Image Text Enhancer
Algorithm Engineering 2026 Project Paper

Daniel Motz
Friedrich Schiller University Jena
Germany
daniel.motz@uni-jena.de

Leonard Teschner
Friedrich Schiller University Jena
Germany
leonard.teschnner@uni-jena.de

Mher Mnatsakanyan
Friedrich Schiller University Jena
Germany
mher.mnatsakanyan@uni-jena.de

Abstract

The five-finger pattern [**?**]:

(1) **Topic and background:** What topic does the paper deal with? What is the point of departure for your research? Why are you studying this now?
(2) **Focus:** What is your research question? What are you studying precisely?
(3) **Method:** What did you do?
(4) **Key findings:** What did you discover?
(5) **Conclusions or implications:** What do these findings mean? What broader issues do they speak to?

Keywords

noise reduction, background removal, image filter, binarization

1 Introduction

1.1 Background

1.2 Related Work

1.3 Outline

This paper is structured as follows: Section 2 provides a detailed description the developed pipeline and its methods. Section 3 demonstrates the performance of our algorithm using experiments. Finally, Section 4 summarises our results and provides an outlook on possible future work.

2 The Algorithm

Many approaches and best practices already exist for improving the quality of scanned images as seen in section 1.2. We have developed a pipeline that combines several of these methods in order to achieve potentially good results. Users can choose which steps to apply from the pipeline. The individual methods of the pipeline are shown in algorithm 1.

Algorithm 1 Image Text Enhance Pipeline

(1) convert image to grayscale
(2) Deskew (if requested)
(3) Contrast enhancement
(4) Denoising
(5) Binarization
(6) Despeckle (if requested)
(7) Morphological operations (if requested)

The individual methods of the pipeline are explained in more detail below.

2.1 Convert image to graysacle

All pipeline methods work on grayscale images. Therefore, the first step is to convert the input image into a grayscale image. This is achieved by applying the weighted sum $Y = 0.299R + 0.587G + 0.114B$, as defined by the International Telecommunication Union [2], to each pixel. The result is a grayscale image in which the brightness value of each pixel, represented by Y, corresponds to that of the original RGB-pixel. All steps in the pipeline are performed on the converted grayscale image in-place after the conversion.

2.2 Deskew

2.3 Contrast enhancement

2.4 Denoising

2.5 Binarization

Sauvola. [3]

$$T_w = m_w * \left(1 + k * \left(\frac{\sigma_w}{R} - 1\right)\right) - delta \tag{1}$$

$$I(x,y) = \begin{cases} black, & i(x,y) < T_w, \\ white, & textotherwise, \end{cases} \tag{2}$$

$$\tag{3}$$

Bataineh.
Adaptive binarization using local threshold [1]

$$T_w = m_w - \frac{m_W^2 - \sigma_W}{(m_g + \sigma_W) \times (\sigma_{adaptive} + \sigma_W)} \tag{4}$$

$$\sigma_{adaptive} = \frac{\sigma_W - \sigma_{min}}{\sigma_{max} - \sigma_{min}} \tag{5}$$

$$I(x,y) = \begin{cases} black, & \text{if } i(x,y) < T_w \\ white, & \text{otherwise} \end{cases} \tag{6}$$

Adaptive window size [1]

$$T_{con} = m_g - k * \frac{m_g^2 * \sigma_g}{(m_g + \sigma_g) * (0.5max_{level} + \sigma_g)} \tag{7}$$

$$I = \begin{cases} black, & i(x,y) \leq T_{con} - \left(\frac{\sigma_g}{2}\right), \\ red, & T_{con} - \left(\frac{\sigma_g}{2}\right) < i(x,y) < T_{exp} + \left(\frac{\sigma_g}{2}\right), \\ white, & i(x,y) \geq T_{con} + \left(\frac{\sigma_g}{2}\right), \end{cases} \tag{8}$$

$$p = \left(\frac{\text{number of black pixels}}{\text{number of red pixels}}\right) \tag{9}$$

$$PW_{size} = \begin{cases} \left(\frac{I_h}{4}, \frac{I_w}{6}\right), & \geq 2.5 or(\sigma_g < 0.1 * max_{level}), \\ \left(\frac{I_h}{4}, \frac{I_w}{6}\right), & 1 < p < 2 - 5 or(I_h + I_w < 400), \\ \left(\frac{I_h}{4}, \frac{I_w}{6}\right), & p \leq 1, \end{cases} \tag{10}$$

$$SW_{size} = \begin{cases} \left(\frac{PW_h}{2}, \frac{PW_w}{2}\right), & \text{red\_pixel > black\_pixel}, \\ WP_{size}, & \text{otherwise}, \end{cases} \tag{11}$$

2.6 Despeckle

2.7 Morphological operations

3 Experiments

4 Conclusions

References

[1] Bilal Bataineh, Siti Abdullah, Khairuddin Omar, and Mohammad Faidzul Nasrudin. 2011. Adaptive Thresholding Methods for Documents Image Binarization, Vol. 6718. 230–239. doi:10.1007/978-3-642-21587-2_25

[2] International Telecommunication Union. 2011. *Recommendation ITU-R BT.601-7: Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios.* Recommendation BT.601-7. ITU Radiocommunication Sector (ITU-R), Geneva, Switzerland.

[3] J. Sauvola and M. Pietikäinen. 2000. Adaptive document image binarization. https://www.sciencedirect.com/science/article/pii/S0031320399000552. *Pattern Recognition* 33, 2 (2000), 225–236. doi:10.1016/S0031-3203(99)00055-2