# Visual Analysis of Chapbooks Printed in Scotland


Abhishek Dutta
Visual Geometry Group (VGG)
Department of Engineering Science,
University of Oxford
Oxford, UK
adutta@robots.ox.ac.uk

Giles Bergel
Visual Geometry Group (VGG)
Department of Engineering Science,
University of Oxford
Oxford, UK
giles.bergel@eng.ox.ac.uk

Andrew Zisserman
Visual Geometry Group (VGG)
Department of Engineering Science,
University of Oxford
Oxford, UK
az@robots.ox.ac.uk

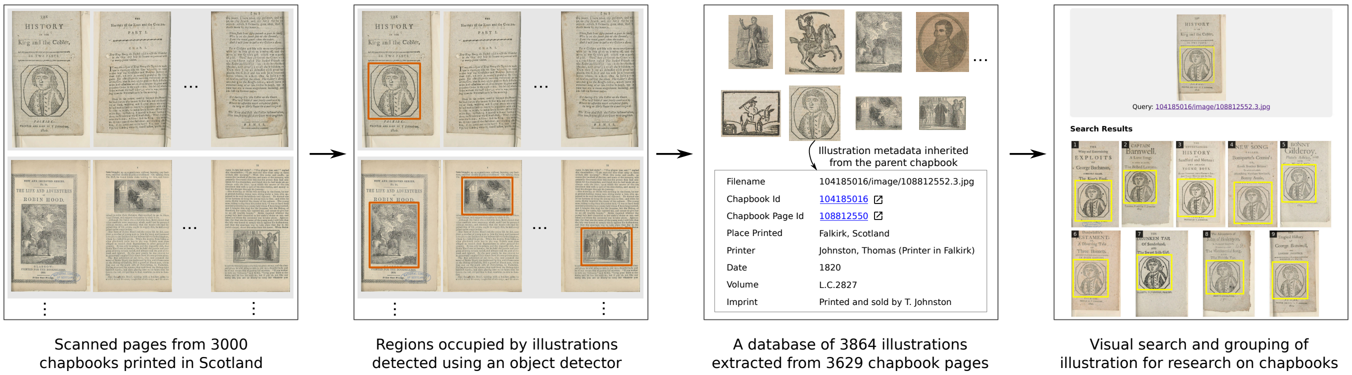




**Figure 1: An overview of the process described in this paper that enables visual analysis of chapbooks printed in Scotland [19]. An object detector is trained to detect illustration regions in pages of 3000 chapbooks. These regions enable the VGG Image Search Engine (VISE) [9] software to visually search and group illustrations based on their visual content. Such visual search capability and grouping of illustrations allows researchers to forensically analyse book pages and metadata within the chapbooks dataset.**


## ABSTRACT

Chapbooks were short, cheap printed booklets produced in large quantities in Scotland, England, Ireland, North America and much of Europe between roughly the seventeenth and nineteenth centuries. A form of popular literature containing songs, stories, poems, games, riddles, religious writings and other content designed to appeal to a wide readership, they were frequently illustrated, particularly on their title-pages. This paper describes the visual analysis of such chapbook illustrations. We automatically extract all the illustrations contained in the National Library of Scotland Chapbooks Printed in Scotland dataset, and create a visual search engine to search this dataset using full or part-illustrations as queries. We also cluster these illustrations based on their visual content, and provide keyword-based search of the metadata associated with each publication. The visual search; clustering of illustrations based on visual content; and metadata search features enable researchers to forensically analyse the chapbooks dataset and to discover unnoticed relationships between its elements. We release all annotations and software tools described in this paper to enable reproduction of the results presented and to allow extension of the methodology described to datasets of a similar nature.




*HIP '21, September 5–6, 2021, Lausanne, Switzerland*



ACM ISBN 978-1-4503-8690-6/21/09…$15.00
https://doi.org/10.1145/3476887.3476893

## CCS CONCEPTS

• **Computing methodologies → Visual content-based indexing and retrieval**.

