Journal of Social Research

E-ISSN: 2828-335x
Vol. 5, No. 11, November 2026, pp. 6042-6052
DOI : 10.55324/josr.v5i11.3599
Journal of Social Research

Computational Analysis of Ancient Serat Yusup Manuscripts: A Multi-Methodological Approach for Small Archaeological Datasets

Restin Alfinda Zai*, Imam Yuadi
Universitas Airlangga Surabaya, Indonesia
Email : restin.alfinda.zai-2025@pasca.unair.ac.id*

Keywords:
Computational Archaeology; Manuscript Analysis; Small Datasets; Clustering; Classification.

This is an open access article under the CC BY-SA license.

Copyright © 2026 by Author. Published by International Journal Labs

ABSTRACT

This study introduced an integrated computational framework for analyzing historical Serat Yusup manuscripts, specifically designed to address the methodological challenges associated with small cultural heritage datasets. The framework was applied to six high-resolution manuscript images obtained from the National Library of Indonesia, from which 23 numerical features were extracted across four categories: dimensional, color, texture, and structural characteristics. The methodology integrated clustering, classification, and anomaly detection techniques adapted to the constraints of limited sample sizes. To address these limitations, the analytical procedures incorporated leave-one-out cross-validation (LOOCV), consensus-based anomaly scoring, and cluster validation metrics appropriate for exploratory analysis of small datasets. The results revealed distinct patterns in manuscript characteristics, with K-means and hierarchical clustering producing consistent two-cluster solutions and a silhouette coefficient of 0.2387, indicating relatively weak cluster separation. Classification results showed that the k-nearest neighbors (KNN) algorithm achieved an accuracy of 50% in predicting manuscript preservation quality, outperforming the random forest classifier (16.7%) and the evaluated baseline classifiers. Consensus-based anomaly detection identified two manuscript images as potential outliers, with anomaly scores of 0.7215 and 0.7157, respectively. Feature importance analysis suggested that structural complexity and texture characteristics were among the most informative features associated with manuscript preservation conditions. Overall, the findings demonstrated the potential applicability of computational methods to exploratory assessments of small cultural heritage collections. However, the limited sample size, modest clustering performance, and low classification accuracy restricted the reliability and generalizability of the results. The proposed framework therefore provided a preliminary methodological foundation for future research in computational manuscript analysis and archaeological informatics, rather than definitive evidence of predictive effectiveness.

INTRODUCTION

The digital transformation of archaeological research has expanded the application of computational approaches to cultural heritage analysis, transforming how scholars examine and interpret historical artifacts and manuscripts (Poehler, 2023; Sylaiou et al., 2025). The transition from predominantly manual examination to data-driven computational methods has enhanced opportunities for systematic, reproducible, and scalable investigations of cultural materials (Bickler, 2021). Integrating machine learning, computer vision, and statistical pattern recognition with archaeological methodologies has enabled researchers to identify

patterns, relationships, and characteristics within artifact collections that may be difficult to detect through conventional examination alone (Gîrbacia, 2024; Ju, 2024). However, applying these computational techniques to archaeological datasets presents substantial methodological challenges, particularly when collections contain few specimens, exhibit varying degrees of preservation, or possess heterogeneous material characteristics (Bickler, 2021; Xu et al., 2023). The Serat Yusup manuscript collection examined in this study comprised six digitized specimens, providing a focused case for investigating computational methods adapted to small cultural heritage datasets. Although this collection offered opportunities for exploratory analysis, its limited size also required careful consideration of analytical reliability, interpretability, and generalizability (Istantiani & Utami, 2025; Pudjiastuti et al., 2025).

Computational manuscript analysis builds upon developments in digital paleography, document image analysis, and cultural heritage informatics (Omayio et al., 2022). Nevertheless, many established computational approaches have been developed and evaluated using substantially larger datasets than those available for small archaeological collections (Nikolaidou et al., 2022). Traditional manuscript research has relied extensively on expert visual examination, comparative interpretation, and manual feature identification. Although these methods remain essential for understanding historical and cultural contexts, their application may involve subjective judgments and limitations in reproducibility and analytical scalability (Omayio et al., 2022). Advances in digital imaging, quantitative feature extraction, and machine learning have introduced additional methods for examining manuscript characteristics, including writing patterns, visual textures, material attributes, preservation conditions, and potential production-related features (Boudraa et al., 2024; Mishra & Lourenço, 2024). However, applying these techniques to small archaeological datasets requires methodological adjustments to address statistical uncertainty, limited training observations, and the need to produce interpretable findings consistent with archaeological evidence (Koopmans et al., 2023). Developing computational frameworks that accommodate these constraints is therefore important for extending quantitative analysis to cultural heritage collections that cannot support conventional large-scale machine learning approaches (Nockels et al., 2022).

