





## Die Kunst des Lesens aus Linien

Ein neuer Blick in die Tiefe technischer Zeichnungen: Ein Deep-Learning-Framework liest CAD-Grafiken so genau wie nie zuvor – erkennt Formtoleranzen, Maße und jedes Detail. Was früher mühsam von Hand geschah, übernimmt nun eine KI, die sich in die Besonderheiten jeder Linie und Beschriftung einfühlt. Diese vielversprechende Technologie steigert nicht nur die Genauigkeit, sondern beschleunigt auch die Verarbeitung von Zeichnungen erheblich. So eröffnet das System neue Wege für Präzision in der Produktion.

### Schlüsselwörter

Deep Learning, Texterkennung (OCR), technische Zeichnungen, Qualitätskontrolle, Objekterkennung, Produkt- und Fertigungsinformationen

**Dr. Xiao Zhao** arbeitet als Postdoktorand am Institute for Machine Learning and Analytics (IMLA) der Hochschule Offenburg und als PLM-Lösungsarchitekt für KI-Anwendungen bei der BCT Technology AG.

**Marko Weber, DI**, ist Leiter der Softwareentwicklung und des Produktmanagements bei der BCT Technology AG.

**Jan Schöffmann, BS**, arbeitet als Softwareentwickler bei der BCT Technology AG.

**Prof. Daniela Oelke** ist Professorin für Maschinenlernen an der Hochschule Offenburg, Deutschland, und Mitglied des Institute for Machine Learning and Analytics (IMLA).

### Kontakt

[xiao.zhao@bct-technology.com](mailto:xiao.zhao@bct-technology.com)

<https://bct-technology.com/>

# Digitalisierung von Raster-Zeichnungen mit Deep Learning

## Wie ein Framework OCR-Software bei der Extraktion von Daten übertrifft

Xiao Zhao, BCT und Hochschule Offenburg, Marko Weber, BCT Technology AG,  
Jan Schöffmann, BCT und Daniela Oelke, Hochschule Offenburg

Die automatische Extraktion von Produkt- und Fertigungsinformationen (Product Manufacturing Information - PMI) aus technischen CAD-Zeichnungen ist eine Voraussetzung für die Fertigung und Qualitätskontrolle in der Produktion. Aufgrund des speziellen Stils von CAD-Zeichnungen und der begrenzten Verfügbarkeit von Trainings- und Testdaten bleibt die Digitalisierung von CAD-Zeichnungen in Rasterbildern eine Herausforderung für Texterkennungssoftware (Optical Character Recognition - OCR). Dieser Beitrag stellt ein neuartiges, auf Deep Learning basierendes Framework vor, das dieses Problem adressiert, indem es Form- und Lagetoleranzen (Geometrical Dimensioning and Tolerancing - GD&T) sowie Bemaßungen in CAD-Zeichnungen lokalisiert und erkennt. Das Framework besteht aus einem zentralen Lokalisierungsmodul und mehreren nachgelagerten Pipelines für einzelne Klassen von PMI. Die Leistungsfähigkeit des Lokalisierungsmoduls, des Netzwerks zur Zeilenerkennung und der einzelnen Pipelines wird anhand realer Datensätze evaluiert. Ihre Leistung wird mit der des OCR-Programms Tesseract verglichen.

sierte Zeichnungen aus, da deren Inhalte mit traditionellen IT-Technologien ohne den Einsatz von OCR-Werkzeugen extrahiert werden können.

Aus historischen Gründen gibt es immer noch eine beträchtliche Anzahl an technischen Zeichnungen in Form von Rasterbildern, die in der industriellen Praxis verwendet werden. Laut [3] werden jährlich 250 Millionen neue Zeichnungen erstellt, und Millionen von älteren Zeichnungen sind weiterhin in Umlauf.