## KEYWORDS

image search, illustration detection, visual grouping, chapbooks, printing, digital scholarship, illustration dataset

**ACM Reference Format:**
Abhishek Dutta, Giles Bergel, and Andrew Zisserman. 2021. Visual Analysis of Chapbooks Printed in Scotland. In *The 6th International Workshop on Historical Document Imaging and Processing (HIP '21), September 5–6, 2021, Lausanne, Switzerland*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3476887.3476893


## 1 INTRODUCTION

The NLS Chapbooks Printed in Scotland dataset [19] contains over 3000 digitised chapbooks, almost all printed in Scotland between the end of the seventeenth to the later nineteenth centuries. During this time, chapbook manufacture was a substantial business: over eighty firms were involved in their production, selling through travelling pedlars, or 'chapmen', to a large market in Scotland and beyond [12, 20, 23].

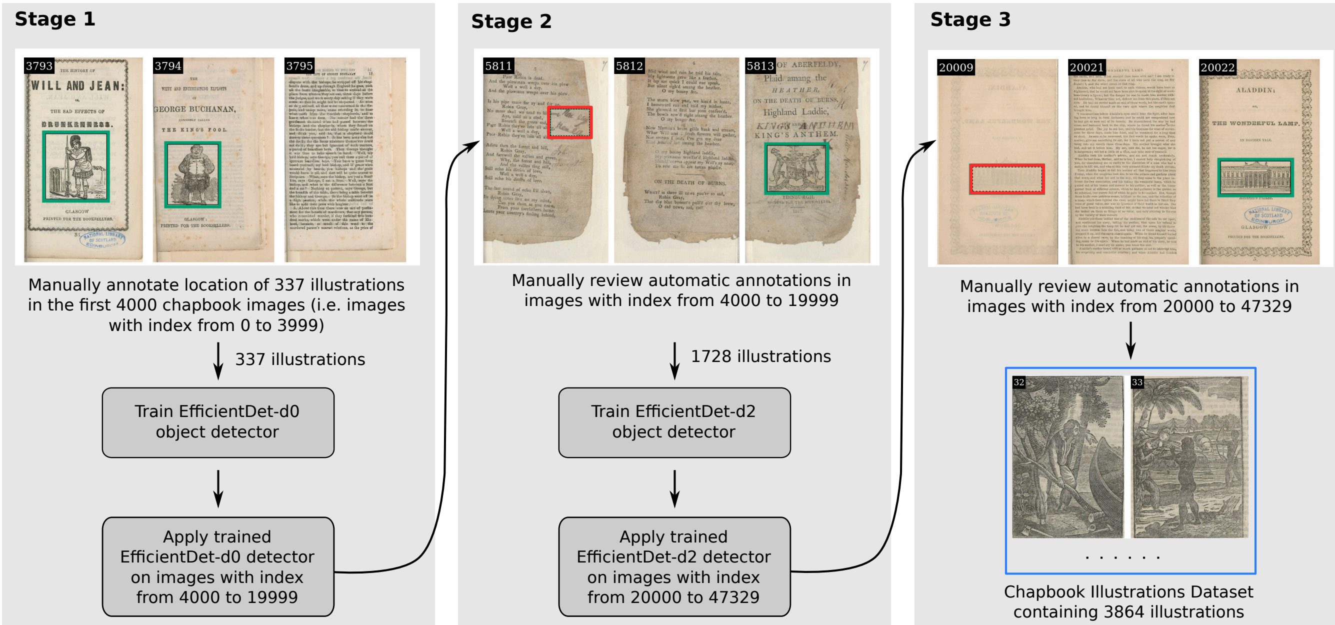



**Figure 2: Visual depiction of our semi-automatic workflow used to detect illustrations. In Stage 1, an initial set of manually-annotated illustrations are used to train an object detector which is then applied to a subset of images. In Stage 2, results from the trained detector obtained in Stage 1 are manually reviewed by a domain expert (see e.g., spurious handwritten notes, shown with red boundary and dashed black outline) resulting in 1728 illustrations. This enlarged set of illustrations is then used to train a more accurate object detector. In Stage 3, a domain expert again reviews the results (e.g., removing the false detection of illustrations on the reverse of pages), resulting in a final dataset of 3864 verified illustrations.**

