

SAMMLUNGSÜBERGREIFEND

Pixelmatch

*KI-unterstützte Digitalisierung der Fotosammlungen
Machura/Meisinger und Mejchar*

Von Karin Kühtreiber, Florian Kleber und Marco Peer

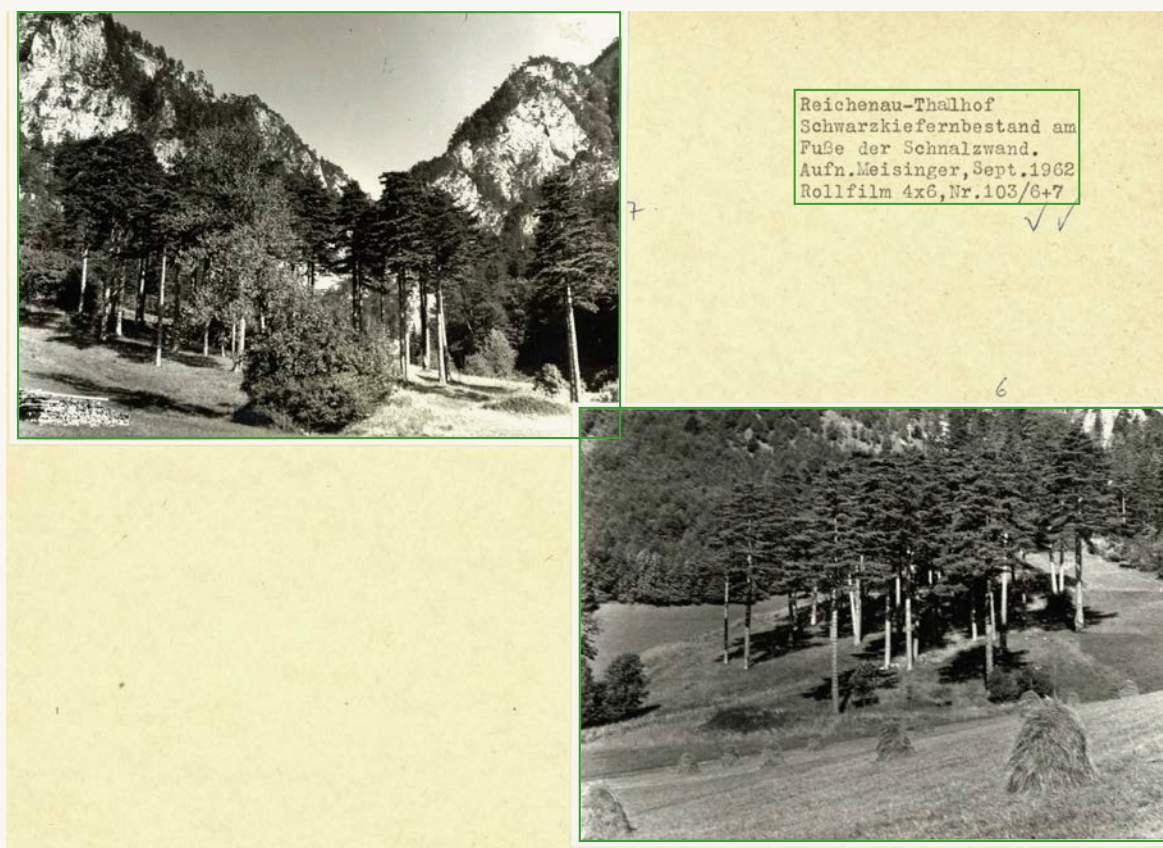


Foto: Landessammlungen NÖ

Zu den Kernzielen der strategischen Weiterentwicklung der Landessammlungen Niederösterreich (LSNÖ) in der für den Zeitraum 2021 bis 2030 definierten Zielsetzungsperiode gehört es, die digitale Transformation der Sammlungsbestände zu forcieren.¹ Zusätzlich zu den laufenden Digitalisierungsmaßnahmen innerhalb der einzelnen Sammlungsbereiche² wurde dieses Vorhaben durch das 2023 erfolgreich eingeworbene Projekt „Digitalisierung ausgewählter musealer Bestände der Landessammlungen Niederösterreich“, finanziert durch das vom Bundesministerium für Kunst, Kultur, öffentlichen Dienst und Sport (seit April 2025: Bundesministerium für Wohnen, Kunst, Kultur, Medien und Sport) ausgeschriebene Förderprogramm „Kulturerbe digital“³, maßgeblich vorangetrieben.⁴ Ziel war es unter anderem, große Teile der naturkundlichen Fotosammlung Lothar Machura und Augustin Meisinger sowie des Nachlasses der Fotokünstlerin Elfriede Mejchar unter Heranziehung von Methoden künstlicher Intelligenz (KI) einerseits möglichst effizient zu erschließen und andererseits

automatisiert mit Metadaten anzureichern. Als Kooperationspartner für diese Aufgaben konnte das Computer Vision Lab⁵ am Institute of Visual Computing and Human-Centered Technology der Technischen Universität Wien gewonnen werden.

Die umfangreiche naturkundliche Fotosammlung⁶ entstand in den 1940er- bis 1960er-Jahren und besteht aus Schwarz-Weiß-Fotos sowie einer nach Sachgebieten geordneten Kartei im Umfang von rund 2.400 Einzelblättern, auf denen neben einem oder mehreren Fotos maschinenschriftliche Informationen zum Aufnahmeort, zu dem Motiv, dem Aufnahmedatum und dem Fotografen festgehalten wurden. Für den Aufbau und die Pflege der Sammlung zeichneten vor allem Lothar Machura (1909–1982) und Augustin Meisinger (1899–1971) verantwortlich, die im Rahmen ihrer Tätigkeit als niederösterreichische Landesbeamte Naturdenkmale, ökologische Schutzgebiete sowie Flora und Fauna fotografisch dokumentierten, aber auch deren Zerstörung und Veränderungen mit der Kamera festhielten. >>