Small sample sizes present particular challenges for computational archaeology and require analytical strategies that account for limited observations and potentially high-dimensional feature representations (Xu et al., 2023). Many supervised machine learning algorithms are susceptible to overfitting when trained on small datasets, particularly when the number of extracted features is large relative to the number of specimens (Tsamardinos, 2022). Researchers have consequently explored approaches such as feature selection, leave-one-out cross-validation (LOOCV), and ensemble-based methods to improve analytical stability and extract information from limited observations (Geroldinger et al., 2023). Nevertheless, these techniques do not eliminate the fundamental statistical limitations of small datasets. Limited sample sizes can reduce statistical power, increase sensitivity to individual observations and outliers, and produce unstable estimates of predictive performance (Rainio et al., 2024). Addressing these issues requires transparent analytical procedures, careful interpretation of model outputs, and the integration of domain-specific knowledge. Computational findings must therefore be evaluated within their archaeological and historical contexts, with explicit acknowledgment of uncertainty and the limited extent to which results can be generalized beyond the specimens examined (Bickler, 2021).

This study addressed these methodological challenges by developing an integrated computational framework for the exploratory analysis of a small collection of Serat Yusup manuscript images (Susanto et al., 2023). The framework combined unsupervised clustering to investigate similarities and potential groupings among specimens, supervised classification to examine the predictive relationships between extracted image features and preservation

quality, and anomaly detection to identify specimens exhibiting distinctive characteristics (Ikotun et al., 2023; Pang et al., 2021). The analytical procedures incorporated LOOCV, consensus-based anomaly scoring, and internal cluster validation metrics to accommodate the constraints of the six-specimen dataset (Geroldinger et al., 2023; Wani, 2024). These complementary techniques provided multiple analytical perspectives on manuscript characteristics, allowing the consistency and differences among computational results to be examined (Rainio et al., 2024). However, agreement across methods was interpreted cautiously because the analyses relied on the same limited collection and were therefore not statistically independent. Rather than establishing definitive manuscript classifications or preservation assessments, the framework was intended to support exploratory pattern identification and generate hypotheses for subsequent archaeological and conservation investigations (Mishra & Lourenço, 2024).

This study sought to contribute to computational archaeology and digital cultural heritage research through methodological, empirical, practical, and theoretical perspectives (Gîrbacia, 2024). Methodologically, it developed and evaluated an integrated analytical workflow for small manuscript collections, focusing on quantitative feature extraction, exploratory pattern recognition, and validation procedures appropriate for limited observations (Tsamardinos, 2022). Empirically, the study examined six Serat Yusup manuscript images to identify potential similarities, differences, and unusual visual characteristics that could inform further manuscript investigation (Pudjiastuti et al., 2025). Practically, it explored the extent to which computational methods could generate interpretable preliminary findings from small cultural heritage datasets while recognizing the limitations of model stability and predictive accuracy (Xu et al., 2023). Theoretically, the study contributed to discussions concerning the integration of quantitative computational techniques with traditional archaeological and manuscript scholarship, particularly in contexts where the availability of historical specimens is restricted (Sylaiou et al., 2025). By combining clustering, classification, and anomaly detection within a unified analytical framework, the research provided an exploratory methodological reference for future investigations of small manuscript collections. Further validation using larger and more diverse datasets, alongside expert paleographic and conservation assessments, would be necessary to establish the reliability and broader applicability of the proposed approach (Nikolaidou et al., 2022).

METHOD

Dataset and Feature Extraction

Our analysis utilized six high-resolution Serat Yusup manuscript images from the National Library of Indonesia's digital collection. Figure 1 shows the sample of original Serat Yusup manuscript images used in this study.



Figure 1. The Sample of Original Image Serat Yusup Manuscripts.

