

Dokumenten Rekonstruktion

Alina Mokra (01325014), Viktoria Petersen (11924496),
Helene Weninger (0809236), Matthias Zezulka (11914298)

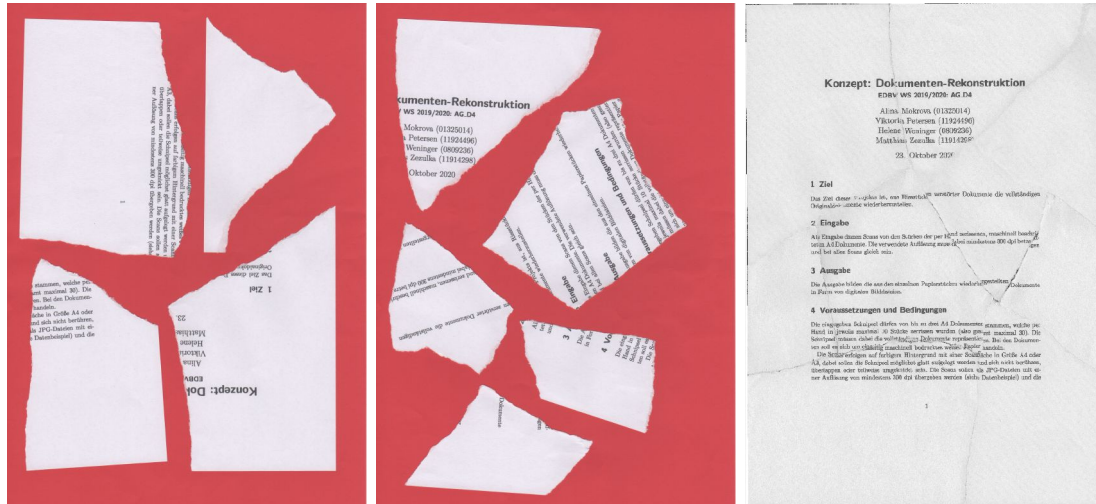


Abbildung 1: Input der Scans als PNGs und Output der Rekonstruktion

Projekt

Das Ziel des Projekts ist es ein maschinell beschriebenes und von Hand zerissenes Dokument über dessen eingescante Schnipsel zu rekonstruieren und wieder lesbar zu machen.

Vorgangsweise

Methodik-Pipeline

1. Connected Component Labeling der Schnipsel und anschließendes Trennen anhand der Bounding Box.
2. Erosion und anschließende Subtraktion vom Original zur Isolierung der Randpixel.
3. Hough Transformation und anhand der geraden Linien Randstücke, Eckstücke und Innenstücke trennen, sowie Rotation aller Rand- und Eckstücke.
4. Distanztransformation zur Identifizierung der Ecken und Trennung der Kanten.
5. Randstücke werden anhand der Standardabweichung der Abstände ihrer einander zugewandten Kanten zusammengesetzt. Es ergeben sich die Ränder der Dokumente.
6. Projektion und Analyse der Innenteile und des fertig zusammengesetzten Außenteils um Zeilenrichtung zu identifizieren und Ausrichtung zu finden
7. Innenstücke werden nach der Projektion in das Außenteil eingefügt. Dabei wird das Außenteil mit den Innenstücken gefaltet, um die optimale Position zu finden.
8. Gegebenfalls nicht einsetzbare Randstücke abschließend ebenfalls mit der Projektion und Faltung einsetzen.
9. Rekonstruiertes Dokument ausgeben.

Ergebnisse

Ergebnis in Abbildung 1.