

# 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.teschner@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

## 1 Introduction

### 1.1 Background

Motivation: - Viele handschriftliche oder gedruckte Dokumente werden heutzutage mit dem Smartphone, oder im Zuge der Digitalisierung mit Druckern gescannt. Dabei entstehen häufig Bilder mit schlechter Qualität, die schwer lesbar, für Mensch und Maschine sind. Schlechte Qualität ist in diesem Kontext durch schlechte Beleuchtung, Schatten, Verzerrungen, Rauschen oder ungleichmäßigen Kontrast gekennzeichnet. - Solche Bilder sind problematisch, wenn sie weiterverwendet werden sollen, beispielsweise um sie zu archivieren, zu drucken oder mittels OCR in maschinenlesbaren Text umzuwandeln.

### 1.2 Related Work

drei Paper wurden uns empfohlen:

- Adaptive Thresholding Methods for Documents Image Binarization [1]: Das Bild wird in Fenster aufgeteilt. Für jedes Fenster wird ein Thresholdingvalue  $T$  berechnet. In jedem Fenster wird die minimum  $\sigma_{min}$  und maximum  $\sigma_{max}$  Standardabweichungen aller Fenster verwendet, um adaptiv den Threshold  $T$  zu bestimmen. Jeder Pixelwert  $i(x, y)$  des Fensters wird anschliessend mit dem Threshold verglichen: Ist der Wert kleiner wird der Pixel schwarz und sonst weiss.

- Adaptive Thresholding using the Integral Image [2]
- Binarization of historical document images using the local maximum and minimum [4]

### 1.3 Our Contributions

- kostenloses Tool/executable um die Qualität von gescannten Bildern zu verbessern, bevor sie beispielsweise versendet, gedruckt oder archiviert werden. - Verbessert die Arbeit der Related Work, durch ... - Kann beispielsweise für die Digitalisierung von Büchern, Dokumenten und handschriftlichen Akten oder Notizen genutzt werden. So wird eine gute, digitale Lesbarkeit und Wiederverwendung ermöglicht. - OpenMP und C++ Implementierung mit CMake Build System.

### 1.4 Outline

Dieses Paper ist wie folgt gegliedert: In Abschnitt 2 beschreiben wir den entwickelten Algorithmus. In Abschnitt 3 zeigen wir die Performance unseres Algorithmus anhand von Experimenten. Abschließend fassen wir in Abschnitt 4 unsere Ergebnisse zusammen und geben einen Ausblick auf mögliche zukünftige Arbeiten.

## 2 The Algorithm

Es gibt bereits viele Ansätze und bewährte Verfahren, um die Qualität von gescannten Bildern zu verbessern. Wir haben uns entschieden, eine Pipeline zu entwickeln, die verschiedene dieser Verfahren kombiniert, um ein möglich gutes Ergebnis zu erzielen. Es ist dem Nutzer überlassen, welche Schritte der Pipeline er anwenden möchte. Die einzelnen Verfahren der Pipeline sind im Algorithmus 1 dargestellt.

---

#### 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)
- 

Im Folgenden werden die einzelnen Schritte der Pipeline näher erläutert:

### 2.1 Convert image to grayscale

Alle Verfahren der Pipeline arbeiten auf Graustufenbildern. Daher ist der erste Schritt der Pipeline die Umwandlung des Eingabebildes in ein Graustufenbild. Dies wird durch eine gewichtete Summe  $Y = 0.299R + 0.587G + 0.114B$  [3] der RGB-Kanäle für jeden Pixel erreicht. Das Ergebnis ist ein Graustufenbild, in dem jeder Pixelwert  $Y$  den Helligkeitswert des entsprechenden Pixels im Originalbild darstellt. Alle Schritte der Pipeline erfolgen nach der Konvertierung in-place, auf dem konvertierten Graustufenbild.

## 2.2 Deskew

## 2.3 Contrast enhancement

## 2.4 Denoising

## 2.5 Binarization

*Sauvola.*

*Bataineh.*

$$T = m_w - \frac{m_w^2 - \sigma_w}{(m_g + \sigma_w) \times (\sigma_{adaptive} + \sigma_w)} \quad (1)$$

$$\sigma_{adaptive} = \frac{\sigma_w - \sigma_{min}}{\sigma_{max} - \sigma_{min}} \quad (2)$$

$$I(x, y) = \begin{cases} \text{black}, & \text{if } i(x, y) < T \\ \text{white}, & \text{otherwise} \end{cases} \quad (3)$$

$$T_{con} = m_g - k * \frac{m_g^2 * \sigma_g}{(m_g + \sigma_g) * (0.5max_{level} + \sigma_g)} \quad (4)$$

$$I(x, y) = \begin{cases} \text{black}, & i(x, y) \leq T_{con} - \left(\frac{\sigma_g}{2}\right), \\ \text{red}, & T_{con} - \left(\frac{\sigma_g}{2}\right) < i(x, y) < T_{exp} + \left(\frac{\sigma_g}{2}\right), \\ \text{white}, & i(x, y) \geq T_{con} + \left(\frac{\sigma_g}{2}\right), \end{cases} \quad (5)$$

$$p = \left( \frac{\text{number of black pixels}}{\text{number of red pixels}} \right) \quad (6)$$

$$PW_{size} = \begin{cases} \left( \frac{I_h}{4}, \frac{I_w}{6} \right), & \geq 2.5or(\sigma_g < 0.1 * max_{level}), \\ \left( \frac{I_h}{4}, \frac{I_w}{6} \right), & 1 < p < 2 - 5or(I_h + I_w < 400), \\ \left( \frac{I_h}{4}, \frac{I_w}{6} \right), & p \leq 1, \end{cases} \quad (7)$$

$$SW_{size} = \begin{cases} \left( \frac{PW_h}{2}, \frac{PW_w}{2} \right), & \text{red\_pixel} > \text{black\_pixel}, \\ WP_{size}, & \text{otherwise}, \end{cases} \quad (8)$$

## 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] Derek Bradley and Gerhard Roth. 2007. Adaptive Thresholding using the Integral Image. *J. Graphics Tools* 12 (01 2007), 13–21. doi:10.1080/2151237X.2007.10129236
- [3] 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.
- [4] Bolan Su, Shijian Lu, and Chew Lim Tan. 2010. Binarization of historical document images using the local maximum and minimum. In *Proceedings of the 9th IAPR International Workshop on Document Analysis Systems* (Boston, Massachusetts, USA) (*DAS '10*). Association for Computing Machinery, New York, NY, USA, 159–166. doi:10.1145/1815330.1815351