Most of the chapbooks are illustrated, as shown in Figure 1 (left). The illustrations range from ornamental devices and simple representations of common objects to more detailed depictions of social events, episodes within well-known stories, portraits of famous individuals, and religious or historical narratives. They were realised by means of woodblock printing surfaces or, during the latter part of this period, derivative metal stereotype plates. The manufacture of a woodblock or stereo plate was a significant investment: each might be reused numerous times and may have become closely associated with a popular title or a specific producer. Some blocks were employed by several generations of family printers, becoming a recognisable part of their brand. The printed illustrations are therefore a valuable source of historical allowing researchers to: attribute the origin of chapbooks in the absence of other evidence; compare all the illustrations used by a particular printer; and study the pattern of block reuse among printers, uncovering business relationships.

In this paper, we describe how modern computer vision tools can greatly assist in addressing these research questions. We describe how an object detector was trained to automatically detect illustrations in chapbooks. We also describe how an image search engine was applied to group illustrations based on their visual similarity. Combining the capabilities of computer vision tools with existing, well-structured metadata, researchers can now address these – and possibly many more – research questions much more systematically. All the software code, text metadata and image annotations described in this paper are available at the following open-source code repository: https://gitlab.com/vgg/nls-chapbooks-illustrations. The methods described in this paper can easily be extended to pursue similar research goals on other image datasets.

In Section 2, we review existing research in the areas of automatic illustration detection and visual search for the study of printed books. Section 3 describes our workflow for detecting the chapbook illustrations. Visual search and grouping of chapbook illustrations is described in Section 4. Section 5 presents an overview of the keyword-based search of the metadata. Conclusions from this study are presented in Section 6.

## 2 RELATED WORK

Scholars of chapbooks have long recognised the importance of their distinctive illustrations, which provide evidence of both their production and audiences [3, 4, 17]. As libraries such as the NLS have digitised their holdings, new possibilities for studying them comparatively and at scale have emerged.

Automatic detection of illustrations in book pages and other forms of documents using computer vision has been widely explored. Hand-crafted features were used by [6] and [7] to separate illustration and text regions in historical documents. We use a deep neural network-based object detector that was trained to detect illustrations in the chapbook pages. Object detectors have also been used by [21] and [24] to automatically detect different graphical elements (e.g., illustrations and tables) in modern documents: however, their method was not evaluated on older materials.

We use the open source VGG Image Search Engine (VISE)[9] software for visual search and grouping of illustrations. VISE builds on earlier, unreleased code that powered the ImageMatch tool described in [5] as applied to early broadside ballad collections in the Bodleian Library in Oxford [16], and extended in the Image Browse and image clustering features described in [6]. VISE has also been used in [8] to visually search a collection of fifteenth-century printed book illustrations.

## 3 DETECTING ILLUSTRATION IN CHAPBOOKS

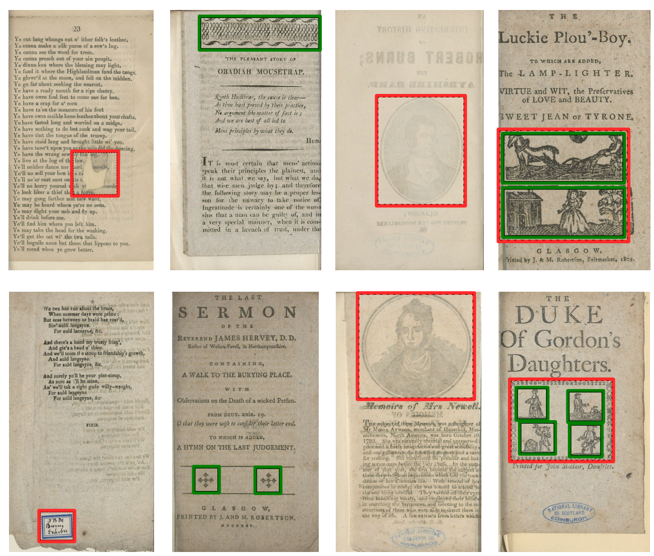
The NLS Chapbooks dataset [19] contains images of 47329 chapbook pages. Most of the pages are unillustrated, while some contain one or more as shown in Figure 1. Before we can pursue a visual analysis, we need to find out: (1) which of the pages contains an illustration and (2) their location on the page. We now describe a workflow that enables illustration detection and localisation with minimal human support.

