

# Buddhist Murals of Kucha on the Northern Silk Road. Ein Versuch der Semi-Automatisierung der Annotierung

**Radisch, Erik**

radisch@saw-leipzig.de

Sächsische Akademie der Wissenschaften zu Leipzig, Deutschland

Die buddhistischen Höhlenkomplexe in der Region Kucha an der nördlichen Seidenstraße (Uigurisches Autonomes Gebiet Xinjiang, VR China) beherbergen beeindruckende Wandmalereien, die etwa aus dem 5. bis 10. Jhd. stammen. Die ersten Hinweise auf eine frühere buddhistische Kultur wurden zu Beginn des 20. Jahrhunderts entdeckt, woraufhin mehrere Länder Expeditionen in das Gebiet schickten, um die einst in der Region vorherrschende Religion zu erforschen. Es war eine Sensation, als verschiedene buddhistische Höhlenkomplexe entdeckt wurden. Damals wurden auch die ersten Fotografien vom Ist-Zustand der Höhlen angefertigt und Teile der Malereien aus den Höhlen entnommen und in die jeweiligen Nationalmuseen gebracht. Heute sind Fragmente der Wandmalereien über die ganze Welt verstreut, was eine Zuordnung zu den einzelnen Ursprungshöhlen sehr schwierig macht (Weitere Informationen: Yaldiz 1987; Popova 2008; Dreyer 2015).

Das hier vorgestellte Projekt hat es sich zur Aufgabe gemacht, die Wandmalereien in situ und die weltweit vorhandenen Einzelstücke zu dokumentieren, zu beschreiben und mit Hilfe von historischen Fotografien wieder in ihren ursprünglichen Kontext zu bringen.<sup>1</sup>

Das Projekt bedient sich moderner Möglichkeiten der Digital Humanities, indem nicht nur eine umfangreiche textliche Beschreibung einzelner Szenen erfolgt, sondern auch die Bildinhalte der sich wiederholenden Darstellungen erfasst und mit digitalen Methoden angereichert werden. Zu diesem Zweck wird das digitale Bildannotationsstool Annotorious<sup>2</sup> (siehe Abbildung 1) genutzt, um die Inhalte direkt mit einer rund 1.000 Einträge umfassenden Taxonomie zu annotieren. Die erarbeiteten Forschungsdaten stehen online frei zur Verfügung.<sup>3</sup>

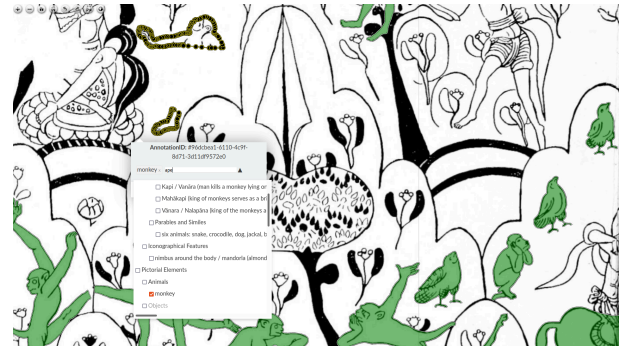


Abbildung 1: Annotieren mit Annotorious

Die Annotation von Objekten im Bild ermöglicht zwar eine wissenschaftliche Nachvollziehbarkeit der identifizierten Objekte, ist aber auch eine sehr umfangreiche und zeitaufwändige Aufgabe. Viele Bildinhalte wiederholen sich häufig in unterschiedlichen Zusammenhängen. Außerdem liegen manchmal mehrere Bilder eines Objekts aus verschiedenen Perspektiven vor oder es gibt Bilder aus der Zeit der Expeditionen, auf denen abgetrennte Teile noch in ihrem ursprünglichen Kontext zu sehen sind. Es besteht also die Notwendigkeit, manchmal sehr ähnliche oder gleiche Objekte mehrfach zu annotieren. Die Übertragung von Annotationen ist jedoch schwierig. Selbst wenn Fotos von denselben Objekten vorhanden sind, können unterschiedliche Blickwinkel und verschiedene Objektive dazu führen, dass die Bilder verzerrt sind. Es ist kaum möglich, diese Aufgabe mit herkömmlichen Computer-Vision-Methoden automatisch durchzuführen.