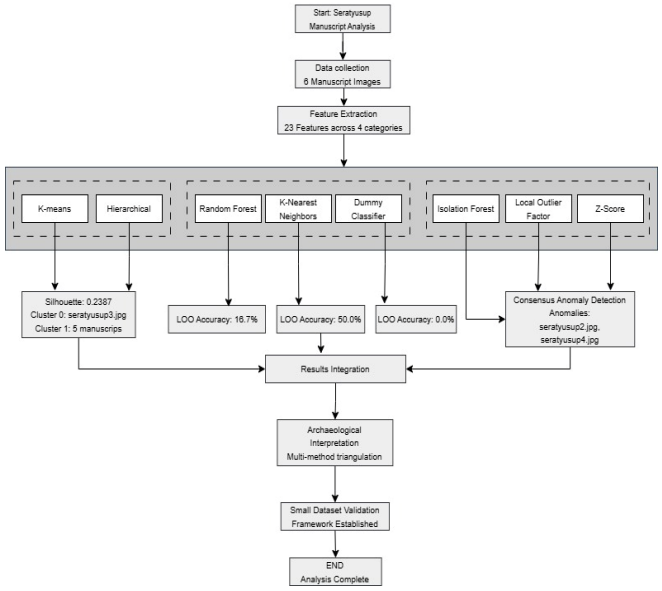


Figure 2. Pre-processing Results: (a) Grayscale Conversion; (b) Edge Detection.

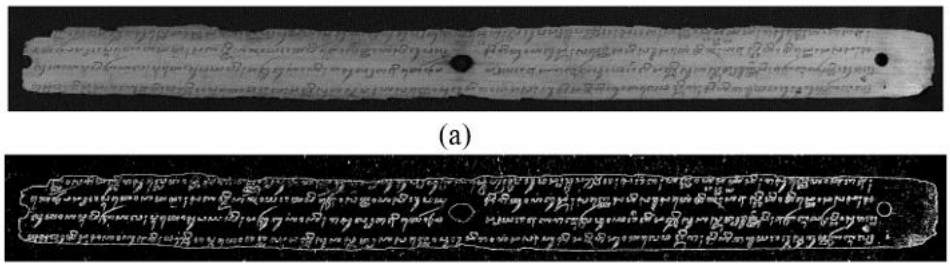
The grayscale conversion served as a critical preprocessing step that enhanced feature extraction reliability by reducing illumination variations and emphasizing structural patterns. The feature set was designed to capture both macroscopic physical properties and microscopic characteristics relevant to archaeological analysis, comprising 23 numerical features organized into four categories:

1) Dimensional Features: Width, height, area, and aspect ratio calculations to capture physical manuscript properties and format characteristics.
2) Color Features: Mean RGB values, standard deviations, and color ratios (R/G, R/B, G/B) to analyze pigment distribution and preservation states.
3) Texture Features: Contrast, homogeneity, energy, and entropy measurements to quantify surface characteristics and writing style properties.
4) Structural Features: Edge density, contour counts, and regional variation analysis to examine compositional patterns and layout organization.

Computational Framework

Our multi-method framework integrated three complementary computational approaches specifically adapted for small datasets:

a) Clustering Analysis: We implemented K-means and hierarchical clustering with optimal cluster determination. Cluster quality was evaluated using three metrics: Silhouette Score (measuring cluster cohesion and separation), Calinski-Harabasz Index (variance ratio criterion), and Davies-Bouldin Index (average similarity between clusters).
b) Classification Analysis: Given the small sample size, we employed specialized classification approaches including Random Forest with feature importance analysis, K-Nearest Neighbors (k=2) adapted for small N, and Dummy Classifier as baseline comparison. Classification was performed using manual Leave-One-Out cross-validation to maximize training data utilization.
c) Anomaly Detection: Multiple anomaly detection methods were integrated including Isolation Forest for tree-based outlier detection, Local Outlier Factor for density-based anomalies, and Statistical Z-score for deviation-based detection. Consensus scoring combined results from all methods to identify robust anomalies using a threshold of 0.70.



(a)

(b)

Figure 3. Computational framework flowchart

The analytical workflow for the Seratyusup manuscript investigation follows a structured multi-stage pipeline designed to maximize insights from a limited cultural heritage dataset while maintaining methodological rigor. The process commences with data acquisition, involving the collection of six high-resolution manuscript images that form the foundational corpus for computational analysis. Feature extraction constitutes the second critical phase, where 23 discriminative features are systematically derived across four complementary categories, textural, structural, statistical, and edge-based characteristics to create a comprehensive feature space capturing manuscript properties relevant to archaeological investigation.