Recent advancements in deep learning have enabled the development of a family of deep neural networks that can automatically recognise common objects like bicycles, umbrellas, buses, teddy bears, apples, etc., in large collections of images, accurately and at speed. These deep networks are called *object detectors*. Since these object detectors are already very good at detecting common objects, they can easily learn to detect novel object categories – like early printed illustrations – using a small number of exemplar images. For this work, we used the EfficientDet [22] object detector. EfficientDet is a family of neural networks for object detection that attains very high accuracy while being light in weight (both model size and computational cost). Pre-trained models of EfficientDet are available under a permissive license that allows unrestricted use in academic research projects and commercial products. EfficientDet can detect 80 commonplace, diverse objects, but as these categories do not include historical printed illustrations, we trained it for this purpose.

We conducted the training of EfficientDet object detector in two stages as shown Figure 2. In Stage 1, we manually annotated a small number of illustrations and used it to train the most light-weight version (i.e., the `efficientdet-d0` model) of the EfficientDet detector. We applied this trained detector on a larger set of images, generating automatic detections whihc were manually verified and (if necessary) corrected by our domain expert (Bergel). This manual verification step allowed us to create a larger training dataset for our Stage 2, in which we trained a more accurate version (i.e., the `efficientdet-d2` model) of the EfficientDet detector that requires a larger set of training examples. The resulting illustration detector was applied to the remaining images, detections from which were, again, manually verified to create the final dataset.

We now present a detailed description of each stage of the workflow depicted in Figure 2. For Stage 1, we manually reviewed the first 4000 images (i.e., from image index 0 to 3999) and annotated the rectangular regions containing illustrations, as shown in Figure 2 (top-left). There were 337 images containing illustrations in the first 4000 images. The (`efficientdet-d0`) model was trained on these 337 images and the trained model was applied to detect illustrations in the next 16000 images As we only had a small number of training images (i.e., 337) we chose the `efficientdet-d0` model which requires learning fewer parameters as compared to other models in the family (e.g. `efficientdet-d2`). Following manual verification, we were able to extract 1601 images containing 1728 illustrations in the first 20000 images in the dataset. In Stage 2, we trained the `efficientdet-d2`) model using a larger dataset consisting of 1728 illustrations spread across 1601 images. The resulting trained model was then applied was applied to the remaining 27329 images, yielding a further 2136 illustrations spread across 2028 images. The final dataset contains 3864 manually-verified illustrations extracted from 3629 images.

We used the open source List Annotator (LISA[1]) – part of the VIA [10] suite of manual annotation tools – to create the initial bounding-boxes and to review and correct EfficientDet-generated annotations. LISA is written in HTML, Javascript and CSS, and runs as an offline application in most modern web browsers, without requiring any installation or setup. Fitting in a single, self-contained HTML page of less than 60 kilobytes, LISA's portability and light footprint allows it to be easily shared. It provided a simple, robust interconnection between the automated and manual processes within the illustration detection pipeline.



