BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 3: e37161
doi: 10.3897/biss.3.37161

OPEN ACCESS

Conference Abstract

Use of Semantic Segmentation for Increasing the Throughput of Digitisation Workflows for Natural History Collections

Abraham Nieva de la Hidalga‡, David Owen‡, Irena Spacic‡, Paul Rosin‡, Xianfang Sun‡

‡ Cardiff University School of Computer Science and Informatics, Cardiff, United Kingdom

Corresponding author: Abraham Nieva de la Hidalga (nievadelahidalgaa@cardiff.ac.uk)

Received: 11 Jun 2019 | Published: 18 Jun 2019

Citation: Nieva de la Hidalga A, Owen D, Spacic I, Rosin P, Sun X (2019) Use of Semantic Segmentation for Increasing the Throughput of Digitisation Workflows for Natural History Collections. Biodiversity Information Science and Standards 3: e37161. https://doi.org/10.3897/biss.3.37161

Abstract

The need to increase global accessibility to specimens while preserving the physical specimens by reducing their handling motivates digitisation. Digitisation of natural history collections has evolved from recording of specimens' catalogue data to including digital images and 3D models of specimens. The sheer size of the collections requires developing high throughput digitisation workflows, as well as novel acquisition systems, image standardisation, curation, preservation, and publishing. For instance, herbarium sheet digitisation workflows (and fast digitisation stations) can digitise up to 6,000 specimens per day; operating digitisation stations in parallel can increase that capacity. However, other activities of digitisation workflows still rely on manual processes which throttle the speed with which images can be published. Image quality control and information extraction from images can benefit from greater automation.

This presentation explores the advantages of applying semantic segmentation (Fig. 1) to improve and automate image quality management (IQM) and information extraction from images (IEFI) of physical specimens. Two experiments were designed to determine if IQM and IEFI activities can be improved by using segments instead of full images. The time for segmenting full images needs to be considered for both IQM and IEFI. A semantic segmentation method developed by the Natural History Museum (Durrant and Livermore

2018) adapted for segmenting herbarium sheet images (Dillen et al. 2019) can process 50 images in 12 minutes.

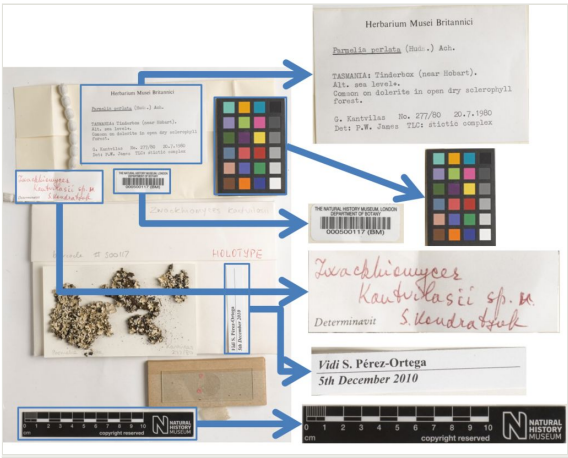

Figure 1.

Example of semantic segmentation of a herbarium sheet. Semantic segmentation entails the identification and classification of image elements. Four classes are targeted for identification use in: labels, barcodes, clour charts and scale. IQM uses labels and barcodes, while IEFI experiments targeted labels and barcodes.

The IQM experiments evaluated the application of three quality attributes to full images and to image segments: colourfulness (Fig. 2), contrast (Fig. 3) and sharpness (Fig. 4). Evaluating colourfulness is an alternative to colour quantization algorithms such as RMSE and Delta E (Hasler and Suesstrunk 2003, Palus 2006), the method produces a value indicating if the image degrades after processing. Contrast measures the difference in luminance or colour that makes an object distinguishable. Contrast is determined by the difference in colour and brightness of the object and other objects within the same field of view (Matkovic et al. 2005, Präkel 2010). Sharpness encompasses the concepts of resolution and acutance (Bahrami and Kot 2014, Präkel 2010). Sharpness influences specimen appearance and readability of information from labels and barcodes. Evaluating the criteria on 56 barcodes and 50 colour charts segments extracted from fifty images took 34 minutes (8 minutes for the barcodes and 26 minutes for colour charts). The evaluation on the corresponding full images took 100 minutes. The processing of individual segments and full images provided results equivalent to subjective manual quality management.

The IEFI experiments compared the performance of four optical character recognition (OCR) programs applied to full images (Drinkwater et al. 2014) against individual segments. The four OCR programs evaluated were Tesseract 4.X, Tesseract 3.X, Abby FineReader Engine 12, and Microsoft OneNote 2013. The test was based on a set of 250 herbarium sheet images and 1,837 segments extracted from them. The results from the experiments show that there is an average OCR speed-up of 49% when using segmented images when compared to processing times for full images (Table 1). Similarly, there was

an average increase of 13% in line correctness (information from lines is ordered and not fragmented (Fig. 5, Table 2). Additionally, the results are useful for comparing the four OCR programs, with Tesseract 3.x offering shortest processing time, while Tesseract 4.X achieving the highest scores for line accuracy (including hand written text recognition). The results suggest that IEFI could be improved by performing OCR using segments rather than whole images, leading to faster processing and more accurate outputs.

Figure 2.

Comparison of results of calculating colurfulness on segmented colour charts from herbarium sheets set (Algorithm from Palus 2006).

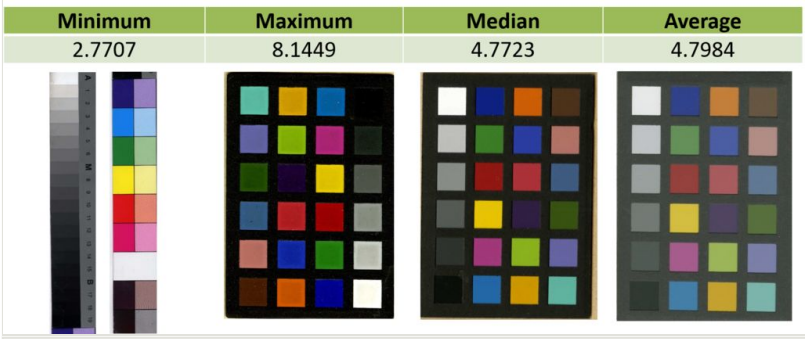

Figure 3.

Comparison of results of calculating contrast on segmented colour charts from herbarium sheets set (Algorithm from Matkovic et al. 2005).

Figure 4.

Comparison of results of calculating sharpness on barcode segments from herbarium sheets set (Algorithm from Bahrami and Kot 2014).

Table 1.

Comparison processing times of three OCR programs applied to full images (250 images) and to individual segments (1,837 segments from those images).

	250 whole images	1,837 segments	Difference (h:m:s)
Tesseract 4.x (Tess 4J)	01:06:05	00:45:02	-00:21:03
Tesseract 3.x (Tess 4J)	00:50:02	00:23:17	-00:26:45
Abbyy FineReader Engine 12	01:18:15	00:29:24	-00:48:51
Processing Time (h:m:s)			

Table 2.

Comparison of line correctness for four OCR programs applied to full images (5 images) and to individual segments (22 segments from those images).

	5 whole images Mean line correctness (%)	22 segments Mean line correctness (%)	Difference (%)
Tesseract 4.x	72.8	75.2	+2.4%
Tesseract 3.x	44.1	63.7	+19.6%
Abbyy FineReader Engine 12	61.0	77.3	+16.3%
Microsoft OneNote 2013	78.9	65.5	-13.4%

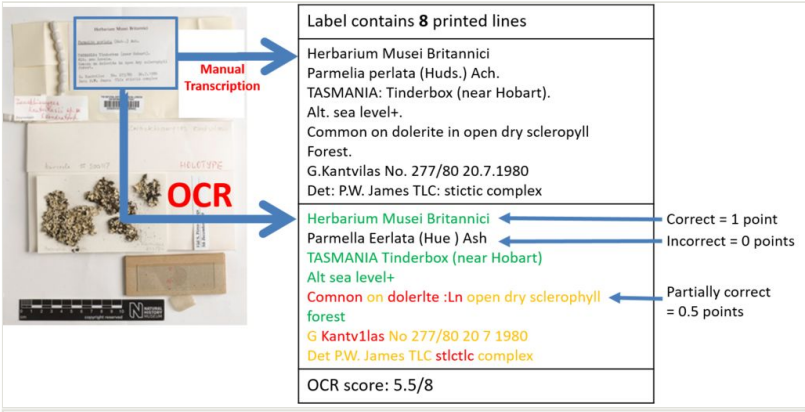

Figure 5.

Line correctness evaluation example. Line correctness means that information from lines is ordered, not fragmented, and not mixed.

The findings support the feasibility of further automation of digitisation workflows for natural history collections. In addition to increasing the accuracy and speed of IQM and IEFI activities, the explored approaches can be packaged and published, enabling automated quality management and information extraction to be offered as a service, taking advantage of cloud platforms and workflow engines.

Keywords

semantic segmentation, image quality management, optical character recognition, digitisation, natural history collections, digital specimens

Presenting author

Abraham Nieva de la Hidalga

Presented at

Biodiversity_Next 2019

Funding program

Horizon 2020 Framework Programme of the European Union

Grant title

ICEDIG – “Innovation and consolidation for large scale digitisation of natural heritage” H2020-INFRADEV-2016-2017 – Grant Agreement No. 777483

References

- Bahrami K, Kot A (2014) A fast approach for no-reference image sharpness assessment based on maximum local variation. IEEE Signal Processing Letters 21 (6): 751-755. https://doi.org/10.1109/LSP.2014.2314487
- Dillen M, Groom Q, Chagnoux S, Güntsch A, Hardisty A, Haston E, Livermore L, Runnel V, Schulman L, Willemse L, Wu Z, Phillips S (2019) A benchmark dataset of herbarium specimen images with label data. Biodiversity Data Journal 7 https://doi.org/10.3897/bdj.7.e31817
- Drinkwater R, Cubey R, Haston E (2014) The use of Optical Character Recognition (OCR) in the digitisation of herbarium specimen labels. PhytoKeys 38: 15-30. https://doi.org/10.3897/phytokeys.38.7168
- Durrant J, Livermore L (2018) Semi-supervised semantic and instance segmentation. https://github.com/NaturalHistoryMuseum/semantic-segmentation. Accessed on: 2018-4-26.
- Hasler D, Suesstrunk S (2003) Measuring colorfulness in natural images. Human Vision and Electronic Imaging VIII https://doi.org/10.1117/12.477378

- Matkovic K, Neumann L, Neumann A, Psik T, Purgathofer W. (2005) Global Contrast Factor-a new approach to Image contrast. Computational Aesthetics 159-168.
- Palus H (2006) Colorfulness of the image: definition, computation, and properties. Conference on Lightmetry and Light and Optics in Biomedicine 709 https://doi.org/10.1117/12.675760
- Präkel D (2010) The visual dictionary of photography. AVA Publishing [ISBN 978-2-940411-04-7] https://doi.org/10.5040/9781350088733