Technische Zeichnungen sind detaillierte, präzise 2D-Darstellungen von Werkstücken oder mechanischen Systemen und wesentliche Werkzeuge für die Gestaltung, Entwicklung und Fertigung mechanischer Teile. Technische Zeichnungen bieten eine visuelle Anleitung, die nicht nur geometrische Informationen wie Formen und Geometrien, sondern auch Textinformationen wie Bemaßungen, Toleranzen, anwendbare Vorschriften (z. B. ISO- oder DIN-Normen) sowie Wärme- und Oberflächenbehandlungen enthält. Obwohl es einen Trend gibt, solche Produkt- und Fertigungsinformationen (PMI) in 3D-CAD-Modelle zu integrieren, werden in der aktuellen industriellen Praxis immer noch 2D-technische Zeichnungen verwendet. Sie dienen als Blaupause für Ingenieur-, Fertigungs- und Qualitätskontrollaufgaben bei groß angelegten Maschinenbauprojekten.

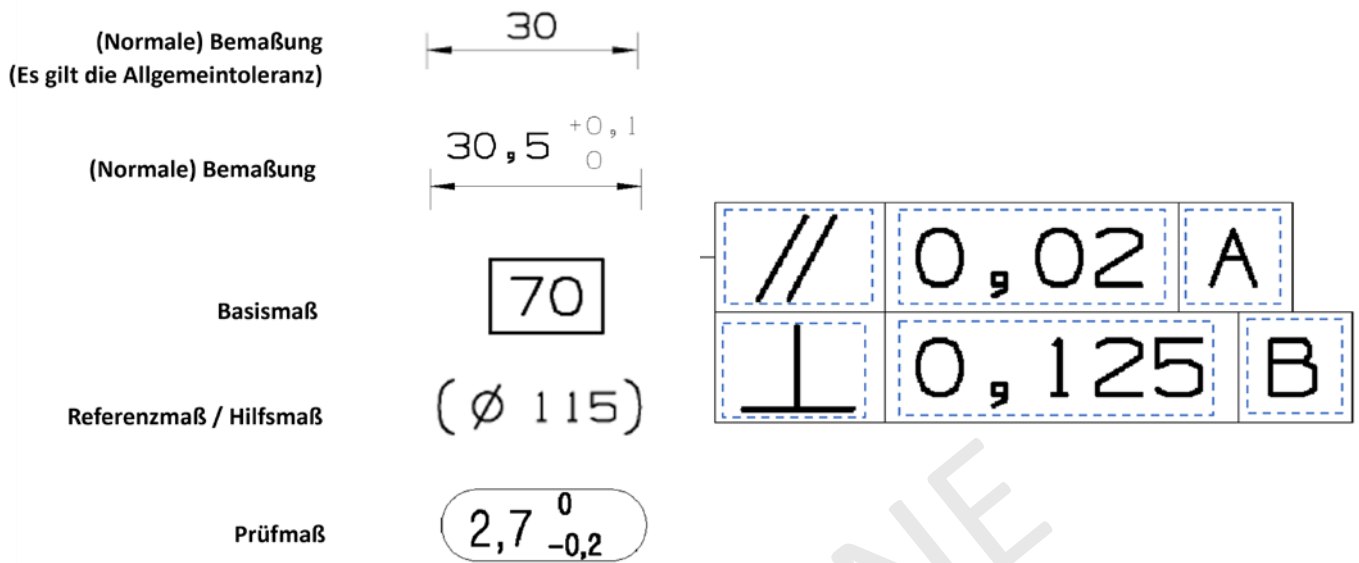
Es gibt drei Arten von technischen Zeichnungen: vektorisierte Zeichnungen, hochqualitative Rasterzeichnungen und niedrigqualitative Rasterzeichnungen [1]. Fortschritte in der CAD-Software ermöglichen es Ingenieuren, vektorisierte Zeichnungen digital zu erstellen, zu modifizieren und zu analysieren. Hochqualitative Rasterzeichnungen enthalten nur Pixelinformationen. Sie werden von CAD-Software generiert, wobei zusätzliche Informationen, wie Toleranzen und Normen, manuell hinzugefügt werden müssen [2]. Niedrigqualitative Rasterzeichnungen sind ältere, papierbasierte Zeichnungen, die durch Einscannen digitalisiert wurden. Diese Arbeit konzentriert sich auf hochqualitative Rasterzeichnungen und schließt vektorisierte