(a) (b) (c) (d)

**Figure 3: False automatic detections are shown with red boundary (dashed black outline) while manually corrected detections are shown in green boundary (solid black outline). (a) False positive detections mostly occur for artefacts like a torn page, or library sticker. (b) False negative detections mostly correspond to book ornaments. (c) Chapbooks were printed on cheap semi-transparent paper and therefore illustrations from the reverse of pages are partially visible. This results in a small number of false positives, which are difficult to avoid. (d) The detector fails to learn the concept of part and whole as preferred by our domain expert.**

[1] https://gitlab.com/vgg/lisa


Figure 4: Example of visual search using an image region as a search query (top) with results (bottom) showing other images that are visually similar to the query.


**Figure 5: Example of a visual group of illustrations that have similar visual content (highlighted using yellow bounding box) and were most likely printed using the same printing surface, probably a stereotype plate.**

Manual review required minimal effort because there were only a small number of failure cases and the LISA tool makes it very easy to quickly review a large number of images and correct them (if required). Some examples of false detections are shown in Figure 3. The majority of false positive detections were caused by images printed on the reverse of a page, but visible on the frontside. These are hard for an automatic detector to avoid, but could potentially be addressed by post-processing techniques incorporating domain knowledge about page layout and book structure. Some printed ornaments (see Figure 3) prompted critical reflection on the nature and boundaries of the "illustration", and on the question of whether all non-textual graphical elements were potentially relevant. The domain expert (Bergel) chose to focus on those visual elements that are most characteristic of the chapbook format, including also the larger head and tailpiece ornaments but excluding ruled lines and section-breaking ornaments – these may however be of interest for the work of printer attribution, in particular. [14]

We recommend that any application of our final illustration detector be followed by a manual review stage (perhaps using the LISA application) as we have described in this paper.

We evaluated the performance of the illustration detector trained on 1728 illustrations appearing in chapbook images from 0 to 19999 using a test set that contained 2136 manually reviewed annotations in chapbook images from 20000 to 47329. The test set did not contain any illustrations seen by the detector during the training phase. Table 1 shows the precision and recall of the illustration detector at three different Intersection-Over-Union (IOU) thresholds. Results show that the detector resulting from Stage 2 is able to detect 89.2% of total illustrations with 97.3% accuracy such that all automatic detections have at least an IOU of 95% with manually annotated bounding boxes.

**Table 1: Performance of Illustration Detector on the Test Set**

| IOU threshold | Precision | Recall |
|---|---|---|
| 0.50 | 0.993 | 0.911 |
| 0.75 | 0.987 | 0.905 |
| 0.95 | 0.973 | 0.892 |

## 4 VISUAL SEARCH AND GROUPING OF ILLUSTRATIONS

Researchers can gain valuable insights into the NLS Chapbooks dataset [19] if they can visually search through the 3864 illustrations contained in this dataset using an image region as a search query (i.e. *visual search*) as shown in Figure 4.

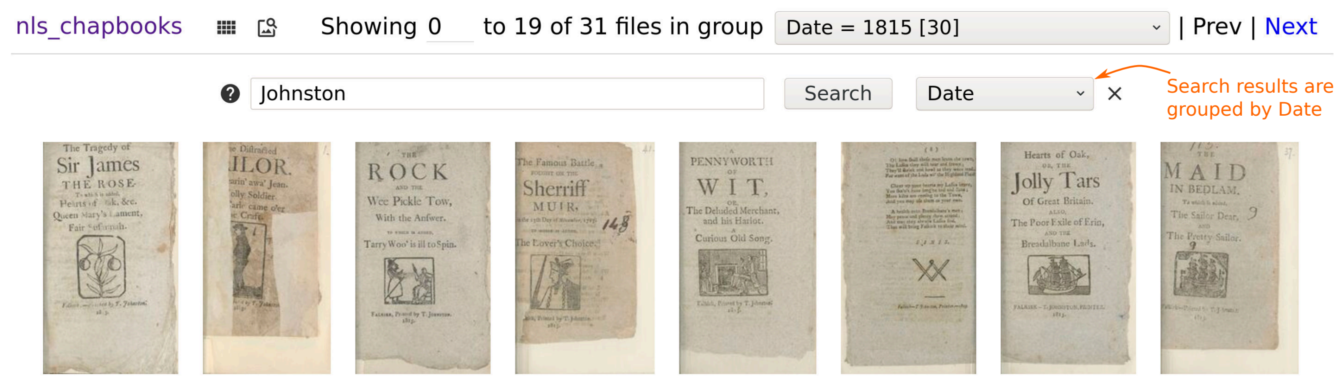



**Figure 6: An example of metadata search. All the chapbook illustrations matching the keyword "Johnston" are grouped by the metadata attribute *Date* and all such illustrations printed in 1815 (i.e. whose group value is 1815) are shown as search results.**