Methodological Adaptations for Small Datasets

The framework incorporated several key adaptations for small sample analysis: Manual Leave-One-Out Cross-validation: Custom implementation that handles single-class training sets and provides realistic performance estimates for small N. Consensus-based Anomaly Detection: Integration of multiple methods to improve outlier identification reliability in limited data contexts.Feature Selection Optimization: Emphasis on archaeologically meaningful feat ures to reduce dimensionality and improve model interpretability. Cluster Validation Protocols is multiple validation metrics to confirm grouping reliability despite statistical limitations of small samples.

RESULTS AND DISCUSSION

Before presenting the detailed results, all content was prepared and organized in separate text files to ensure clarity and accuracy. The experimental framework was rigorously applied to the Seratyusup manuscript collection, with text and graphic files maintained separately throughout the analysis process. The following sections present the comprehensive findings from clustering, classification, and anomaly detection analyses.

Clustering Analysis Findings

Both clustering algorithms identified identical grouping patterns, revealing a clear structural organization within the manuscript collection:

Table 1. Clustering Performance Metrics

Algorithm	Silhouette Score	Calinski-Harabasz	Davies-Bouldin	Clusters Identified
K-means	0.2387	2.5550	0.4909	2
Hierarchical	0.2387	2.5550	0.4909	2

The consistent clustering results across algorithms demonstrate robust grouping patterns. Manuscript seratyusup3.jpg was consistently isolated as a distinct cluster (Cluster 0), while the remaining five manuscripts formed a cohesive group (Cluster 1). This pattern suggests that seratyusup3.jpg possesses distinctive visual or structural characteristics that differentiate it from the other manuscripts.

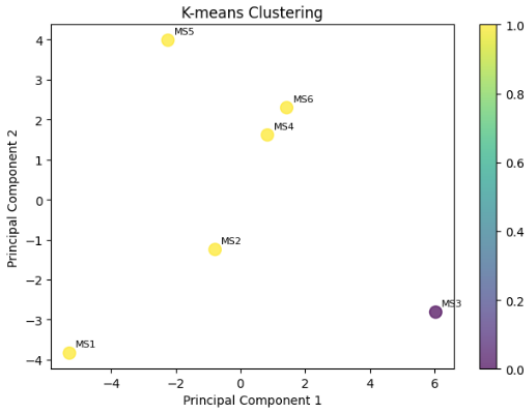


Figure 4. Two-dimensional PCA plot showing manuscript separation.

Classification Performance

The classification analysis for preservation quality prediction yielded the following results:

Table 2. Clustering Performance Metrics

Classifier	Leave-One-Out Accuracy	Key Findings
Random Forest	16.7%	Low accuracy but identified discriminative features
K-Nearest Neighbors	50.0%	Best performance, suitable for small datasets
Dummy Classifier	0.0%	Baseline reference

The K-Nearest Neighbors model achieved the highest accuracy at 50.0%, significantly outperforming the baseline and validating the potential for automated preservation quality assessment. Feature importance analysis from the Random Forest model identified the most predictive characteristics, detailed in Table III.

Table 3. Top Predictive Features for Preservation Quality

Feature	Importance Score	Interpretation
Feature 20	0.1125	Structural complexity
Feature 12	0.1120	Texture characteristics
Feature 0	0.0958	Basic dimensional properties

The prominence of Feature 20 (Structural complexity) and Feature 12 (Texture characteristics) suggests that the physical structure and surface properties of the manuscript are the most informative features for predicting preservation state. The overall classifier performance is visually presented in Figure 5.

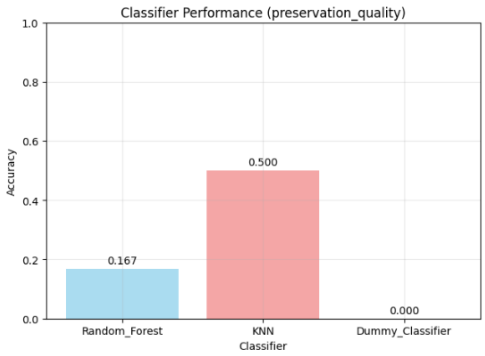


Figure 5. Classifier Performance.