Diese technischen Zeichnungen enthalten nur Pixelinformationen, und es bleibt eine Herausforderung für typische CAD-Software, geometrische und textuelle Informationen wie Bemaßungen, Toleranzen und Wärmebehandlungen aus ihnen zu extrahieren. Eine Lösung, die die Extraktion dieser Informationen aus technischen Zeichnungen ermöglicht, könnte den gesamten Produktionsprozess, einschließlich Mess- und Qualitätskontrollaufgaben, automatisieren [2].

## Digitalisierung von Rasterzeichnungen

Um zu verstehen, was es bedeutet, Rasterzeichnungen zu digitalisieren, ist ein Verständnis des Aufbaus dieser technischen Zeichnungen erforderlich. Wie bereits erwähnt, enthalten die Zeichnungen Produkt- und Fertigungsinformationen (PMI) zusätzlich zu den geometrischen Informationen. Die Digitalisierung technischer Zeichnungen bedeutet daher, dass sowohl der Typ des Informationsobjekts (wie Bemaßungen, Toleranzen, Informationen zu Materialien usw.) als auch deren Inhalt extrahiert werden muss. Das Ziel besteht darin, die technische Zeichnung in ein strukturiertes Format (z. B. eine JSON- oder XML-Datei) zu konvertieren, das dann in nachgelagerten CAD/CAM-Softwaretools geladen und verarbeitet werden kann.

Bild 1 zeigt Beispiele für zwei wichtige Arten von PMI: Bemaßungen sowie Form- und Lagetoleranzen (GD&T). Wie zu sehen ist, kann das Aussehen der Bemaßungen



**Bild 1:** Beispiele für zwei wichtige Arten von Produkt- und Fertigungsinformationen.

variieren, je nachdem, ob Standard- oder explizite Toleranzen gelten. GD&T hingegen wird immer in einem tabellenähnlichen Format dargestellt, das aus einer oder mehreren Zeilen und Spalten bestehen kann. Zellen in GD&T enthalten Symbole, Nennwerte und Toleranzen und manchmal einen Verweis. Dieser Artikel konzentriert sich auf diese beiden Arten von PMI. Das vorgeschlagene Framework kann jedoch erweitert werden, um auch andere Arten von PMI zu erkennen.

## Konventionelle OCR-Software funktioniert nicht gut

Eine offensichtliche Lösung wäre, einfach Standard-OCR-Software für diese Aufgabe zu verwenden. Allerdings wird OCR-Software normalerweise mit normalem Textmaterial wie Büchern, Zeitungen, handschriftlichen Dokumenten usw. trainiert und nicht mit technischen Zeichnungen. Sind mit technischen Zeichnungen daher gute Ergebnisse zu erwarten?

Um diese Frage zu untersuchen, wurden Experimente mit OCR-Software durchgeführt. In einem Vorversuch schnitt Tesseract OCR [4] am besten ab, weshalb diese Softwarelösung im Detail für die Aufgabe untersucht wurde.

Typischerweise müssen zwei Schritte durchgeführt werden, wenn Zeichnungen digitalisiert werden: Zuerst muss das Dokument segmentiert werden, d. h. die Bereiche, die relevante Objekte enthalten (z. B. GD&T oder Bemaßungen), müssen lokalisiert werden. Anschließend müssen die einzelnen Zeichen, die in diesem Bereich abgebildet sind, erkannt und extrahiert werden.

Tesseract OCR folgt diesem zweistufigen Ansatz. In diesem Test wird die Software daher anhand von drei verschiedenen Aufgaben evaluiert:

1. Lokalisierung: Wie gut gelingt es Tesseract, die Bereiche von GD&T oder Bemaßungen, zwei der am häufigsten verwendeten PMI in Qualitätskontrollaufgaben, zu finden?
2. Zeilenerkennung: Wie gut gelingt es Tesseract, die Zeichen innerhalb der entsprechenden Bildbereiche zu erkennen?
3. Lokalisierung + Zeilenerkennung: Wie gut schneidet Tesseract bei der Gesamtaufgabe ab?

