarization of Severely Degraded Ancient Bamboo Slips: Dataset and Baseline

Chongsheng Zhang Wanwan Fu Qilong Li Samra Zafar Zhanshuo Zhang
Qiyan Li Gaojuan Fan Christian Heumann

Abstract

Bamboo and wooden slips were the primary writing materials in ancient China. However, due to the lengthy corrosion or weathering, texts on the bamboo 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. We build HanBamboo, a specialized Bamboo Slips dataset with pixel-level annotations for research on binarization of severely degraded ancient manuscripts. It comprises 1,000 infrared Bamboo Slips exhibiting low-contrast ink traces characterized by significant fading and diminished visibility. As a technical baseline, we propose MRA-Net, which is an Mamba-based Encoder-Decoder framework for degraded Bamboo Slips binarization. Our dataset and code are available at: https://github.com/cszhangLMU/MRA-Net.

Main Contributions

- Construct a specialized benchmark dataset for binarizing severely degraded Bamboo Slips.
- Develop a technical baseline for such a binarization task.

Dataset

We obtain the infrared Bamboo Slips images from domain experts in Bamboo Slips studies. Out of the 7,000 *Xuanquan Han Slips* I∼III Bamboo Slips unearthed in *Gansu* province in northwestern China, we select 1,000 infrared images, and divide them into 4 subsets, which are *minor*, *moderate*, *severe* and *extreme* degradation groups, based on the different levels of ink/text degradation, with varying degrees of legibility (from easy to hard) of the texts. For each subgroup of Bamboo Slips, we equally split it into the training and test sets.

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.

For infrared Bamboo Slips in the *minor degradation* group, we first utilize automated annotation software to produce initial pixel-level annotation results, then manually check and correct a small portion of inaccurate annotations. But for the remaining three groups, we have to manually make all the pixel-level annotations.