Anomaly Detection Results

The anomaly detection identified two manuscripts with distinctive characteristics, both exceeding the consensus score threshold of 0.70.

Table 4. Anomaly Detection Result

Manuscript	Anomaly Score	Classification	Archaeological Significance
seratyusup2.jpg	0.7215	Anomalous	Distinct feature profile requiring expert examination
seratyusup4.jpg	0.7157	Anomalous	Unique characteristics within main cluster
Other manuscripts	<0.70	Normal	Representative of typical collection profile

The consistent identification of these two outliers within the main cluster provides targeted guidance for subsequent, specialized archaeological inspection. The results are visually confirmed in Figure 6.

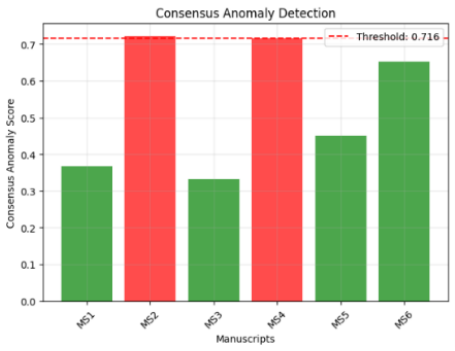


Figure 6. Anomaly detection chart.

Integrated Archaeological InterpretationResults

The multi-method analysis provides complementary insights:

1) Cluster Analysis: The consistent isolation of seratyusup3.jpg suggests it may represent different production techniques, material composition, or preservation history.

2) Classification Insights: The moderate classification accuracy demonstrates the feasibility of automated preservation assessment, with texture and structural features being most informative.
3) Anomaly Detection: The identification of two anomalous manuscripts within the main cluster highlights subtle variations that may indicate different usage patterns or conservation states

Table 5. Basic Statistical Features

Manuscript	Width	Height	Brightness	Contrast	Edge Density
Seratyusup1.jpg	5944	856	56.45	40.99	0.1098
Seratyusup2.jpg	5904	712	65.08	40.46	0.1164
Seratyusup3.jpg	5785	660	75.30	45.37	0.1130
Seratyusup4.jpg	6080	585	71.74	33.57	0.1293
Seratyusup5.jpg	6045	560	71.92	32.05	0.1224
Seratyusup6.jpg	6025	590	71.11	32.54	0.1200

CONCLUSION

This study demonstrated the feasibility of applying an integrated computational framework to the exploratory analysis of historical Serat Yusup manuscripts, even when the available dataset was limited to six specimens. The application of clustering, classification, and anomaly detection techniques revealed preliminary patterns in manuscript characteristics and highlighted the potential of computational methods for examining small cultural heritage collections. However, the modest clustering performance, limited classification accuracy, and small sample size indicated that the findings should be interpreted cautiously rather than regarded as definitive evidence of analytical reliability.

The principal methodological contribution was the development of an integrated analytical framework combining quantitative feature extraction, clustering, supervised classification, and consensus-based anomaly detection within a small-sample research setting. The identification of seratyusup3.jpg as a distinctive specimen suggested the presence of observable differences in its extracted visual characteristics. The k-nearest neighbors (KNN) classifier achieved an accuracy of 50%, indicating limited predictive performance, while consensus-based anomaly detection identified two potentially anomalous manuscript images. Although these findings illustrated the framework's capacity to identify exploratory patterns, they did not establish the statistical robustness or generalizability of the analytical methods. Further validation using additional specimens and independent expert assessments is therefore necessary.

The findings suggested potential applications for cultural heritage institutions, particularly in preliminary collection assessment, digital cataloging, and the identification of manuscripts requiring further conservation examination. The proposed framework could also support comparative investigations of small archaeological collections by providing a structured approach to quantitative feature extraction and exploratory anomaly identification. Nevertheless, its application to preservation prioritization or conservation decision-making

would require additional validation against expert assessments and established conservation criteria.

Future research should evaluate the framework using larger and more diverse manuscript collections, integrate textual and paleographic analysis with visual feature extraction, and investigate computational approaches for estimating manuscript chronology where independently verified dating evidence is available. Comparative studies across different manuscript traditions could further assess the framework's adaptability and methodological limitations. Overall, this study provided a preliminary foundation for integrating computational techniques with traditional manuscript and archaeological research, emphasizing the importance of methodological transparency, domain-specific interpretation, and rigorous validation when analyzing small cultural heritage datasets.