We use the open-source VGG Image Search Engine (VISE) [9] software to create a publicly-available[2] visual search engine based on the dataset created in Section 3. VISE operates by first applying a Hessian Affine region detector to identify a set of image regions that can be consistently detected even when the image content is scaled, rotated or translated (i.e. scale and affine invariant regions point). Next, it extracts a 128 dimensional feature vector for each detected region and uses K-means clustering technique to learn a set of cluster points (i.e. set of visual words) that can be used to efficiently and more robustly represent visual content in detected regions of an image. All the image regions detected in 3864 illustrations are represented in terms of these visual words. Visual search now boils down to efficient search and matching between the visual words assigned to the query region and visual words corresponding to all the images in the dataset. VISE uses an efficient data structure to precompute and store the visual words associated with each image in the dataset, which enables efficient storage and fast lookup of all images that contain a particular visual word. A more detailed explanation of the visual search process is present in Section 2.1.1.2 of [2].

The visual search feature provides a powerful analytical tool to researchers, enabling new discoveries at speed and at scale. In some cases, illustration matching supported an earlier attribution of a place of origin made by National Library of Scotland curators. In other cases, matching indicated an unrecorded place of origin; an unknown association between printers apparently sharing a common block.

For example, the illustration contained in the search query region of Figure 4 (top) appears on a chapbook containing no information about its origins, but for which the NLS catalogue includes a conjectured place and date of origin – Glasgow, in approximately 1790. VISE matches it with a very similar block in use on several chapbooks printed by the firm of J. & M. Robertson of Glasgow, including one with a publication date of 1809 [15]. This information is printed on the title-page and is therefore listed with confidence within the metadata in the NLS Chapbooks dataset [19]. The matches (indexed from 1 to 4) for this search query are shown in Figure 4 (bottom). Experience has shown that close visual inspection is necessary to disambiguate the copies of blocks that are present within corpora such as chapbook illustrations: computer vision-aided registration and comparison, which is provided in VISE, is a valuable tool in distinguishing between such copies. For the given example, we observe that the illustration contained in the match index 4 contains irregular, damaged areas that are consistent with the block used for the search query illustration, but also some additional regions indicating progressive deterioration of the original surface. This allows us to infer that the two publications share the same surface; that it was a wooden block rather than a metal stereotype plate; that the publication of the chapbook in the match index 4 preceded that of the chapbook in the search query; and that both were, indeed, in all likelihood printed in Glasgow.

This example illustrates how visual search empowers the researcher to pursue forensic investigations on image datasets with missing or uncertain metadata. Future systems could benefit from the establishment of training sets distinguishing between images derived from unique blocks or from multiple copies of a block. External evidence could also be further integrated, either in a rules-based system enforcing (for example) known dates or, depending on the confidence behind such evidence, as hypotheses that visual matching and comparison might usefully test.

Since we know the location of each illustration in the chapbook images, it is possible to exhaustively run search queries similar to the example shown in Figure 4 for all the 3864 illustrations. This procedure generates a *match graph* whose nodes corresponds to a unique illustration and whose edges depict a visual match between the search query node and match node. We only retain match results above a certain score threshold to ensure that only high-quality matches are represented in the match graph. The depth first search graph exploration procedure applied to the match graph allows us to discover all the visual groups (i.e. groups of illustrations that have similar visual content) in the illustration dataset. Figure 5 shows an example of such a visual group: a complete list of all such groups is available online[3].