Insgesamt standen 14 echte technische Zeichnungen für die Auswertung zur Verfügung. Für jede Aufgabe wurde ein Datensatz erstellt, bei dem menschliche Annotatoren die korrekte Lösung bestimmt haben (genannt „Ground-Truth“-Datensatz). Da Tesseract keine separaten Klassen für GD&T oder Bemaßungen speziell für Zeichnungen hat, musste eine heuristische Regel angewendet werden, um das Ergebnis von Tesseract mit der „Ground-Truth“ zu vergleichen. Die heuristische Regel prüft die „Intersection over Union“ (IoU) jeder vorhergesagten Begrenzungsbox mit allen „Ground-Truth“-Begrenzungsboxen. IoU ist eine Metrik zur Messung der Lokalisierungsgenauigkeit vorhergesagter Begrenzungsboxen. Wenn die maximale IoU größer ist als ein Schwellenwert (hier auf 0,05 festgelegt), übernimmt die vorhergesagte Begrenzungsbox die Klasse der „Ground-Truth“-Begrenzungsbox, mit der sie die maximale IoU hat. Andernfalls wird die vorhergesagte Begrenzungsbox als „Hintergrund“ betrachtet.



	Messgrößen	Daten	Tesseract			Framework		
			GD&T	Bemaßung	Wert	GD&T	Bemaßung	Wert
Lokalisierung	Präzision	14 echte Zeichnungen:	0.852	0.717	-	0.822	0.696	-
	Recall	Pro Zeichnung: 4984*3531 bis 13200*10200 Pixel	0.294	0.258	-	0.862	0.565	-
	F1		0.437	0.379	-	0.841	0.624	-
	IoU		0.279	0.234	-	0.726	0.453	-
Zeilenerkennung	CER	1195 Zeilen	-	-	0.084	-	-	0.088
	WER		-	-	0.290	-	-	0.358
	CER	216 GD&T-Zeilen	-	-	0.512	-	-	0.135
	WER		-	-	0.674	-	-	0.243
Lokalisierung + Zeilenerkennung	CER	10 echte Zeichnungen:	0.911	1.036	-	0.161	0.705	-
	Präzision	497 Zeichen bei GD&T	0.411	0.353	-	0.876	0.591	-
	Recall	3209 Zeichen bei Bemaßung	0.350	0.183	-	0.923	0.473	-
	F1		0.378	0.241	-	0.899	0.525	-

**Bild 2:** Evaluationsergebnisse.

Bild 2 fasst die Auswertungsergebnisse zusammen. Für Aufgabe 1 (Lokalisierung) werden Precision, Recall, F1 und IoU auf Pixelebene der Lokalisierungsergebnisse als Metriken verwendet, um die Leistung zu messen. Precision misst, welcher Anteil der zugewiesenen Kennzeichnungen richtig ist, während Recall misst, welcher Anteil der Objekte, die eine bestimmte Kennzeichnung erhalten sollten, diese tatsächlich auch erhalten hat. Da es einen Kompromiss zwischen diesen beiden Kennzahlen gibt, wird auch das F1-Maß berechnet, das beide Kennzahlen kombiniert. Wie zu sehen ist, erreicht Tesseract eine relativ hohe Precision und einen niedrigen Recall. Dies bedeutet, dass die erkannten Bereiche in vielen Fällen korrekt sind, gleichzeitig aber viele Bereiche mit relevanten Informationen überhaupt nicht gefunden werden. Die niedrigen Werte der IoU spiegeln die geringe Überlappung zwischen den „Ground-Truth“-Bounding Box und den von Tesseract vorhergesagten Bounding Boxen wider.