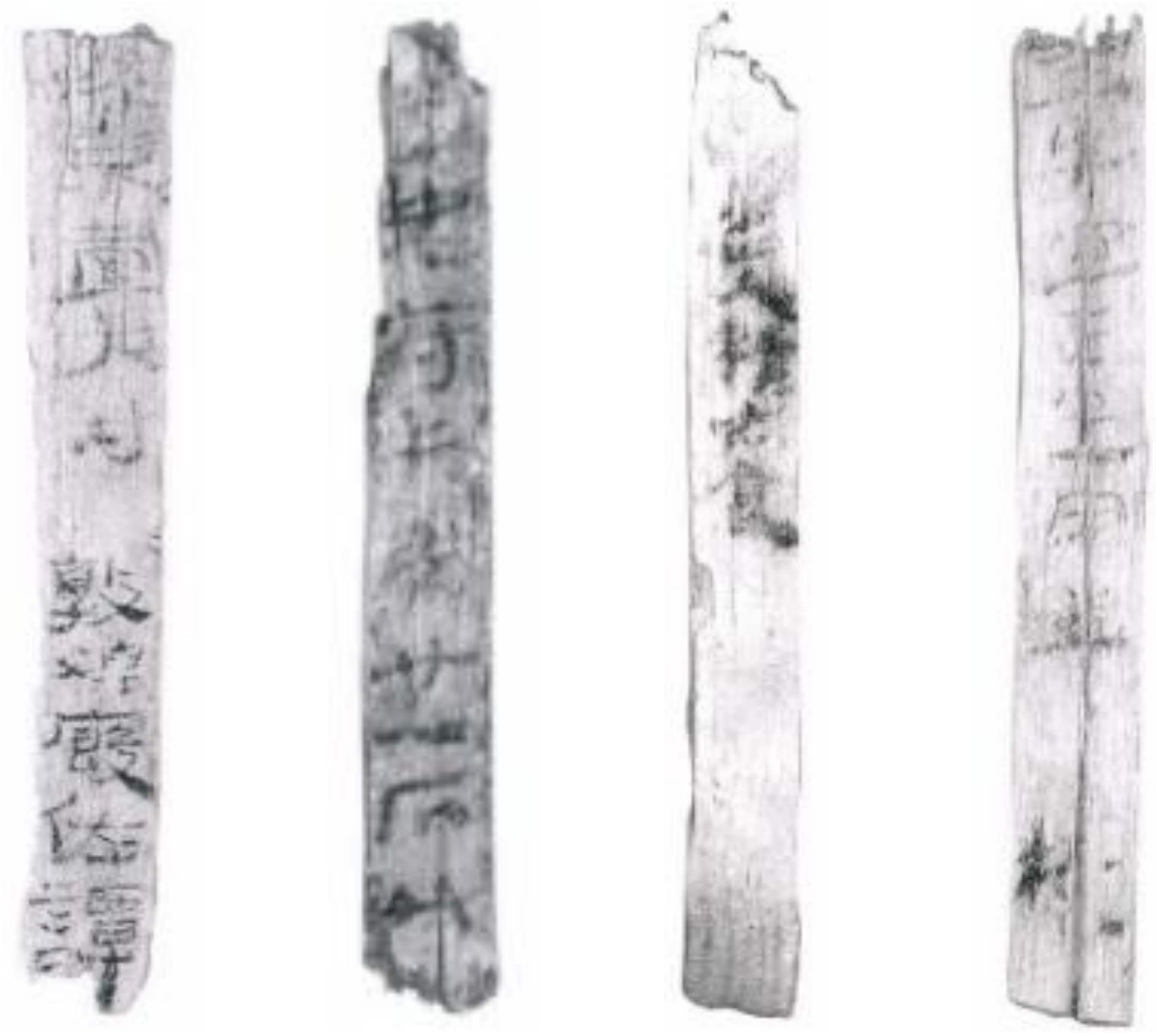

Figure 1. Four severely degraded ancient Bamboo Slips, with challenges such as uneven ink distribution, corrosion of the slip surface, ink bleeding, and fading of the inscriptions.

Figure 2. Comparison between automated annotation results and manually labeling results.

Methodology

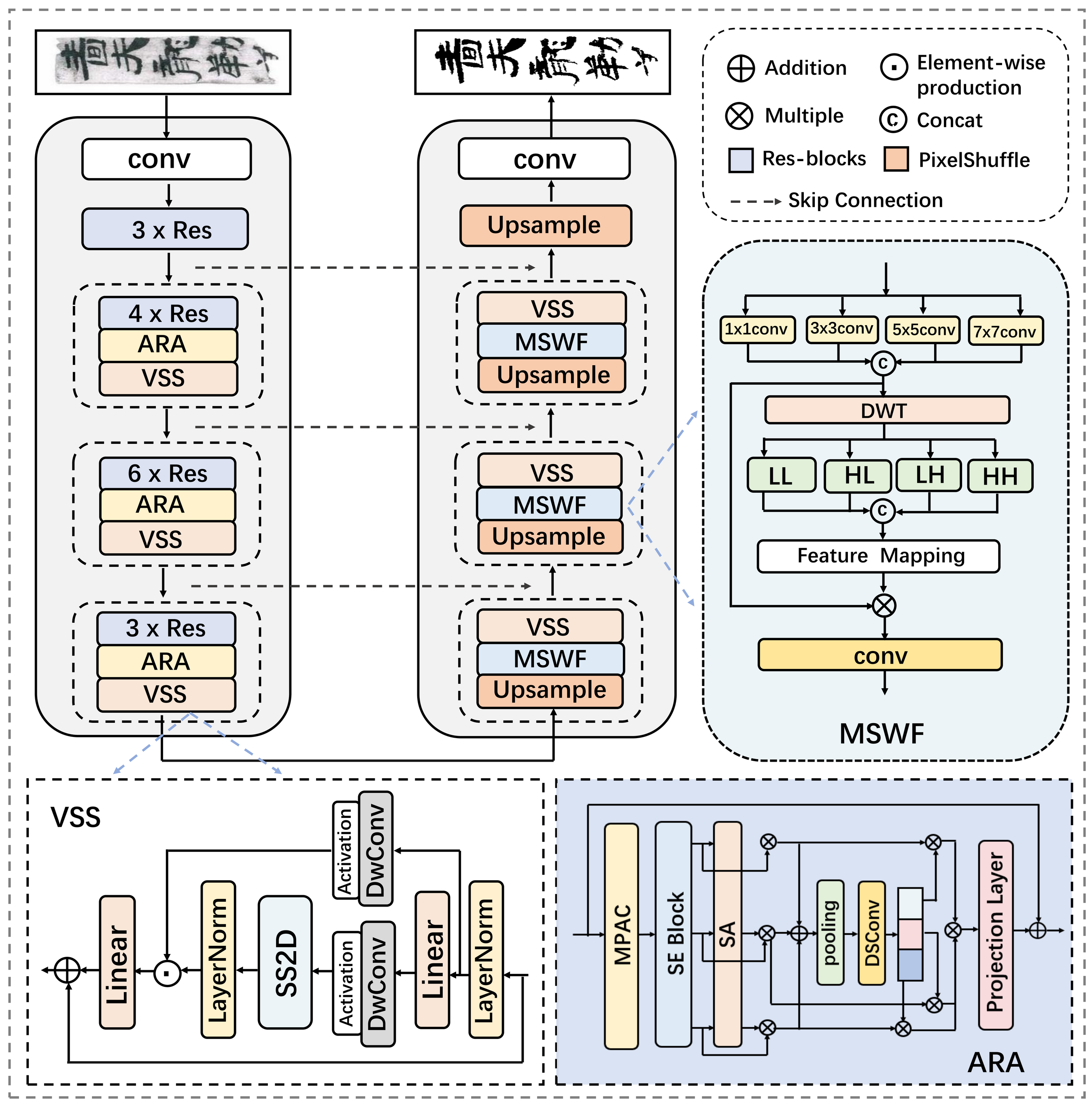

Figure 3. Overall architecture of MRA-Net.