[2] http://meru.robots.ox.ac.uk/nls_chapbooks/

[3] http://meru.robots.ox.ac.uk/nls_chapbooks/visual_group?group_id=illustration-group

## 5 TEXT METADATA SEARCH

As well as page-images the NLS chapbooks dataset [19] also contains rich metadata (e.g. printer, author, place of publication date, publisher) in a METS-XML format. The Library also supplied us with additional JSON metadata for regularised forms of such metadata, matching the terms attached to each item within the Library's public interface [18]. The chapbook illustrations dataset – described in Section 3 – inherits these metadata and therefore VISE is able to support keyword-based metadata search on the chapbook illustrations. The ability to easily move between visual and metadata-search modes was invaluable for leveraging the complementary strengths of each mode. As VISE contains a documented API, it may be integrated with various editorial workflows and data sources.

VISE uses the SQLITE [13] database engine to store and retrieve text metadata. The full-text search extension in SQLITE enables the text search bar and the `GROUP BY` clause of SQLITE is used to group the search results by a certain attribute. Figure 6 shows an example of metadata search by grouping search results using the *Date* metadata attribute. All the software code required to import metadata from NLS chapbooks dataset to VISE is available at [11].

## 6 CONCLUSION

Illustration detection and matching are of great assistance to historians of printed books. This assistance is further magnified when the visual analysis can complement prior analyses stored within well-curated metadata. As an example of "digital scholarship" – a relatively new service within libraries and museums – the project showed the value of open curated datasets for computational analysis, and the potential for such analyses to work for the mutual advancement of both research and curation. [1]

In this paper, we describe a workflow to detect, match and annotate printed illustrations within the NLS Chapbooks Printed in Scotland dataset. We publish the trained model for automatic illustration detection which may well be suitable for obtaining illustrations in similar material without retraining.

We also describe the use of VGG Image Search Engine (VISE) for visual search using both image regions and keyword-based metadata. We publicly release all the software tools and annotations described in this paper at the following code repository: https://gitlab.com/vgg/nls-chapbooks-illustrations. Since the NLS Chapbooks dataset [19] is already available from the Library's Data Foundry, interested readers should be able to reproduce all the results presented in this paper. Based on the results presented in this paper, we envisage further productive collaborations between computer vision scientists and domain experts in the field of book history.

## ACKNOWLEDGMENTS

G. Bergel was supported by the National Librarian of Scotland's Fellowship in Digital Scholarship for 2020/21. This research is funded by the EPSRC Programme Grants Seebibyte EP/M013774/1 and VisualAI EP/T028572/1, and a Royal Society Research Professorship.

## REFERENCES

[1] Sarah Ames. 2021. Transparency, provenance and collections as data: the National Library of Scotland's Data Foundry. *LIBER Quarterly* 31, 1 (2021).

[2] Relja Arandjelović. 2013. *Advancing large scale object retrieval.* Ph.D. Dissertation. University of Oxford.

[3] Iain Beavan. 2015. The chapbooks and broadsides of James Chalmers III, printer in Aberdeen: some re-discoveries and initial observations on his woodcuts. *Journal of the Edinburgh Bibliographical Society* 10 (2015), 29–86.

[4] Iain Beavan. 2019. Lines of defence: thoughts on Scottish chapbook title-page woodcuts and their functions. *Publishing History* 81 (2019), 41–5.

[5] Giles Bergel, Alexandra Franklin, Michael Heaney, Relja Arandjelović, Andrew Zisserman, and Donata Funke. 2013. Content-based image recognition on printed broadside ballads: The Bodleian Libraries' ImageMatch Tool. In *Proceedings of the IFLA World Library & Information Congress.* http://library.ifla.org/209

[6] Joon Son Chung, Relja Arandjelović, Giles Bergel, Alexandra Franklin, and Andrew Zisserman. 2014. Re-presentations of art collections. In *European Conference on Computer Vision.* Springer, 85–100.

[7] Rafi Cohen, Abedelkadir Asi, Klara Kedem, Jihad El-Sana, and Itshak Dinstein. 2013. Robust Text and Drawing Segmentation Algorithm for Historical Documents. In *Proceedings of the 2nd International Workshop on Historical Document Imaging and Processing* (Washington, District of Columbia, USA) *(HIP '13).* Association for Computing Machinery, New York, NY, USA, 110–117. https://doi.org/10.1145/2501115.2501117

[8] Cristina Dondi, Abhishek Dutta, Matilde Malaspina, and Andrew Zisserman. 2020. *The Use and Reuse of Printed Illustrations in 15th-Century Venetian Editions.* Edizioni Ca' Foscari - Digital Publishing, Chapter 30, 839–869. https://doi.org/10.14277/978-88-6969-332-8/030

[9] Abhishek Dutta, Relja Arandjelović, and Andrew Zisserman. 2021. *VGG Image Search Engine.* Retrieved May 05, 2021 from https://www.robots.ox.ac.uk/~vgg/software/vise/

[10] Abhishek Dutta and Andrew Zisserman. 2019. The VIA Annotation Software for Images, Audio and Video. In *Proceedings of the 27th ACM International Conference on Multimedia* (Nice, France) *(MM '19).* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3343031.3350535

[11] Abhishek Dutta and Andrew Zisserman. 2021. *Code for Visual Analysis of Chapbooks Printed in Scotland.* Retrieved May 05, 2021 from https://gitlab.com/vgg/nls-chapbooks-illustrations

[12] Adam Fox. 2013. 'Little Story Books' and 'Small Pamphlets' in Edinburgh, 1680–1760: The Making of the Scottish Chapbook. *Scottish Historical Review* 92, 2 (2013), 207–230.

[13] Richard D Hipp. 2020. *SQLite.* https://www.sqlite.org/index.html

[14] Scottish Book Trade Index. [n.d.]. Entered According To Order. Retrieved May 14, 2021 from https://data.cerl.org/sbti/002219

[15] Scottish Book Trade Index. [n.d.]. Robertson, James and Matthew. Retrieved May 14, 2021 from https://data.cerl.org/sbti/006372

[16] Bodleian Libraries. [n.d.]. Bodleian Ballads Online. Retrieved May 14, 2021 from http://ballads.bodleian.ox.ac.uk

[17] Barry McKay. 1998. Cumbrian chapbook cuts: some sources and other versions. In *The Reach of Print: Making, Selling and Using Books*, Peter Isaac and Barry Mckay (Eds.). St. Paul's Bibliographies, Winchester, 65–83.

[18] National Library of Scotland. [n.d.]. Chapbooks printed in Scotland. Retrieved May 14, 2021 from https://digital.nls.uk/chapbooks-printed-in-scotland/

[19] National Library of Scotland. 2019. *Chapbooks printed in Scotland.* Retrieved May 05, 2021 from https://doi.org/10.34812/vb2s-9g58

[20] G Ross Roy. 1974. Some notes on Scottish chapbooks. *Scottish Literary Journal* 1 (1974), 50–60.

[21] Ranajit Saha, Ajoy Mondal, and C V Jawahar. 2019. Graphical Object Detection in Document Images. In *2019 International Conference on Document Analysis and Recognition (ICDAR).* 51–58. https://doi.org/10.1109/ICDAR.2019.00018

[22] Mingxing Tan, Ruoming Pang, and Quoc V Le. 2020. Efficientdet: Scalable and efficient object detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition.* 10781–10790.

[23] Alastair R Thompson. 1972. Chapbook printers. *The Bibliotheck; a Scottish Journal of Bibliography and Allied Topics* 6, 3 (1972), 76.

[24] Xiaohan Yi, Liangcai Gao, Yuan Liao, Xiaode Zhang, Runtao Liu, and Zhuoren Jiang. 2017. CNN Based Page Object Detection in Document Images. In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, Vol. 01. 230–235. https://doi.org/10.1109/ICDAR.2017.46