

Classifying Medieval Manuscripts by Pen and Support

Sharva Gogawale^{*} , Omer Ventura^{*}, Daria Vasyutinsky-Shapira^{*} , Berat Kurar-Barakat^{*} , Gal Grudka^{*}, Mohammad Suliman^{*}, Iddo Hakim^{*}, and Nachum Dershowitz^{*}

^{*} School of Computer Science and AI, Tel Aviv University, Ramat Aviv, Israel

Abstract

We present a machine-learning approach for classifying medieval Hebrew manuscripts by two key material attributes: writing support (whether the substrate is paper or parchment) and writing implement (quill pen vs. reed calamus). Our work contributes to the emerging field of computational codicology, offering tools to aid paleographers in the large-scale analysis of digitized manuscripts. Our datasets—derived from existing digitized repositories—

have been carefully annotated and balanced to capture the range of material and stylistic variation. For both classification tasks, we employ convolutional neural networks tailored to their respective challenges: identifying broad substrate textures and capturing fine-grained stroke morphology. The support classifier achieved an accuracy of 91% and demonstrated reliable performance even on visually ambiguous examples. Likewise, the implement classifier was 91.5% accurate. These findings show that computational analysis can aid and, in some cases, surpass manual paleographic methods in analyzing historical manuscripts. This work highlights the potential of computational tools to assist scholars in large-scale analysis of digitized corpora, aiding manuscript dating, provenance research, and the study of scribal practices.

Keywords: manuscripts, pen, support, substrate, paper, parchment, quill, calamus, classifiers, CNN

1 Introduction

The National Library of Israel, as part of its Ktiv project,¹ has been digitizing most of the surviving approximately one hundred thousand Hebrew-character manuscripts that sit in libraries and collections in Europe and elsewhere [14]. About ten million images are expected with the completion of the project in the near future. Unfortunately, precious few of the manuscripts include the date or location of its writing.

For centuries, parchment served as the primary material for a wide array of media, ranging from sumptuously decorated religious texts to practical everyday documents. This writing support

remains an invaluable source of historical, artistic, cultural, and even biological information [11]. The use of skins—those of calves, sheep and goats—for the writing of scrolls was standard among Jews already at around the time of the canonization of the Hebrew Bible two millennia ago. The type of processing of the skin depended on the type of manuscript. Most of the Dead Sea scrolls are written on *gevil*, a hide processed for writing on one side only, as required by the scroll format. Medieval manuscripts, which often take the form of a codex (bound book form), were made from parchment, with the leather skin processed for writing on both sides. The even earlier type of writing support, papyrus—made of leaves of the papyrus sedge, was still used for

Sharva Gogawale, Omer Ventura, Daria Vasyutinsky-Shapira, Berat Kurar-Barakat, Gal Grudka, Mohammad Suliman, Iddo Hakim, and Nachum Dershowitz. “Classifying Medieval Manuscripts by Pen and Support.” In: *Computational Humanities Research 2025*, ed. by Taylor Arnold, Margherita Fantoli, and Ruben Ros. Vol. 3. Anthology of Computers and the Humanities. 2025, 1349–1359. <https://doi.org/10.63744/tL5xGcaScd42>.

© 2025 by the authors. Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0).

¹ <https://www.nli.org.il/en/discover/manuscripts/hebrew-manuscripts>



Figure 1: Sample patches of four medieval manuscripts in Hebrew characters. Top left: calamus on parchment; top right: quill on parchment; bottom left: calamus on paper; bottom right: quill on paper. They are written in square and semi-cursive script modes, in Hebrew, Aramaic, and/or Judeo-Arabic.

many Judaean desert documents in classical times, but, by the Middle Ages, parchment gradually became the standard writing support due to its flexibility, which made it better suited to the codex structure while lending itself to writing on both sides of a leaf.

Thus, in the Middle Ages, Hebrew manuscripts were written on two kinds of substrates only: parchment made from skins of sheep and cattle, or paper, produced from linen rags, or sometimes—in the Mediterranean—from cotton. The use of parchment preceded the use of paper by hundreds of years, and in the Islamic lands the transition from the expensive parchment to cheap paper happened much earlier than in Christian countries.

From the 11th century onward, paper becomes the standard writing support for Hebrew manuscripts in the Mediterranean, and later in North Africa and Spain. Although, in the entire corpus of the dated medieval Hebrew manuscripts, the amounts of those written on parchment and on paper are nearly equal, for some regions and regional script types the situation is different. There are, in fact, only 28 known, dated Oriental Hebrew manuscripts on parchment (about 8% of all dated Oriental manuscripts), and all of them were copied before the early 14th century. On the other hand, in Ashkenaz (Germany and Central Europe), the percentage of dated manuscripts on parchment is 82%, and reaches 100% in some centuries. See the book on Hebrew codicology by Malachi Beit-Arié [2, pp. 217 ff.].

Recognizing the writing support of a manuscript can in many cases, as indicated above, contribute to its accurate dating and to the determination of the place of copying. The visual appearance and characteristics of paper and parchment (especially of the hair side) are often highly informa-

tive. The cases when a single manuscript is written partly on paper and partly on parchment, that is, paper quires wrapped in parchment bifolia, are also indicative. This practice was more pronounced in Byzantium and was never adopted in the Middle East or in Ashkenaz.

Finally, when one recognizes that some leaves of a manuscript are written on parchment and others on paper but the two are not intermixed, this suggests that it may be a collection of disparate works bound together. Many such items are yet to be identified, so they are still often mistakenly cataloged as a single work.

Differentiation between paper and parchment can easily be performed *in situ* when one can touch the manuscript pages. Determining the writing support by image is sometimes easy, for example when one sees hair follicles on the parchment, but sometimes the two dimensional nature of an image makes this task challenging. This is one of the cases when an algorithm can be more efficient than a human researcher.

Calamus, a reed pen, was the principle writing instrument in the Orient. In Europe, the quill (a trimmed flight feather of a large bird) was introduced and beginning in the 6th century gradually replaced the calamus, which led to a significant change in writing techniques. In the Italian regional type of medieval Hebrew script, both implements are attested. Thus, calamus and quill are associated with different script types. In addition, there may be subtle differences in the morphology (shape) and ductus (flow) of the script depending on the writing implement.

We have prepared datasets of images of medieval Hebrew manuscripts, labeled by writing material—whether paper or parchment—and by writing implement—whether quill or reed pen, and have trained neural networks to classify images of manuscript pages into each of the two binary categories. Figure 1 displays a few manuscript patches showcasing the diversity of writing tools and substrates captured in our datasets.

2 Related Work

In recent years, there have been extensive efforts to analyze historical manuscripts using computational tools. These focused primarily on script classification, dating, and the like. Codicological features have been measured computationally or with computational aids in [6; 18], for example.

A thorough investigation into the physical characteristics of the writing medium helps provide a more comprehensive understanding of their provenance and historical context. This should include identifying the main components of the support, writing, and miniature materials, and assessing the evolution of their conservation state in correlation with intrinsic and environmental factors that have impacted their physical condition and visual appeal. The substrate and the writing tool, despite their critical role in understanding historical contexts and scribal practices across different cultural and historical contexts, have received comparatively less attention.

2.1 Support Analysis

Advancements in analytical techniques are continuously enhancing our understanding of the materials used for writing. Characterizing them is essential for proper preservation and serves as important evidence for codicological studies. For example, X-ray fluorescence (XRF) spectroscopy has been applied in [8] for compositional analysis, particularly for elemental mapping of inks, pigments, parchment, and paper. An X-ray diffraction (XRD) study on 18th–19th-century South American manuscripts incorporated XRF to analyze their composition [15]. Non-invasive investigations have also been conducted using multi- and hyper-spectral imaging, Fourier transform infrared (FTIR) spectroscopy, and XRF spectroscopy to obtain crucial information on parchment manufacturing techniques and original materials [7]. Furthermore, FTIR has been specifically applied to investigate degradation in parchments of various origins [3]. In a similar vein, attenuated

total reflectance FTIR, micro-Raman, and XRF spectroscopies were employed to analyze fragments of different Yemenite manuscripts on both parchment and paper [4]. In [13], a multispectral thresholding and energy minimization (MTEM)-based pipeline for Dead Sea scroll fragments was developed, enabling the distinction and segmentation of ink from parchment by leveraging the characteristic spectral signatures of the parchment. A classifier based on the two-dimensional Fourier transform was proposed in [9; 16] to capture surface-level texture characteristics for distinguishing material substrates.

2.2 Implement Analysis

Traditional techniques for determining the specific writing tool employed (such as quill or calamus) rely on paleographic observations of stroke morphology, pressure variations, and ink flow characteristics. Foundational paleographic studies, such as [5], establish links between writing implements and stroke characteristics, showing that instruments like quills and reed pens produce distinct visual characteristics. These differences are seen in the thickness and fluidity of the strokes and stem from factors such as the design of the nib, ink retention, and the way the writer holds and presses the tool while writing.

In [19], frameworks were introduced to extract tool-specific features such as pen pressure, writing movement, and stroke-width variability using computational feature extraction and machine-learning techniques.

3 Datasets

The manuscript images used in this work originate from the SfarData² and HebrewPal³ projects, each of which provides images of Hebrew manuscripts from a variety of historical sources, along with substantial metadata.

SfarData is the unique database developed by the Hebrew Paleography Project. Since 1965, a group of researchers under the direction of the late Malachi Beit-Arié conducted a pioneering effort to catalog and analyze all dated or datable medieval Hebrew manuscripts produced before 1540. During decades of work, these manuscripts were studied *in situ* in over 250 public and private collections worldwide, and meticulously described codicologically (describing the writing support) with some paleographic features (describing the script). Since the 1990s, the handwritten data has been gradually digitized and, after a series of transformations, SfarData today exists in the form of an online database housed by the National Library of Israel [1].

HebrewPal is the brainchild of paleographer, Judith Olszowy-Schlanger. Launched in 2022, it is being developed to include exemplars of Hebrew books, documents, and epigraphy from all regions of the Jewish diaspora. Images are annotated with a wealth of paleographic data, including fine-grained features of select letters. Traditional Hebrew paleography is being combined with the latest developments in digital technology as part of an ERC-funded project, “MiDRASH – Migrations of Textual and Scribal Traditions via Large-Scale Computational Analysis of Medieval Manuscripts in Hebrew Script,” a computational-humanities effort aiming to reconstruct medieval Jewish book culture [21].

Our dataset images were carefully selected by the paleographer on our team to constitute a representative sample of the diversity in writing tools and substrates, ensuring that both material properties (paper or parchment) and writing characteristics (quill or calamus) are well-represented for robust classification. To ensure accuracy of the annotations, we used SfarData and HebrewPal as sources of metadata for the support and implement tags, respectively.

² https://sfardata.nli.org.il/#/startSearch_En

³ <https://www.hebrewpalaeography.com>

Items	Quantity
Manuscripts	1,250
Paper	456
Parchment	794

Table 1: Writing support dataset statistics.

Items	Quantity
Manuscripts	80
Quill/Calamus (count)	40/40
Patches	720
Lines (estimate)	4,320

Table 2: Writing implement dataset statistics.

n.b. *The datasets will be gladly shared upon request for non-commercial research purposes. The source code for our implementation is publicly available on GitHub.*

3.1 Support Dataset

For the substrate classification task (see Table 1), we curated a corpus of 1,250 manuscript images, based on material annotations extracted from SfarData and the corresponding images from Ktiv. Black-and-white images were excluded to ensure consistency in visual characteristics. Among these, 456 manuscripts were identified as paper and 794 as parchment. To mitigate class imbalance, two page images were sampled from each paper manuscript and one from each parchment manuscript, resulting in 912 paper and 794 parchment page images (a total of 1,706 page images) derived from 1,250 manuscripts.

3.2 Implement Dataset

For the writing-tool classification task (see Table 2), we curated a corpus based on annotations from HebrewPal and high-resolution color images obtained from Ktiv. The dataset includes 80 manuscripts, of which 40 were written with a quill and 40 with a calamus. One page image was sampled from each manuscript, resulting in a balanced dataset comprising 40 quill and 40 calamus page images derived from distinct manuscripts.

4 Method

Our methodology is centered on two parallel classification pipelines to analyze medieval Hebrew manuscripts, each with an approach tailored to its specific target binary classification task. For *writing implement classification* (quill vs. calamus), we employ a patch-based deep-learning model, allowing the model to focus more on localized visual details (fine-grained morphological features of the script), such as stroke thickness, curvature, and ink dispersion. In contrast, *writing support classification* (parchment vs. paper) requires a more global perspective. The objective is to analyze the broader textural characteristics of the substrate itself, emphasizing background texture and material characteristics, which are best assessed across different areas of the page.

4.1 Support Classification: Paper vs. Parchment

For the writing support classification task, we aim to distinguish manuscript pages based on whether they were written on paper or parchment. Given the nature of this problem, we adopted a page-level approach leveraging a convolutional neural network (CNN) to capture global surface characteristics and material textures indicative of the underlying substrate. The entire manuscript page was used as input, allowing the model to learn from visual cues present in regions surrounding the text, such as the exposed substrate around letterforms, as well as at the page edges where material properties like texture and fiber patterns are more pronounced. The encoder is designed to learn the subtle material signals that differentiate parchment from paper, focusing on variations that may be challenging for human observers without physical examination.

We employed an EfficientNet-based architecture [20] pretrained on ImageNet, which we fine-tuned for binary classification of paper versus parchment. EfficientNet was chosen for its compound scaling strategy, which balances depth, width, and resolution, making it well-suited to capture fine-grained distinctions in manuscript images while maintaining computational efficiency. The images were resized to a consistent resolution, selected to retain the visual granularity of substrate textures critical for classification, and normalized pixel intensity values to standardize input distributions across the dataset. The resulting classifier is able to capture differences that are difficult for human experts to replicate manually at scale, providing a scalable and generalizable approach for automated material identification across a wide range of manuscripts.

4.2 Implement Classification: Quill vs. Calamus

To automatically differentiate between a manuscript written using a quill or with a calamus, we adapted a deep learning framework to this specific task. The core backbone of the method is a CNN that learns to distinguish the subtle visual features associated with each tool. High-resolution manuscript images were preprocessed, with each page being divided into smaller, non-overlapping patches of 350×350 pixels. This expands the dataset by several orders of magnitude while preserving meaningful visual features.

Paleographic studies suggest that patches containing around five lines of text are adequate for identifying the script type. Accordingly, the patch generation process ensured consistent text scale and inclusion of sufficient text lines [10]. To ensure that only informative and content-rich patches are retained, each candidate patch undergoes a filtering step based on pixel intensity statistics: (i) Patches with too much white space or low variation

(ii) A statistical threshold is applied using heuristics like entropy or pixel mean/variance to detect and remove low-texture regions. This significantly improves the signal-to-noise ratio in the dataset, ensuring that models train only on visually meaningful data.

For the CNN architecture, we fine-tuned a ResNet18 [12] model pretrained on ImageNet, applying standard image transformations (resizing, normalization). Notably, standard data augmentation techniques such as random rotations or flips were deliberately avoided to preserve the original orientation and directionality of the script’s strokes, which are crucial features for distinguishing between the flexible marks of a quill and the more rigid, angular lines of a calamus.

To mitigate the potential risk of overfitting, we employed several strategies at both the data and training levels. First, we used a patch-based approach to expand our dataset to 720 distinct image samples, providing the model with more varied examples. Second, we enforced a strict manuscript-level split, ensuring no manuscript overlap between the training and test sets. This validation strategy ensures the model’s ability to generalize to entirely unseen manuscripts, rather than just new sections of familiar ones. Finally, during training, we further regularized the model using early stopping and weight decay.

5 Results

We evaluate the performance of our approaches using standard classification metrics including accuracy, precision, recall, and F_1 -score, providing a comprehensive assessment of model effectiveness.

5.1 Support Classification

Paleographers are not trained to determine the type of writing support solely from an image of a manuscript page. As with other codicological tasks, it is performed physically, by touching and

Class	Precision	Recall	F_1 -score	Quantity
Paper	0.96	0.87	0.91	136
Parchment	0.86	0.96	0.91	120
Total accuracy = 91%				

Table 3: Writing support classification metrics.

Actual	Predicted	
	Paper	Parchment
Paper	118	18
Parchment	5	115

Table 4: Writing support confusion matrix.

turning (and smelling) the pages of a codex. This makes it hard to evaluate algorithm results compared to human performance.

For most types of writing support, especially for parchment with noticeable signs on the hair side or for paper with visible texture, an image suffices to differentiate between them. However, Ashkenazi manuscripts since the end of the 13th century (that is, in most cases, manuscripts produced in Germany—because of the 14th century expulsions of the Jews from France) were written on parchment processed in such a way that the hair side and the flesh side became virtually indistinguishable. In such cases, as well as in the numerous cases of very fine smooth paper with almost no signs of texture, differentiating between writing supports by image only, and without the use of paleographic features, is problematic. For this reason and for this specific task, we believe that the algorithm is at least as effective as—and probably superior to—a flesh-and-blood scholar.

Our EfficientNet-based support classifier achieved a total accuracy of 91%, demonstrating strong performance even on challenging manuscript images where visual cues distinguishing paper and parchment are subtle.

Table 3 reports aggregate classification metrics, while Table 4 presents the confusion matrix summarizing correct and incorrect predictions at the page level.

5.2 Implement Classification

At the patch level, the ResNet18-based writing-implement classifier achieved strong performance with an accuracy of 88.5%, precision of 88.6%, and F_1 -score of 87.1%. To obtain page-level predictions, patch-wise outputs were aggregated using a majority voting strategy, yielding an accuracy of 91.5%, precision of 93.5%, and F_1 -score of 90.9%. These results, summarized in Figure 2, demonstrate that local visual cues effectively contribute to writing implement attribution, while aggregation improves robustness significantly at the page level.

To interpret the model’s predictions, we visualize the most salient features using Grad-CAM [17], generating localization maps that highlight the regions of the manuscript images that were responsible for the models’ decisions. See Figure 3. These visualizations confirm that the models are indeed focusing on key stylistic cues relevant to this classification task. However, the visualizations also hint that the model is focusing on different features than those to which a human paleographer would pay attention.

The Grad-CAM visualizations suggest that the writing-implement model pays attention primarily to stroke morphology and junctions. The highlighted regions in the heatmaps often cover relatively broad areas, so not every activation can be tied to a single discrete paleographic feature. However many of the high activation zones coincide with serifs, hairlines, and ascenders of the letter *lamed* with flags. These are the regions where differences between calamus and quill

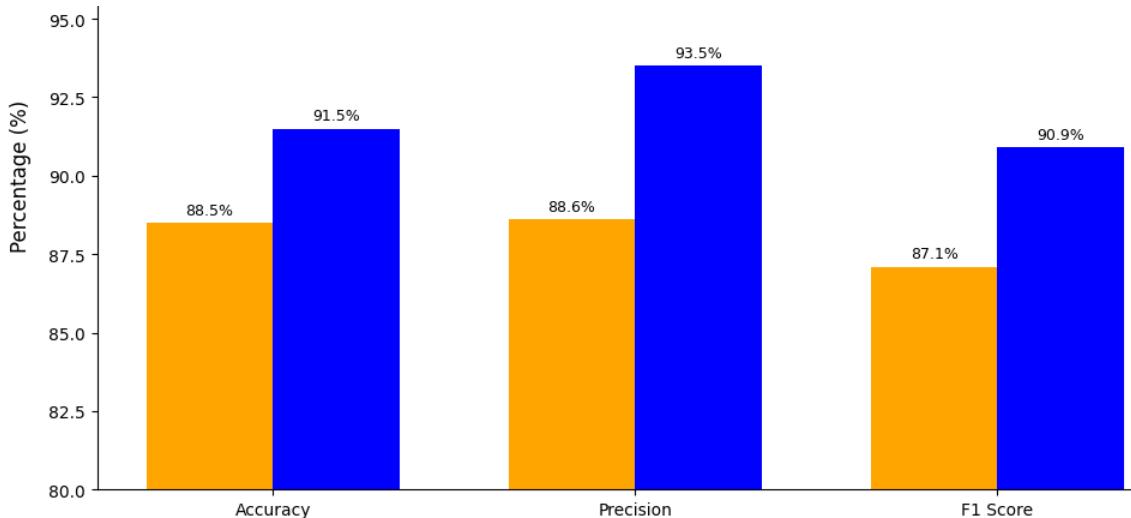


Figure 2: Model evaluation metrics: Patch-level (orange) and page-level (blue) performance of writing-implement classification.

become most apparent. The difference is—first and foremost—the ability of a quill to produce very thin lines that are connected to much thicker ones (shading), and special forms such as “fish-tails,” hairlines, etc. This alignment suggests that the model is, at least in part, learning genuine implement-related features.

6 Conclusion

In this study, we developed a machine-learning framework to classify medieval Hebrew manuscripts by writing support (paper vs. parchment) and writing implement (quill vs. calamus). By curating carefully annotated datasets extracted from digitized repositories (SfarData, Hebrew-Pal, and Ktiv) and employing convolutional neural networks tailored to each task, we demonstrated strong classification performance, with accuracy exceeding 90% for both of the tasks. The results show that algorithmic methods can match or surpass human expertise in certain codicological tasks where tactile examination is not feasible, providing a scalable approach for large-scale manuscript analysis.

While our models were trained exclusively on Hebrew-character manuscripts, the underlying tasks of textural and stroke-morphology analysis are not inherently script-dependent. Therefore, we believe that this approach shows promise for generalization and could be applied to manuscripts in other languages and scripts as well. Additional material substrates, such as papyrus, with its distinctive fibrous structure, could easily be incorporated into such a classification scheme.

This research opens pathways for combining material analysis with script and content analysis, ultimately contributing to a deeper and more systematic understanding of medieval manuscript production across cultures.

Acknowledgments

We are grateful to the HebrewPal team for their contribution to the construction of the dataset. We would also like to thank the reviewers for their valuable comments and suggestions, which helped improve the quality of this paper.

This research was funded by the European Union (ERC, MiDRASH, Project No. 101071829. Principal investigators: Nachum Dershowitz, Tel Aviv University; Judith Olszowy-Schlanger,



Figure 3: Grad-CAM visualizations highlighting activations focused around letter forms that strongly influence writing-implement classification.

EPHE-PSL; Avi Shmidman, Bar-Ilan University, and Daniel Stökl Ben Ezra, EPHE-PSL). Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

References

- [1] Beit Arié, Malachi. “SFARDATA: The Henri Schiller Codicological Database of the Hebrew Palaeography Project, Jerusalem”. In: *Gazette du Livre Médiéval* 25 (automne 1994), pp. 24–29. DOI: 10.3406/galim.1994.1280.
- [2] Beit-Arié, Malachi. *Hebrew Codicology: Historical and Comparative Typology of Medieval Hebrew Codices based on the Documentation of the Extant Dated Manuscripts until 1540 using a Quantitative Approach*. I. Goldberg, trans. Jerusalem: The Israel Academy of Sciences and Humanities, 2021. DOI: 10.25592/uuhfdm.9349.
- [3] Bicchieri, Marina, Monti, Michela, Piantanida, Giovanna, Pinzari, Flavia, and Sodo, Armida. “Non-destructive Spectroscopic Characterization of Parchment Documents”. In: *Vibrational Spectroscopy* 55 (Mar. 2011), pp. 267–272. DOI: 10.1016/j.vibspec.2010.12.006. URL: https://www.academia.edu/19761407/Non_destructive_spectroscopic_characterization_of_parchment_documents.
- [4] Bicchieri, Marina, Monti, Michela, Piantanida, Giovanna, and Sodo, Armida. “Non-destructive Spectroscopic Investigation on Historic Yemenite Scriptorial Fragments: Evidence of Different Degradation and Recipes for Iron Tannic Inks”. In: *Analytical and Bioanalytical Chemistry* 405 (Jan. 2013), pp. 2713–2721. DOI: 10.1007/s00216-012-6681-4. URL: https://www.researchgate.net/publication/234105213_Non-destructive_spectroscopic_investigation_on_historic_Yemenite_scriptorial_fragments_Evidence_of_different_degradation_and_recipes_for_iron_tannic_inks.
- [5] Bischoff, Bernhard. *Latin Palaeography: Antiquity and the Middle Ages*. Dáibhí Ó Cróinín, and David Ganz, trans. Cambridge, UK: Cambridge University Press, 1990. DOI: 10.1017/CBO9780511809927.

- [6] Causo, Serena. “Enhancing Data Collection on the Materiality of Papyri: The Measurement Tool”. In: *Digital Papyrology III*, ed. by N. Reggiani. De Gruyter, Dec. 2025, pp. 257–284. DOI: 10.1515/9783111070162-015.
- [7] Cornea, Ioana Maria, Ghervase, Luminița, Ratoiu, Lucian, and Răduan, Roxana. “Application of Spectroscopic and Hyperspectral Imaging Techniques for Rapid and Nondestructive Investigation of Jewish Ritual Parchment”. In: *Frontiers in Materials* 7 (Dec. 2020). Art. 601339. DOI: 10.3389/fmats.2020.601339.
- [8] de Viguerie, Laurence, Rochut, Sophie, Alfeld, Matthias, Walter, Philippe, Astier, Sophie, Gontero, Valérie, and Bouc'h, Florence. “XRF and Reflectance Hyperspectral Imaging on a 15th Century Illuminated Manuscript: Combining Imaging and Quantitative Analysis to Understand the Artist’s Technique”. In: *Heritage Science* 6 (Apr. 2018). Art. 11. DOI: 10.1186/s40494-018-0177-2.
- [9] Dhali, Maruf A, Reynolds, Thomas, Alizadeh, Aylar Ziad, Nijdam, Stephan H, and Schomaker, Lambert. “Pattern Recognition Techniques in Image-Based Material Classification of Ancient Manuscripts”. In: *International Conference on Pattern Recognition Applications and Methods*. Springer. 2023, pp. 124–150.
- [10] Droby, Ahmad, Kurar Barakat, Berat, Vasyutinsky Shapira, Daria, Rabaev, Irina, and El-Sana, Jihad. “VML-HP: Hebrew Paleography Dataset”. In: *Proceedings of the 16th International Conference on Document Analysis and Recognition (ICDAR 2021, Lausanne, Switzerland)*, ed. by J. Lladós, D. Lopresti, and S. Uchida. *Lecture Notes in Computer Science*, vol. 12824. Cham: Springer, Sept. 2021, pp. 205–220. DOI: 10.1007/978-3-030-86337-1_14. URL: https://www.researchgate.net/publication/354359715_VML-HP_Hebrew_Paleography_Dataset.
- [11] Fiddyment, Sarah, Teasdale, Matthew D., Vnouček, Jiří, Lévéque, Élodie, Binois, Annelise, and Collins, Matthew J. “So you Want to do Biocodicology? A Field Guide to the Biological Analysis of Parchment”. In: *Heritage Science* 7 (June 2019). Art. 35. DOI: 10.1186/s40494-019-0278-6.
- [12] He, Kaiming, Zhang, Xiangyu, Ren, Shaoqing, and Sun, Jian. “Deep Residual Learning for Image Recognition”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90. URL: https://openaccess.thecvf.com/content_cvpr_2016/papers/He_Deep_Residual_Learning_CVPR_2016_paper.pdf.
- [13] Kurar-Barakat, Berat and Dershowitz, Nachum. “Segmentation of Ink and Parchment in Dead Sea Scroll Fragments”. Nov. 2024. DOI: 10.48550/arXiv.2411.10668. arXiv: 2411.10668 [cs.CV].
- [14] National Library of Israel. “Digitized Hebrew Manuscripts”. Dec. 2021. URL: <https://web.nli.org.il/sites/nlis/en/manuscript>.
- [15] Obregón, Celina Luízar, Jara, Marco A. Zamalloa, Arizapana, Flor L. Rojas, Huayllani, Yuri J. Chura, Bellido, Janet F. Gonzales, and Olivera, Jorge Olivera. “XRF Elemental Analysis of Inks in South American Manuscripts from 1779 to 1825”. In: *Heritage Science* 9 (Nov. 2021). Art. 147. DOI: 10.1186/s40494-021-00619-9.
- [16] Reynolds, Thomas, Dhali, Maruf, and Schomaker, Lambert. “Image-Based Material Analysis of Ancient Historical Documents”. In: *Proceedings of the 12th International Conference on Pattern Recognition Applications and Methods (ICPRAM)*. Vol. 1. SCITEPRESS – Science and Technology Publications, 2023, pp. 697–706. DOI: 10.5220/0011743700003411. URL: <https://arxiv.org/pdf/2203.01042>.

- [17] Selvaraju, Ramprasaath R., Cogswell, Michael, Das, Abhishek, Vedantam, Ramakrishna, Parikh, Devi, and Batra, Dhruv. “Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization”. In: *International Journal of Computer Vision* 128 (Oct. 2019), pp. 336–359. DOI: 10.1007/s11263-019-01228-7. URL: https://ramprs.github.io/static/docs/IJCV_Grad-CAM.pdf.
- [18] Shweka, Roni, Choueka, Yaacov, Wolf, Lior, and Dershowitz, Nachum. “Automatic Extraction of Catalog Data from Digital Images of Historical Manuscripts”. In: *Literary and Linguistic Computing* 28 (Feb. 2013), pp. 315–330. DOI: 10.1093/llc/fqt007. URL: <https://d1wqtxts1xzle7.cloudfront.net/30887601/fqt007-libre.pdf>.
- [19] Srihari, Sargur N., Cha, Sung-Hyuk, Arora, Hina, and Lee, Sangjik. “Individuality of Handwriting”. In: *Journal of Forensic Sciences* 47 (July 2002), pp. 856–872. DOI: 10.1520/JFS15447J. URL: https://cedar.buffalo.edu/NIJ/nij_ref.pdf.
- [20] Tan, Mingxing and Le, Quoc V. “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks”. Sept. 2020. DOI: 10.48550/arXiv.1905.11946. arXiv: 1905.11946 [cs.CV].
- [21] Vasyutinsky Shapira, Daria, Kurar-Barakat, Berat, Gogawale, Sharva, Suliman, Mohammad, and Dershowitz, Nachum. “MiDRASH – A Project for Computational Analysis of Medieval Hebrew Manuscripts”. In: *EUROGRAPHICS Workshop on Graphics and Cultural Heritage*, ed. by M. Corsini, D. Ferdani, A. Kuijper, and H. Kutlu. Sept. 2024. DOI: 10.2312/gch.20241267.