Für Aufgabe 2 (Zeilenerkennung) besteht der Testdatensatz aus 1411 Zeilenbildern (einer einzigen Zeichenfolge in einer Zeile). Die Leistung von Tesseract wurde basierend auf einem Satz von Zeilenbildern bewertet, die nur GD&T-Informationen enthalten, und einem weiteren Satz

von allen Zeilenbildern, die irgendeine Textinformation enthalten. Da es darum geht, die in den Bildern dargestellten Zeichen zu extrahieren, werden der Zeichenfehler rate (Character Error Rate - CER) und der Wortfehler rate (Word Error Rate - WER) als Bewertungsmaßstäbe verwendet. Während Tesseract bei einem Testdatensatz, der alle Arten von Zeilenbildern enthält, recht gut abschneidet, sinkt seine Leistung erheblich, wenn es speziell auf GD&T-Informationen angewendet wird.

Der Gesamttest (Aufgabe 3), der über 3700 Zeichen umfasst ist hinsichtlich aller Bewertungskriterien unbefriedigend. Dies war aufgrund der relativ niedrigen Leistungsfähigkeit bei den Einzelaufgaben (Lokalisierung und Zeilenerkennung) zu erwarten. Man erinnere sich an den niedrigen Recall von Aufgabe 1: Wird ein relevanter Bereich überhaupt nicht erkannt, können dessen Informationen nicht extrahiert werden, was zu einer schlechten Leistung bei nachfolgenden Aufgaben führt.

Zusammenfassend lässt sich sagen, dass die Leistung der Standard-OCR-Software Tesseract nicht ausreicht, um technische Zeichnungen zu digitalisieren. Ein Grund dafür könnte sein, dass technische Zeichnungen spezielle Ob-

jekte wie Form- und Lagetoleranz (GD&T), Bemaßungen und Sonderzeichen wie  $\angle$  70  $\circ$   $\perp$   $\sim$   $\phi$   $\frac{F}{L}$   $\frac{M}{P}$   $\frac{S}{T}$   $\frac{U}{V}$  enthalten, die in normalen Dokumenten nicht vorkommen und auch nicht in den Trainingsdaten gängiger OCR-Software verwendet werden.

Technische Zeichnungen unterscheiden sich erheblich von gewöhnlichen Dokumenten, da sie umfangreiche weiße Flächen, geometrische Formen, spezielle Schriftarten und rein binäre Schwarz-Weiß-Pixel aufweisen. Es ist wahrscheinlich, dass das Modell von Tesseract während des Trainings nicht mit diesen Eigenschaften in Berührung kam und daher nicht in der Lage war, sich an sie anzupassen.

## Deep Learning zur Digitalisierung von Zeichnungen

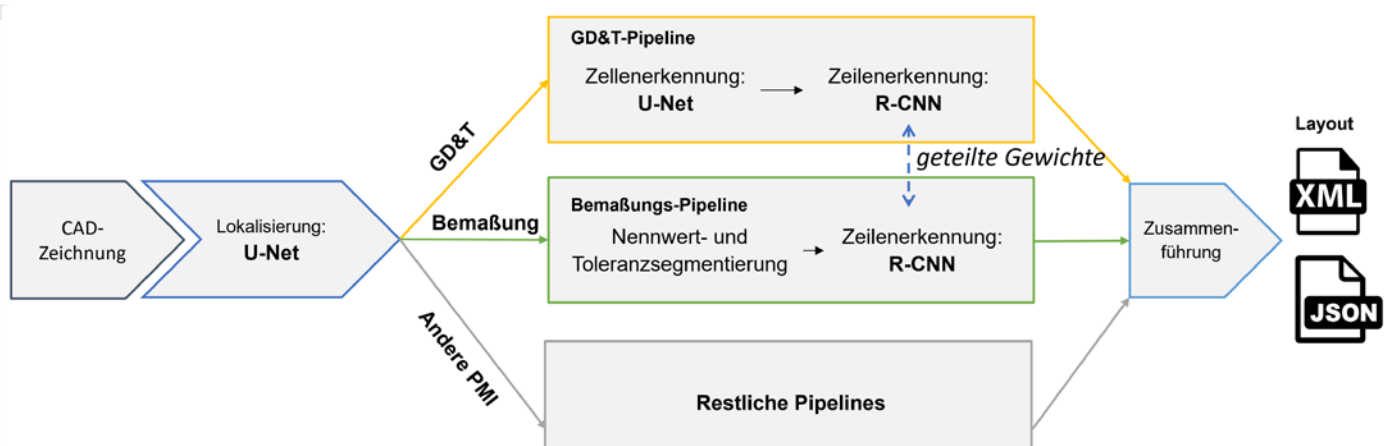
Um die Probleme handelsüblicher OCR-Software zu überwinden, wurde ein neues Framework entwickelt, das sich auf die Digitalisierung von technischen Zeichnungen mit Schwerpunkt auf Form- und Lagetoleranz (GD&T) sowie Bemaßung konzentriert. Das Framework besteht aus einem zentralen Lokalisierungsmodul und mehreren nachgelagerten Pipelines. Das Lokalisierungsmodul segmentiert relevante Objekte aus der gesamten technischen Zeichnung. Je nach Art des segmentierten Objekts wird die entsprechende nachgelagerte Pipeline aufgerufen. Jede Pipeline verwendet das vom Lokalisierungsmodul erzeugte Bild als Eingabe und erkennt den textbasierten Inhalt des jeweiligen segmentierten Bildes. Im letzten Schritt werden die Ergebnisse der verschiedenen Pipelines zusammengeführt, was das endgültige Erkennungsergebnis der gesamten technischen Zeichnung bildet. Das vorgeschlagene Framework basiert auf [1]. Im Vergleich zu diesem Ansatz wird jedoch ein zentralisierter, auf Deep

Learning basierender Lokalisierungsschritt verwendet, und die GD&T-Pipeline wurde um einen Zellenerkennungsschritt und einen Zeilenerkennungsschritt erweitert. Bild 3 veranschaulicht das Framework.

Das Lokalisierungsmodul wird mit einem U-Net [5] implementiert, einem semantischen Segmentierungnetzwerk, das Bilder auf Pixelebene klassifiziert. Ein Nachbearbeitungsschritt dient zur Ermittlung der Bounding Boxen und ihrer Ausrichtung. Die ermittelten Bounding Boxen dürfen Winkel haben, um sicherzustellen, dass auch nicht-horizontal ausgerichtete PMI lokalisiert werden können. Das Netzwerk wird mit einem internen Datensatz trainiert, der echte CAD-Zeichnungen enthält. Die Positionen der relevanten Objekte wurden von menschlichen Annotatoren gekennzeichnet. Im Gegensatz zu anderen gängigen Netzwerken zur Erkennung von Objekten wie YOLO oder Faster R-CNN, die in anderen Arbeiten [6-9] eingesetzt werden, bietet ein U-Net den Vorteil, dass Bilder mit beliebig großen Bemaßungen direkt verarbeitet werden können. Es sei angemerkt, dass der Vergleich der Leistung der implementierten Netzwerke mit anderen in der Literatur veröffentlichten Netzwerken für typische Aufgaben der Computer Vision, wie der Objekterkennung, nicht Teil dieses Beitrags ist.

Analog zu [10] besteht die GD&T-Pipeline aus einem Zellenerkennungsschritt und einem Zeilenerkennungsschritt. Die Aufteilung der GD&T-Pipeline in zwei Schritte wird durch die Beobachtung motiviert, dass Tabellenränder innerhalb von GD&T-Objekten die Erkennungsleistung verringern. Für den Zellenerkennungsschritt wird ein U-Net [5] verwendet, das auf generierten Daten trainiert wurde. Ebenso beginnt die Bemaßungs-Pipeline mit einem Segmentierungsschritt, der die Positionen von Nennwerten und Toleranzen lokalisiert. Für diesen Schritt wird in dieser Pipeline Tesseract OCR eingesetzt.

**Bild 3:** Vorgeschlagenes Framework zur Extraktion von Produkt- und Fertigungsinformationen aus Zeichnungen.



In beiden Fällen führt der Vorverarbeitungsschritt zu Zeilenbildern (Segmente, die eine einzelne Zeichenfolge enthalten), die dann an das gemeinsame Zeilenerkennungsmodell weitergeleitet werden, um den Inhalt zu erkennen. Für diesen Schritt wird keine handelsübliche OCR-Software verwendet. Stattdessen wird ein R-CNN [11] von Grund auf mit einem synthetischen Datensatz von etwa 8 Millionen Bildzeilen trainiert. Grund dafür ist die große Anzahl ungewöhnlicher Zeichen, die in technischen Zeichnungen vorkommen, und die geringe Leistung handelsüblicher OCR-Software hierfür. Ein solches Framework bietet den Vorteil, dass die Lokalisierungsstufe und die einzelnen Pipelines modular gestaltet und implementiert werden können. Einzelne Pipelines können aktualisiert und ersetzt werden, falls in Zukunft fortschrittlichere Technologien verfügbar sind. Zusätzliche Pipelines können ebenfalls in das bestehende Framework integriert werden, wenn weitere Objekttypen erkannt werden sollen.

## Evaluation des Deep-Learning-Frameworks

Um einen Vergleich zu ermöglichen, wird das vorgeschlagene Deep-Learning-Framework für dieselben Aufgaben wie Tesseract evaluiert. Die Ergebnisse der Bewertung sind ebenfalls in Bild 2 dargestellt.

Wie zu sehen ist, verschlechtert sich der Precision-Wert für Aufgabe 1 (Lokalisierung) beim Framework leicht um 3,0 % (GD&T) und 2,1 % (Bemaßungen). Gleichzeitig zeigt sich jedoch eine signifikante Verbesserung des Recall-Wertes (von 29,4 % auf 86,2 % für GD&T und von 25,8 % auf 56,5 % für Bemaßungen) und der IoU (von 27,9 % auf 72,6 % für GD&T und von 23,4 % auf 45,3 % für Bemaßungen). Ein höherer Recall ist für die Aufgabe wichtig, da in den nachfolgenden Schritten auch nichts aus Segmenten extrahiert werden kann, die überhaupt nicht gefunden wurden. Diese Informationen würden in der digitalisierten Version der technischen Zeichnung am Ende fehlen. Wie erwartet, spiegelt sich dies auch in höheren Werten für den IoU-Score und das F1-Maß wider. Die Precision- und Recall-Werte für GD&T sind bereits zufriedenstellend, jedoch streben wir eine weitere Verbesserung der Lokalisierungsleistung für Bemaßungen an.

Tesseract hatte im Test große Schwierigkeiten mit der zweiten Aufgabe, nämlich der Erkennung von Zeichen in jedem Zeilensegment, insbesondere bei den Segmenten in GD&T. Hier verringert sich durch Einsatz des selbst entwickelten R-CNN, das für die Zeilenerkennungsaufgabe auf speziell generierten Zeichnungsdaten trainiert wurde, die Character Error Rate (CER) von 51,2 % auf 13,5 % und die Word Error Rate (WER) von 67,4 % auf 24,3 %. Gleichzeitig muss eine leichte Verschlechterung der

Worterkennungsrates für den gesamten Datensatz in Kauf genommen werden.

Die erreichten Verbesserungen spiegeln sich auch in den Evaluationsergebnissen von Aufgabe 3 wider, in der das Gesamtsystem getestet wird. Für beide Arten von produktionsbezogenen Informationen wurden signifikante Verbesserungen erzielt. Dennoch muss festgestellt werden, dass die Leistung für die Kategorie Bemaßungen weiterhin unter den Erwartungen liegt. Eine mögliche Erklärung ist, dass das neue System noch nicht gut genug darin ist, Elemente dieses Typs zu segmentieren. Dies könnte auf eine unzureichende Menge an Trainingsdaten zurückzuführen sein, was es dem tiefen neuronalen Netzwerk erschwert, die Besonderheiten dieser variantenreichen Art von produktionsbezogenen Informationen zu erkennen.