Im Jahr 2013 übernahmen die LSNÖ im Rahmen einer Schenkung den Vorlass der Fotokünstlerin Elfriede Mejchar⁷ (1924–2020), von dem nun durch das geförderte Projekt mehr als 9.000 Werke digitalisiert werden konnten, in den Sammlungsbereich Kunst nach 1960. Die Künstlerin war von 1947 bis 1984 Mitarbeiterin des Bundesdenkmalamts, wo sie als Fotografin mit der Dokumentation von Baudenkmalern und Kunstwerken betraut war. Daneben übernahm sie externe Auftragsarbeiten. Davon ist insbesondere die fotografische Dokumentation von Industriedenkmalern in Österreich für das Institut für Kunstgeschichte, Denkmalpflege und Industriearchäologie der Technischen Universität Wien hervorzuheben, die in zwei Publikationen⁸ Eingang fand.

Für beide Fotosammlungen galt es in einem ersten Schritt ein automatisiertes Image Matching/Retrieval⁹ durchzuführen. Im Fall der naturkundlichen Sammlung mussten die Fotos auf den Karteikarten im Gesamtbestand der Negative und Positive selektiert werden; im Projektteil Mejchar sollten die Fotos in den Publikationen mit dem Sammlungskonvolut abgeglichen werden. Mittels KI-unterstützter Bilddetektierung (Object Detection) unter Heranziehung des Deep Learning Models YOLOv8¹⁰ erfolgte zunächst eine Separierung der Inhalte der Karteikarten bzw. der Bücher in die Bereiche „Bild“, „gedruckter Text“ und „Handschrift“. Für das anschließende Image Matching/Retrieval wurde für jedes Bild mithilfe eines Self-supervised Vision Transformer Models¹¹ (unter Verwendung von DINOv2¹²) ein Encoding¹³ berechnet und wurden die Bilder im Suchdatensatz der Gesamtsammlung nach Ähnlichkeit gereiht. Um die Ergebnisse nachvollzieh- und kontrollierbar zu machen, wurden Visualisierungen in Form von Bildreihen aus sechs Fotos erstellt: Beginnend mit dem gesuchten Bild (grüne Ziffer), folgten absteigend nach Ähnlichkeit die gefundenen Fotos (rote Ziffer), sodass das Bild mit der größten Übereinstimmung jeweils an erster Stelle neben dem gesuchten Foto zu finden war. Eine quantitative Evaluierung auf Grundlage von 100 naturkundlichen Bildern ergab eine korrekte Matching-Rate von zirka 85 Prozent, wobei im Falle der falschen Treffer der be-

