

Report der automatischen Bildqualitätsanalyse		
Merkmal	Anzahl	Hinweis
1. Bilder entdeckt (Layout Parser)	4	Können sowohl Rastergrafiken als auch Vektorgrafiken sein
2. Bilder entdeckt (Analyse PDF Struktur)	4	Nur Rastergrafiken
3. Bilder im JPEG Format	2	Können aufgrund von Kompressionen eine geringe Qualität aufweisen
4. Bilder mit weniger als 300 PPI	4	Empfohlene Anzahl für Printmedien
5. Bilder zu nah am Rand	1	Werden ggf. nicht vollständig gedruckt
6. Bilder mit hohem Anteil an Blockartefakten	1	Bildqualität kann durch Blockartefakte stark eingeschränkt sein
7. Bilder mit Rechtschreibfehlern	3	Überprüfen Sie die annotierten Bilder auf Rechtschreibfehler
8. Bilder mit Metadaten	0	Metadaten können sensible Informationen enthalten (GPS Koordinaten etc.)
9. Bilder im Internet gefunden	0	Beachten Sie eine richtige Quellenangabe
10. Bilder mit kleiner Schriftgröße (< 8px)	3	Beachten Sie eine adequate Schriftgröße für eine bessere Leserlichkeit
11. Bilder mit hoher Dateigröße (> 3000 KB)	0	Vielzahl an großen Bildern kann die PDF Dateigröße sehr groß werden lassen
12. Bilder mit geringem Farbkontrast (< 3 nach WCAG Standard)	1	Achten Sie auf eine gute Farbwahl

# Testdokument zur automatischen Qualitätsanalyse von Abbildungen in Publikationen und Abschlussarbeiten

An der Fachhochschule Dortmund  
im Fachbereich Informatik

von

**Daniel Schall**

## Vorwort

Dieses Dokument soll das Layout einer typischen Abschlussarbeit nachstellen und enthält mehrere Abbildungen, welche Fehler beinhalten. Ziel ist es, einen ersten Prototyp zu erstellen, welcher zur Erkennung dieser Fehler Ansätze aufzeigt. Die jeweiligen Fehler der Abbildungen werden in diesem Testdokument beschrieben. Die Ergebnisse aus den resultierenden Tests in Zusammenhang mit diesem Testdokument, werden in der dazugehörigen Bachelorarbeit näher beschrieben. Anzumerken ist, dass die Inhalte der Abbildungen (Wahl der dargestellten Daten) keine Relevanz haben und ohne Bezug zur Realität gewählt worden sind. Die Abbildungen wurden mithilfe der Programmiersprache Python und der Bibliothek „Matplotlib“<sup>1</sup> erstellt. Dieses Dokument wurde in Microsoft Word (Version 2208) erstellt und als PDF-Datei exportiert.

---

<sup>1</sup> Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. Computing in Science & Engineering 9 (3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>.

## Abbildungen und Erklärungen

Die Nachfolgende Abbildung 1 ist die erste Abbildung für den Prototypentest. Sie weist verschiedene Qualitätsmängel auf, die folgend beschrieben sind.

Der erste Mangel ist die zu kleine Schriftgröße der Überschrift. Beim Erstellen der Grafik in Matplotlib wurde eine Schriftgröße von 8 Punkten gewählt. Aufgrund von Größenanpassungen bzw. Skalierungen von Grafiken zur Anpassung an das Dokument können diese Schriftarten noch kleiner (bei Verkleinerung der Grafik) oder größer (bei Vergrößerung der Grafik – eventuell mit einhergehen von Unschärfen) ausfallen. Ein weiterer Mangel ist ein Rechtschreibfehler in der angezeigten Legende („Deutshcland“ anstelle korrekterweise „Deutschland“). Der nächste anzumerkende Punkt ist das Fehlen der Beschriftung der X-Achse. Auch die Positionierung der Abbildung ist zu nah am linken, sowie rechten Rand. Der letzte visuelle anzusprechende Punkt ist die Farbwahl der Linien im Diagramm. Diese werden mit den Farben „Rot“ und „Gelb“ dargestellt. Die Farbe Gelb ist auf dem weißen Hintergrund nur schwer erkennbar.

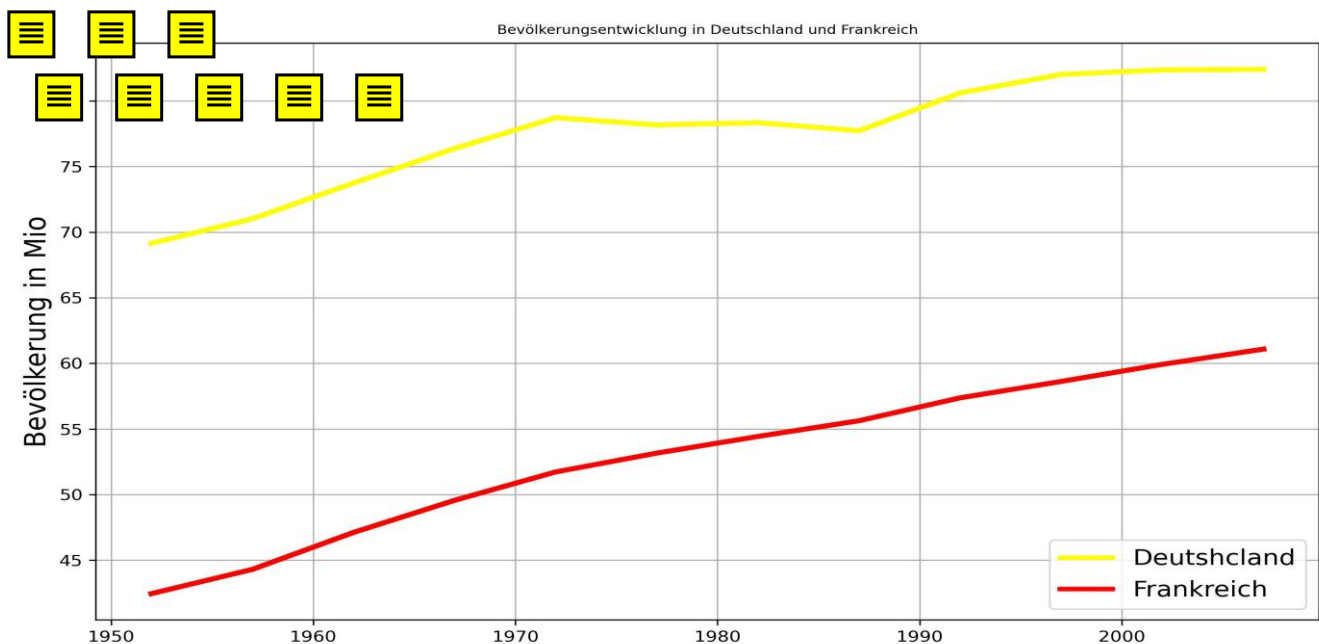


Abbildung 1 Erste Testabbildung mit verschiedenen Qualitätsmängel

Weitere Eigenschaften der Abbildung 1 die nicht unbedingt als Mängel einzustufen sind, dennoch Potential zur Verbesserung bieten, sind das Dateiformat und die Metadaten der Abbildung. Diese können über eine genaue Untersuchung der direkten Einbettung in der PDF-Datei ermittelt werden. Bei der Abbildung handelt es sich um das JPEG-Format. Dieses ist ein verlustbehaftetes Format, d.h. bei Kompressionen kann es zu Qualitätsverlusten kommen. Die Metadaten der Abbildung 1 beinhalten konkrete Angaben über den Ersteller und das Urheberrecht.

Ein weiterer Test ist die Erkennung von Linienstärken in Diagrammen. Für diesen Test wurde Abbildung 2 im PNG-Format angefertigt. Hingegen den Mängeln von Abbildung 1 ist in dieser Abbildung zu sehen, dass die Linienstärke der blauen Linie zu gering ist und Schwierigkeiten beim Lesen verursachen kann. Vor allem beim Übergang zum Printmedium kann diese geringe Linienstärke zu weiteren Problemen führen.

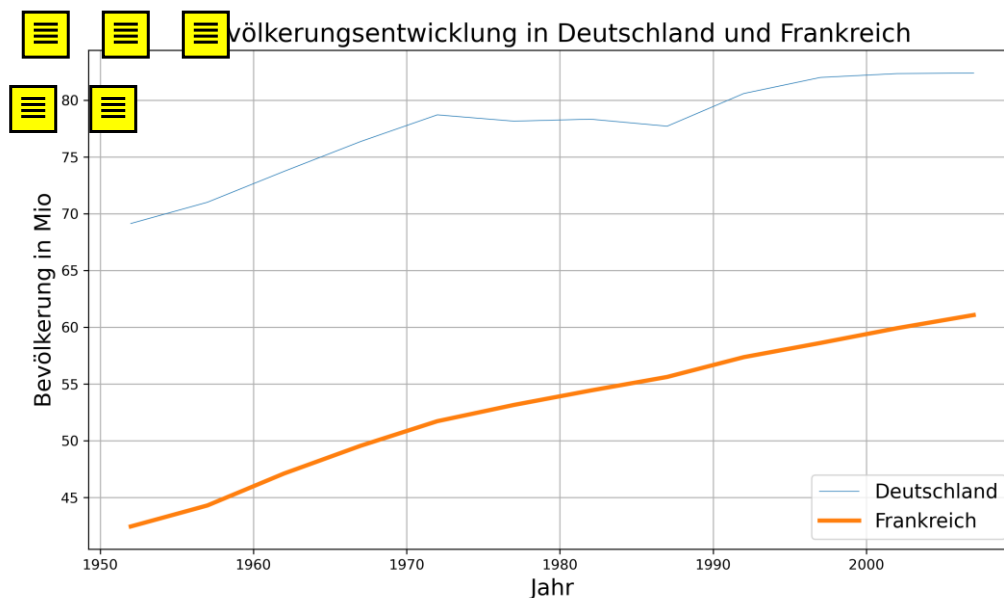


Abbildung 2 Zweite Testabbildung mit geringer Linienstärke

Die nächste Testabbildung (Abbildung 3) ist analog zur Abbildung 2, jedoch sind beide Linienstärken in dieser Abbildung gleich und es handelt sich um das JPGE-Format. Auffallend sind vor allem die starken Kompressionsartefakte, die sich besonders um die Schriften und Linien gebildet haben. Diese Artefakte entstehen bei einer zu hohen Kompression. Zur Erzeugung der Kompression und den Artefakten wurde die Webseite „<https://onlinejpgtools.com/>“<sup>2</sup> benutzt.

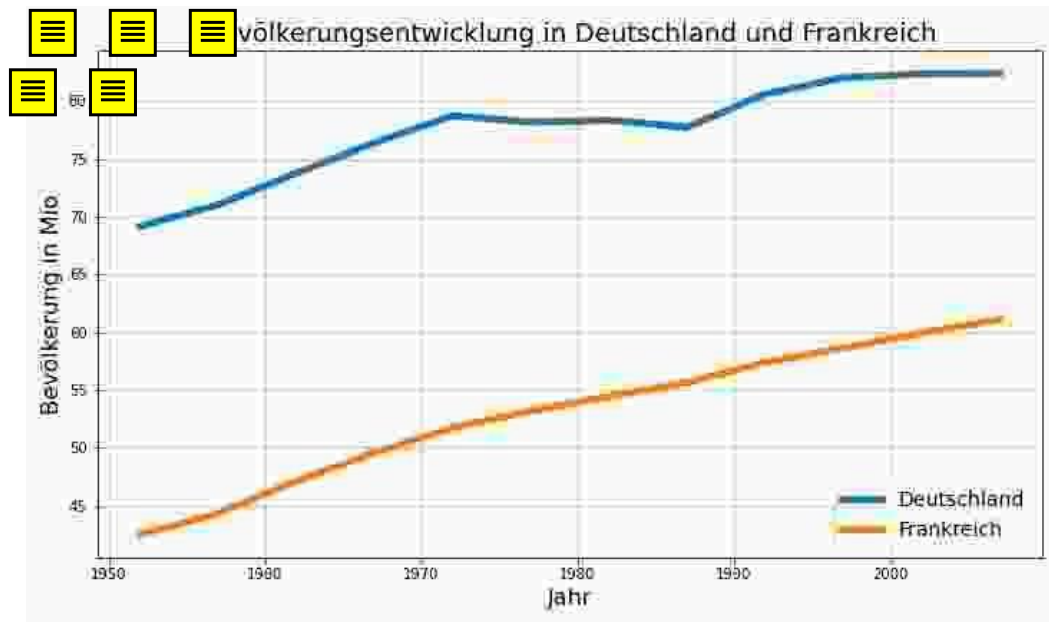


Abbildung 3 Testabbildung mit Kompressionsartefakten

Die Abbildung (Abbildung 4) wurde vom „Statistischen Bundesamt (Destatis)“<sup>3</sup> hinzugezogen. Diese dient dem Test zur Ermittlung von Abbildungen, die aus dem Internet stammen. Dafür wird eine sogenannte „umgekehrte Bildsuche“ durchgeführt.

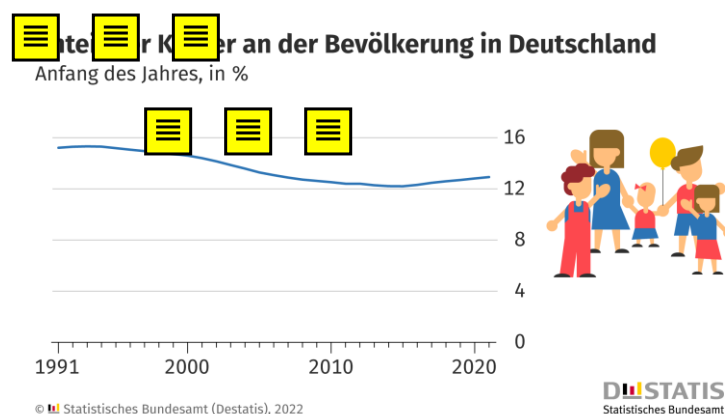
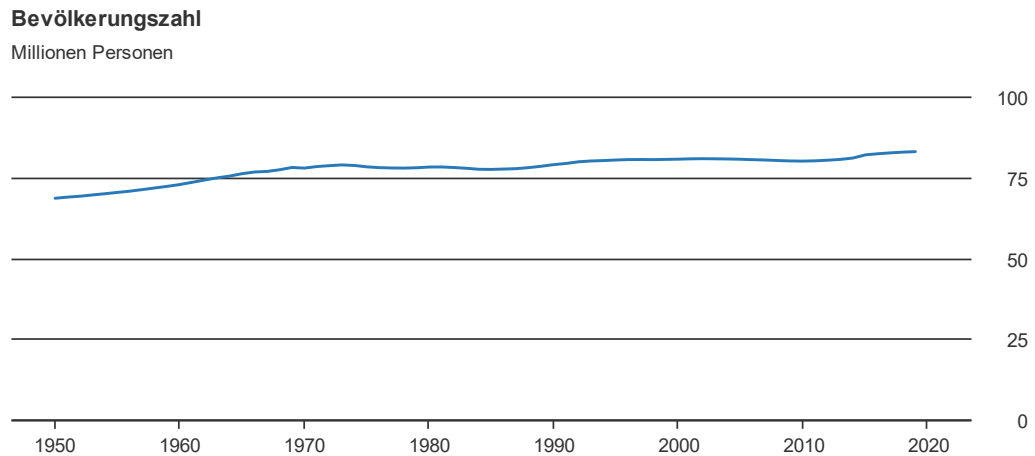


Abbildung 4 Testabbildung für die umgekehrte Bildsuche

<sup>2</sup> Erzeugung der Kompression mit „onlinejpgtool“ unter „<https://onlinejpgtools.com/add-jpg-artifacts>“. Letzter Abruf 25.09.2022

<sup>3</sup> Abbildung 4 abgerufen unter „[https://www.destatis.de/DE/Presse/Pressemitteilungen/Zahl-der-Woche/2022/PD22\\_22\\_p002.html](https://www.destatis.de/DE/Presse/Pressemitteilungen/Zahl-der-Woche/2022/PD22_22_p002.html)“. Letzte Abruf 25.09.2022

Die letzte Abbildung 5 wurde als Vektorgrafik (SVG-Format) erstellt und eingebunden. Auch diese wurde vom „Statistischen Bundesamt (Destatis)“<sup>4</sup> hinzugezogen. Hier soll getestet werden, ob Vektorgrafiken erkannt werden können.



Zwischen 1991 und 2010 rückgerechnete Werte auf Basis des Zensus 2011.

© Statistisches Bundesamt (Destatis), 2022

Abbildung 5 Testabbildung erstellt als Vektorgrafik

<sup>4</sup> Abbildung 5 abgerufen unter „<https://www.destatis.de/DE/Themen/Querschnitt/Demografischer-Wandel/textbaustein-taser-blau-bevoelkerungszahl.html?nn=238640>“. Letzte Abruf 03.10.2022