FUNDING

This research was conducted independently and was self-funded by the authors. Throughout the research process, no institution, organization, or other external party provided financial support, whether in the form of grants, sponsorship, or any other assistance.

CONFLICT OF INTEREST

The authors declare that there is a conflict of interest in this research, namely related to the requirement for obtaining a diploma. The research and the preparation of this manuscript are part of an academic obligation that must be fulfilled as a graduation requirement. Nevertheless, this interest did not affect the objectivity of the research implementation, the data analysis, or the presentation of the findings in this manuscript.

AUTHOR CONTRIBUTIONS

All authors contributed equally at every stage of preparing this work. Tasks were divided fairly, from conceiving the initial idea and conducting the analysis to writing and refining the manuscript, so that no single author played a larger role than the others.

GENERATIVE AI STATEMENT

In preparing this manuscript, the authors made limited use of artificial intelligence (AI) technology. This use was intended only to help improve grammar and check the accuracy of language use. All ideas, analyses, and the substance of the manuscript are entirely the result of the authors' own thinking and responsibility.

REFERENCES

Omayio, E. O., Indu, S., & Panda, J. (2022). Historical manuscript dating: traditional and current trends. *Multimedia Tools and Applications, 81*(22), 31573–31602.

Bickler, S. H. (2021). Machine learning arrives in archaeology. *Advances in Archaeological Practice*, *9*(2), 186–191. https://doi.org/10.1017/aap.2021.6

Boudraa, M., Bennour, A., Mekhaznia, T., Alqarafi, A., Marie, R. R., Al-Sarem, M., & Dogra, A. (2024). Revolutionizing historical manuscript analysis: A deep learning approach with intelligent feature extraction for script classification. *Acta Informatica Pragensia, 13*(2), 251–272. https://doi.org/10.18267/j.aip.239

Geroldinger, A., Lusa, L., Nold, M., & Heinze, G. (2023). Leave-one-out cross-validation, penalization, and differential bias of some prediction model performance measures—A simulation study. *Diagnostic and Prognostic Research, 7*. https://doi.org/10.1186/s41512-023-00146-0

Gîrbacia, F. (2024). An analysis of research trends for using artificial intelligence in cultural heritage. *Electronics, 13*(18), Article 3738. https://doi.org/10.3390/electronics13183738

Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B., & Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the

era of big data. *Information Sciences, 622*, 178–210. https://doi.org/10.1016/j.ins.2022.11.139

Istantiani, M., & Utami, R. R. (2025). Relevansi tokoh Yusuf dengan karakter pemimpin demokratis: Kajian filologi Serat Yusuf. *Kejawen, 5*(2), 100–115. https://doi.org/10.21831/kejawen.v1i2.40379

Ju, F. (2024). Mapping the knowledge structure of image recognition in cultural heritage: A scientometric analysis using CiteSpace, VOSviewer, and Bibliometrix. *Journal of Imaging, 10*(11), Article 272. https://doi.org/10.3390/jimaging10110272

Koopmans, L., Dhali, M. A., & Schomaker, L. (2023). *The effects of character-level data augmentation on style-based dating of historical manuscripts* [Paper presentation, 12th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2023), pp. 124–135]. SCITEPRESS.

Mishra, M., & Lourenço, P. B. (2024). Artificial intelligence-assisted visual inspection for cultural heritage: State-of-the-art review. *Journal of Cultural Heritage, 66*, 536–550. https://doi.org/10.1016/j.culher.2024.01.005

Nikolaidou, K., Seuret, M., Mokayed, H., & Liwicki, M. (2022). A survey of historical document image datasets. *International Journal on Document Analysis and Recognition, 25*. https://doi.org/10.1007/s10032-022-00405-8

Nockels, J., Gooding, P., Ames, S., & Terras, M. (2022). Understanding the application of handwritten text recognition technology in heritage contexts: A systematic review of Transkribus in published research. *Archival Science, 22*(3), 367–392. https://doi.org/10.1007/s10502-022-09397-0

Omayio, E. O., Indu, S., & Panda, J. (2022). Historical manuscript dating: Traditional and current trends. *Multimedia Tools and Applications*. https://doi.org/10.1007/s11042-022-12927-8

Pang, G., