Confidence Score for Unsupervised Foreground Background Separation of Document Images

Soumyadeep Dey, Pratik Jawanpuria
{soumyadeep.dey,pratik.jawanpuria}@microsoft.com

1. Introduction

Foreground-background separation is an important problem in document image analysis. Popular unsupervised binarization methods (such as the Sauvola's algorithm) employ adaptive thresholding to classify pixels as foreground or background. In this work, we propose a novel approach for computing confidence scores of the classification in such algorithms. This score provides an insight of the confidence level of the prediction. The computational complexity of the proposed approach is the same as the underlying binarization algorithm. Our experiments illustrate the utility of the proposed scores in various applications like document binarization, document image cleanup, and texture addition.

2. Computation of scores

The score for each pixel is computed using the following Equations.

$$C_W^b(p) = \begin{cases} \frac{I(p)-T_W(p)}{\max(I)-T_W(p)} & \text{if } I(p) > T_W(p) \\ 1 - \frac{T_W(p)-I(p)}{T_W(p)-\min(I)} & \text{otherwise} \end{cases} \quad (1)$$

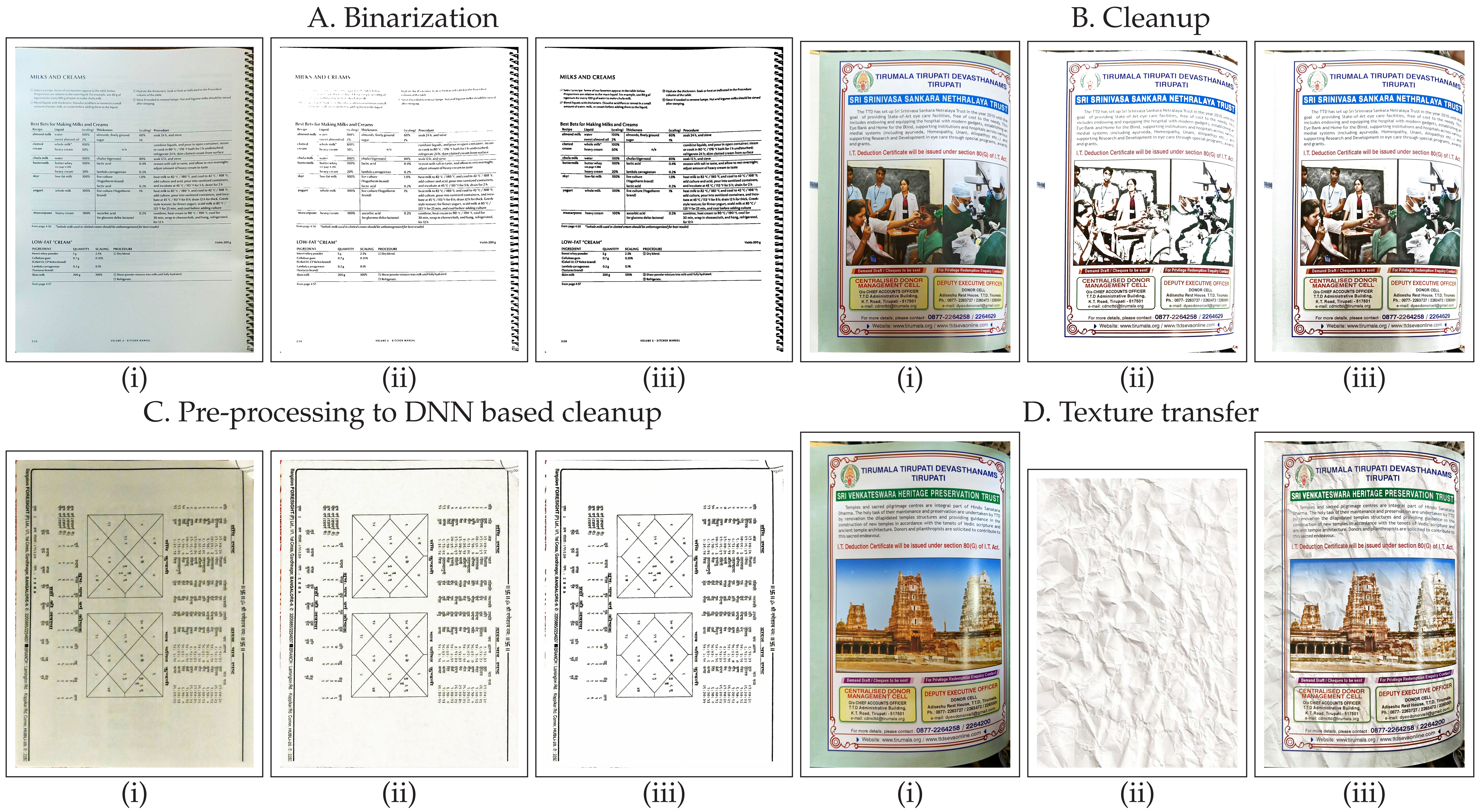
$$C_W^f(p) = 1 - C_W^b(p) \quad (2)$$

Here, max(I) and min(I) represent maximum and minimum value of any pixel of an input image I, respectively. It should be noted that the confidence score lies in the interval $[0, 1]$. The proposed confidence scores can be generated with any adaptive thresholding approach. For empirical comparison, we considered Sauvola's thresholding algorithm [2] as the base method. The threshold is computed for each pixel ($T_W(p)$) using the Eq 3, where, for an input image I, $R = \frac{\max(I)-\min(I)}{2}$.

$$T_W(p) = m_W^p \times [1 + k \times (\frac{s_W^p}{R} - 1)] \quad (3)$$

The threshold is computed for each pixel (p) based on a window W of size $n \times n$ surrounding it, where m_W^p, s_W^p respectively represent mean and standard deviation of W around pixel p, and k lies between $0 \le k \le 1$.

3. Applications of Confidence Scores



(A,B,C,D)(i) input images; A(ii) binary image using [2]; A(iii) improved binary image using the proposed scores; B(ii) image cleanup using [2]; B(iii) image cleanup using the proposed scores; C(ii) image cleanup using [9]; C(iii) image cleanup using [9] with proposed score based pre-processing; D(ii) texture to be transferred to D(i); D(iii) texture transferred image using proposed scores.

4. References

[1] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Trans. on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.

[2] J. Sauvola and M. Pietikäinen. Adaptive document image binarization. *Pattern Recognition*, 33:225–236, 2000.

[3] G. Lazzara, and T. Géraud. Efficient multiscale sauvola's binarization. IJDAR, **17**(2), 105–123. 2014.

[4] X. Peng, H. Cao, K. Subramanian, R. Prasad, and P. Natarajan. Exploiting stroke orientation for crf based binarization of historical documents. ICDAR 2013.

[5] T. Mondal, T. Coustaty, M. Gomez-Krämer, and P. J. Ogier. Learning free document image binarization based on fast fuzzy c-means clustering ICDAR 2019.

[6] G. D. Vo and C. Park. Robust regression for image binarization under heavy noise and nonuniform background. *Pattern Recognition*. **81**, 224 – 239. 2018.

[7] N. Liu, D. Zhang, X. Xu, W. Liu, D. Ke, L. Guo, S. Shi, H. Liu, and L. Chen. An iterative refinement framework for image document binarization with bhattacharyya similarity measure. ICDAR 2017.

[8] X. Peng, C. Wang, and H. Cao Document binarization via multi-resolutional attention model with drd loss. ICDAR 2019.

[9] S. Dey and P. Jawanpuria. Light-weight Document Image Cleanup using Perceptual Loss. ICDAR 2021.