Aus diesem Grund wird im Rahmen des Projekts derzeit versucht, mit den bereits vorgenommenen Annotationen Region Based Convolutional Neural Networks (RCNNs)<sup>4</sup> zu trainieren, um in Zukunft zumindest Teile der Annotation halbautomatisch (die Annotierenden werden die Möglichkeit haben, die gefundenen Regionen des RCNNs einzusehen und diese anzunehmen oder gegebenenfalls zu verbessern) durchführen zu können.

Bislang wurden RCNNs in den Digital Humanities vor allem zur Identifizierung, Lokalisierung und Ordnung von Objekten in Bildern eingesetzt (siehe z.B.: Howanitz et al. 2019; Arnold/Tilton 2019; Duhaime 2019; Duhaime 2019; Helm et al. 2021). Ihre Verwendung für eine halbautomatische Annotation ist zumindest zur Kenntnis des Autors dieses Posterproposals noch nicht umgesetzt worden. Da die Ränder automatisch erkannten Annotationen oft ausfransen oder nicht das gesamte Objekt erkannt wird, mag es auch gewagt sein, einen solchen Versuch zu starten.

Die Voraussetzungen des Kucha-Projekts sind sehr gut. Es existieren bereits über 9.000 Polygone, die in insgesamt etwa 12.000 Annotationen verwendet wurden (ein Polygon kann mit mehreren Elementen der Taxonomie verknüpft sein). Einige Objekte wurden mehrer hundert Mal annotiert. Es gibt jedoch auch einige Probleme, die zu berücksichtigen sind. So gibt es beispielsweise zwei grundlegend verschiedene Arten von Bildern im Korpus: zum einen Fotografien (historische und moderne), zum anderen Zeichnungen der Malereien. Die Erkennung auf Fotografien dürfte deutlich schwieriger sein, da hier die

Malereien oft in sehr schlechtem Zustand sind und selbst von einem geübten Auge nur schwer zu identifizieren sind.

Es wurden verschiedene Experimente durchgeführt, die die prinzipielle Nutzbarkeit von RCNNs für eine semi-automatisierte Annotation testen sollten. Diese werden im Poster näher präsentiert. In einem ersten Experiment wurden alle Trainingsdaten zusammen trainiert ( $mAP_{IoU=0.75} = 5.85^5$ ). Ein zweites Experiment teilte die Daten nach Bildarten auf ( $mAP_{IoU=0.75} = 3.40$  für Fotos und  $mAP_{IoU=0.75} = 4.04$  bei Zeichnungen). Es war dabei auffällig, dass die Zeichnungen bessere Resultate erzielten, die aber niedriger als die Resultate des ersten Experimentes waren. Die folgenden beiden Experimenten wurden deswegen vorerst auf Zeichnungen beschränkt. Das dritte Experiment konzentrierte sich auf Klassen mit mehr als 50 Annotationen, was zwar den Großteil der Klassen außen vor ließ, aber vielversprechende Ergebnisse erzeugte ( $mAP_{IoU=0.75} = 23.06$ ). Deswegen wurden in einem vierten Experiment alle Klassen von menschlichen Abbildungen zu einer Metaklasse zusammengelegt. Dieses letzte Experiment funktionierte besonders vielversprechend ( $mAP_{IoU=0.75} = 59.99$ ; ein Beispiel kann in Abbildung 2 eingesehen werden). Zur Bewertung der Experimente kann der Output der Testbilder online eingesehen werden.<sup>6</sup> Bis zur DHd im nächsten Jahr sollen einige weitere Experimente durchgeführt werden und ein erster Prototyp zur Anwendung kommen und im Rahmen des Posters präsentiert werden.