gründete Verdacht bestand, dass es sich um Karteikartenbilder handelte, zu denen im Bestand der Gesamtsammlung keine entsprechenden Positive oder Negative vorhanden sind, womit die tatsächliche Trefferquote als höher anzunehmen ist.

Der zweite Schritt der automatisierten Datenerfassung betraf die Aufbereitung der Textdaten auf den naturkundlichen Karteikarten, die als Metadaten für die jeweiligen Digitalisate benötigt wurden. Hierfür wurde mithilfe von Transkribus¹⁴ eine Texterkennung (Optical Character Recognition [OCR], Handwritten Text Recognition [HTR]), anschließend mittels ChatGPT 3.5¹⁵ eine Named Entity Recognition¹⁶ vorgenommen und der Fließtext in eine strukturierte Form (JSON-Format) überführt. Die Ergebnisse präsentierten sich als gute Grundlage für die gewünschten Metadaten, erforderten allerdings eine eingehende redaktionelle Nachbearbeitung für jeden Datensatz. Hier erwiesen sich die uneinheitlich notierten Angaben auf den Karteikarten, wie beispielsweise unterschiedliche Schreibvarianten der Fotografennamen oder des Datums, als nachteilig für fehlerfreie Ergebnisse. Insgesamt konnten durch das Image Matching/Retrieval und die automatisierte Erfassung der Metadaten mehr als 2.600 Digitalisate der naturkundlichen Fotosammlung erschlossen und veröffentlicht werden.

Schließlich galt es für beide Sammlungskonvolute anhand von Abfragen gängiger Geodatenbanken automatisiert Georeferenzen zu erheben, um die Sammlungsbestände in Zukunft auch über Karten durchsuchbar zu machen. Hier lieferten die Geocoder APIs von Google Maps und Google Places¹⁷ die besten Ergebnisse, während etwa GeoNames¹⁸ für den hier vorliegenden Datenbestand eine hohe Fehlerquote aufwies. Eine manuelle Überprüfung und allfällige Korrektur der abgefragten Georeferenzen war allerdings auch hier unerlässlich, da die Automatisierung nicht nur bei gleichlautenden Ortsnamen an ihre Grenzen stieß, sondern auch bei historischen Flurnamen oder abgekommenen Gebäuden (z. B. Wienerberger Ziegelöfen) keine oder falsche Treffer lieferte.



Beispiele für Bildreihen der Fotomatchings aus den Sammlungen Machura (Reihe 1 zu Inv.Nr. NK-F-2006; Reihe 2 zu Inv.Nr. NK-F-2729) und Mejchar (Reihe 3 zu Inv.Nr. KS-19106/4//497; Reihe 4 zu Inv.Nr. KS-19106/4//584)

Foto: Landessammlungen NÖ

Als Fazit kann festgehalten werden, dass das KI-basierte Image Matching/Retrieval für beide Fotosammlungen sehr gute Resultate erbrachte und die Digitalisierungsarbeiten maßgeblich beschleunigte. In gleicher Weise lieferten die KI-gestützte Erfassung der Textdaten und die automatisierte Abfrage der Georeferenzen einen maßgeblichen Beitrag zur effizienten Erschließung der gewünschten Metadaten. Wie bei allen Einsatzbereichen der künstlichen Intelligenz gilt aber auch hier: Die neuen Technologien können die Expertisen und Erfahrungen sachkundiger Fachkräfte nicht ersetzen – sie werden auch in Zukunft unverzichtbar sein.

¹ Vgl. Amt der Niederösterreichischen Landesregierung, Abteilung Kunst und Kultur (Hrsg.): Strategie der Landessammlungen Niederösterreich 2021 bis 2030. St. Pölten 2021, S. 15, 18, https://landessammlungen-noe.at/files/news/bilder/Landessam_STRe_151121.pdf, abgerufen am 2.4.2025.

² Vgl. <https://landessammlungen-noe.at/de/online.html>, abgerufen am 13.12.2024.

³ Vgl. Förderprogramm „Kulturerbe digital“, www.bmwkms.gv.at/themen/kunst-und-kultur/schwerpunkte/digitalisierung/foerderprogramm-kulturerbe-digital.html, 1. Ausschreibung, abgerufen am 3.12.2024.

⁴ Vgl. dazu auch Armin Laussegger, Sandra Sam: Ein Jubiläum und vieles mehr. Die Landessammlungen Niederösterreich und das Zentrum für Museale Sammlungswissenschaften. In: dies. (Hrsg.), Im Bestand. Sammlungswissenschaftliche Einblicke. Tätigkeitsbericht 2023 der Landessammlungen Niederösterreich und des Zentrums für Museale Sammlungswissenschaften. Veröffentlichungen aus den Landessammlungen Niederösterreich, Nr. 10. St. Pölten 2024, S. 14–15. <https://cvi.tuwien.ac.at/>, abgerufen am 9.12.2024.

⁶ Ausführlich zur Sammlung: Erich Steiner: Blick ins Land. Fotografien von Lothar Machura und Augustin Meisinger. Katalog des Niederösterreichischen Landesmuseums, N.F. 512. St. Pölten 2013; sowie Ronald Lintner, Christian Dietrich: Blick in die Natur. Bewahren für die Zukunft. In: Jahrbuch für Landeskunde von Niederösterreich, N.F. 89, 2023. St. Pölten 2024, S. 267–273.

⁷ Z. B. Anton Holzer u. a. (Hrsg.): Elfriede Mejchar. Grenzgängerin der Fotografie. München 2024.

⁸ Vgl. Manfred Wehdorn, Ute Georgeacopol-Winischhofer: Baudenkmäler der Technik und Industrie in Österreich, Bd. 1: Wien, Niederösterreich, Burgenland. Wien – Köln – Graz 1984; Manfred Wehdorn, Ute Georgeacopol-Winischhofer, Paul W. Roth: Baudenkmäler der Technik und Industrie in Österreich, Bd. 2: Steiermark und Kärnten. Wien – Köln – Weimar 1991.

⁹ Image Matching sucht in einer Bilddatenbank nach dem ähnlichsten Treffer des eingegebenen Suchbildes anhand von Übereinstimmungen (z. B. Formen oder Farben). Bei Image Retrieval ist das Suchergebnis eine gereichte Liste der ähnlichsten Bilder.

¹⁰ <https://yolov8.com/>, abgerufen am 13.12.2024.

¹¹ Self-Supervised Learning ist ein Ansatz des maschinellen Lernens, bei dem ein Modell aus automatisch generierten Pseudo-Labels oder Strukturen innerhalb unbeschrifteter Daten lernt, um später Aufgaben mit weniger beschrifteten Daten effektiv lösen zu können.

¹² <https://dino2.metademolab.com/>, abgerufen am 10.12.2024.

¹³ Darunter wird eine numerische Repräsentation eines Bildes verstanden, die seine wesentlichen Merkmale beschreibt.

¹⁴ www.transkribus.org/, abgerufen am 13.12.2024.

¹⁵ <https://chatgpt.com/>, abgerufen am 17.12.2024.

¹⁶ Named Entity Recognition bezeichnet eine Technik im Bereich der natürlichen Sprachverarbeitung, die dazu dient, bestimmte Entitäten in unstrukturiertem Text zu identifizieren und zu klassifizieren.

¹⁷ <https://developers.google.com/maps/documentation?hl=de>, abgerufen am 17.12.2024.

¹⁸ www.geonames.org/export/web-services.html, abgerufen am 17.12.2024.