## Der Schlüssel zum Erfolg

Die Digitalisierung von historischen technischen Zeichnungen ist und bleibt eine wichtige Aufgabe in der Automatisierung von Produktionsprozessen. Konventionelle Standardsoftware erzielt oft keine guten Ergebnisse, vermutlich weil die zugrunde liegenden Modelle auf anderen Daten trainiert wurden und daher nicht gut an die besonderen Merkmale von Zeichnungsdaten angepasst sind. Daher ist es sinnvoll, spezialisierte Software für diese Aufgabe zu entwickeln und zu nutzen. Ein mehrstufiges Framework, in dem verschiedene Arten von produktionsbezogenen Informationen separat behandelt werden, bietet die Möglichkeit, die spezifischen Eigenschaften jeder PMI anzugehen und damit die Qualität erheblich zu verbessern. Darüber hinaus ermöglicht die vorgeschlagene modulare Struktur, einzelne Komponenten jederzeit auszutauschen, wenn bessere technische Lösungen verfügbar werden.

*Das dem vorliegenden Artikel zugrunde liegende Projekt wurde vom Bundesministerium für Bildung und Forschung (BMBF) sowie dem Land Baden-Württemberg unter der Fördernummer 03FHP127B gefördert. Die Autoren sind verantwortlich für den Inhalt dieser Veröffentlichung.*

## Literatur

- [1] Toro, J.; Wiberg, A.; Tarkian, M.: Optical Character Recognition on Engineering Drawings to Achieve Automation in Production Quality Control. In: *Frontiers in Manufacturing Technology 3* (2023).
- [2] Scheibel, B.; Mangler, J.; Rinderle-Ma, S.: Extraction of Dimension Requirements from Engineering Drawings for Supporting Quality Control in Production Processes. In: *Computers in Industry 129*, S. 103442 (2021).
- [3] Henderson, T.: *Analysis of Engineering Drawings and Raster Map Images*. Springer 2014.

- [4] Kay, A.: Tesseract: An Open-Source Optical Character Recognition Engine. In: Linux Journal 159 (2007).
- [5] Ronneberger, O.; Fischer, P.; Brox, T.: U-net: Convolutional Networks for Biomedical Image Segmentation. In: Medical Image Computing and Computer-Assisted Intervention 9351, S. 234–241 (2015).
- [6] Schlagenhauf, T.; Netzer, M.; Hillinger, J.: Text Detection on Technical Drawings for the Digitization of Brown-Field Processes. In: 16th CIRP Conference on Intelligent Computation in Manufacturing Engineering 118, S. 372–377 (2023).
- [7] Haar, C.; Kim, H.; Koberg, L.: Ai-Based Engineering and Production Drawing Information Extraction. In: Flexible Automation and Intelligent Manufacturing: The Human-Data-Technology Nexus. Springer, S. 374–382 (2023).
- [8] Nguyen, T. u. a.: Object Detection and Text Recognition in Large-Scale Technical Drawings. In: Proceedings of the 10th International Conference on Pattern Recognition Applications and Methods. SciTePress, S. 612–619 (2021).
- [9] Elyan, E.; Jamieson, L.; Ali-Gombe, A.: Deep Learning for Symbols Detection and Classification in Engineering Drawings. In: Neural Networks 129, S. 91–102 (2020).
- [10] Hashmi, K. A. u. a.: Current Status and Performance Analysis of Table Recognition in Document Images with Deep Neural Networks. In: IEEE Access 9, S. 87663–87685 (2021).
- [11] Shi, B.; Bai, X.; Yao, C.: An End-To-End Trainable Neural Network for Image-Based Sequence Recognition and Its Application to Scene Text Recognition. In: IEEE Transactions on Pattern Analysis and Machine Intelligence 39, S. 2298–2304 (2017).