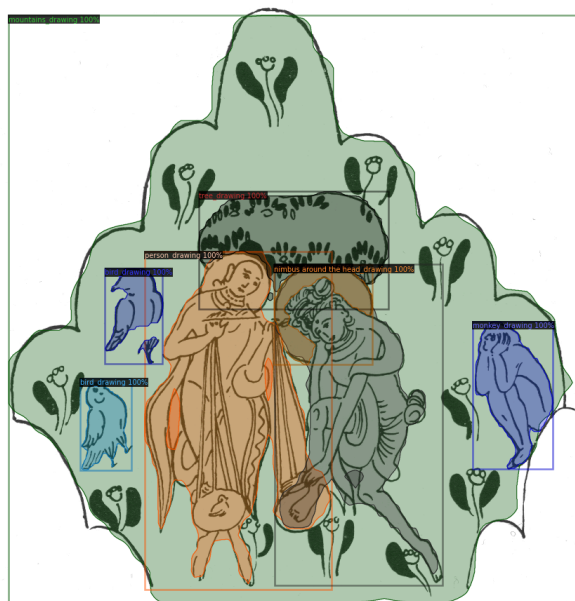


Abbildung 2: Beispiel-Output für Experiment 4

## Fußnoten

1. Wissenschaftliche Bearbeitung der buddhistischen Höhlenmalereien in der Kucha-Region der nördlichen Seidenstrasse (SAW-Leipzig). Datenbank: <https://kucha-saw-leipzig.de>.

2. <https://recogito.github.io/annotorious/>. Das Projekt hat seine eigene Reihe «Leipzig Kucha Studies». Erster Band: Konczak-Nagel/Zin 2020

3. <https://kuchatest.saw-leipzig.de/>

4. Das Projekt nutzt hierfür detectron2 (Wu et al. 2019).

5. Die Maßeinheit gibt die Mean Average Precision bei einer Mindestübereinstimmung einer Region mit dem Goldstandard von 75% an.

6. <https://github.com/erikradisch/examplePics>

## Bibliographie

**Arnold, T. and Tilton, L.** 2019. "Distant viewing: analyzing large visual corpora." In *Digital Scholarship in the Humanities* 36(1), DOI: <https://doi.org/10.1093/llc/fqz013>.

**Dreyer, C.** 2015. "Abenteuer Seidenstraße: Die Berliner Turfan-Expeditionen 1902–1914." Leipzig: Seemann.

**Duhaime, D.** 2019. "PixPlot." <https://github.com/YaleDHLab/pix-plot>.

**Helm, W., Schmideler, S., Im, C., Mandl, T., Kollmann, S. and Müller, L.** 2021. "Wie sich die Bilder ähneln. Vom Zufallsfund zur systematischen Forschung im Bereich der automatisierten Bildähnlichkeitssuche." In *Fabrikation von Erkenntnis: Experimente in den Digital Humanities*, ed. Burghardt, M., Dieckmann, L., Steyer, T., Trilcke, P., Walkowski, N.-O., Weis, J. and Wuttke, U. ZfdG (Sonderband 5). DOI: 10.26298/melusina.8f8w-y749-wsdb.

**Howanitz, G., Bermeitinger, B., Radisch, E., Gassner, S., Rehbein, M. and Handschuh, S.** 2019. July 11. "Deep Watching - Towards New Methods of Analyzing Visual Media in Cultural Studies." *Digital Humanities 2019* (DH2019), Utrecht, Netherlands. <https://zenodo.org/record/3326470#.Y5uLSKfMIYw>.

**Konczak-Nagel, I. and Zin, M.** 2020. "Essays and Studies in the Art of Kucha" (Leipzig Kucha Studies 1). New Delhi: Dev Publishers.

**Popova, I. F. (ed.)** 2008. "Russian Expeditions to Central Asia at the Turn of the 20th Century: Collected Articles." St Petersburg: Slavia Publishers.

**Wu, Y., Kirillov, A., Massa, F., Lo, W. and Girshick, R.** 2019. "Detectron2." <https://github.com/facebookresearch/detectron2>.

**Yaldiz, M.** 1987. "Archäologie und Kunstgeschichte Chinesisch-Zentralasiens (Xinjiang)." *Handbuch der Orientalistik, Abteilung 7, Kunst und Archäologie, Band 3, Innerasien*. Leiden: Brill.