Key Results

Method	FM↑	pFM↑	PSNR↑	DRD↓
Otsu	69.52	70.16	13.72	30.91
Sauvola	68.77	69.63	13.65	24.96
Gatos	73.12	73.60	14.49	16.79
Suh	83.94	82.78	17.27	5.86
DP-LinkNet	83.96	83.89	17.84	4.81
cGANs	83.14	83.26	17.53	5.04
Zhao	81.84	81.52	17.25	6.48
DocEnTr	83.30	83.55	16.43	6.11
FourBi	84.13	84.48	17.81	4.62
MRA-Net (ours)	**84.87**	**85.16**	**17.98**	**4.33**

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

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

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

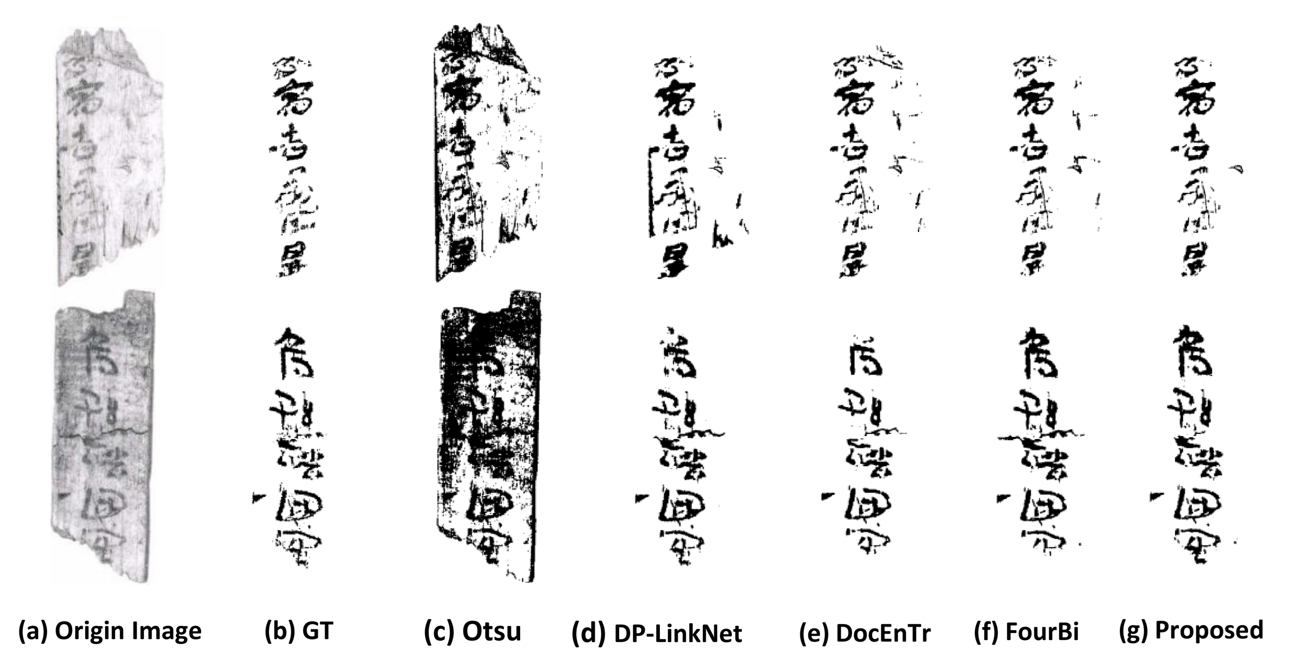

Figure 4. Visual comparisons between MRA-Net and existing methods on HanBamboo.

Conclusions

We present HanBamboo, a specialized Bamboo Slips dataset with high-quality pixel-level annotations, for promoting research in severely degraded Bamboo Slips binarization, which lays the foundation for relevant tasks such as text restoration. We also propose a technical baseline for this binarization task. In future research, we will investigate text impainting on degraded Bamboo Slips.

 Henan University & LMU Munich Email:cszhang@henu.edu.cn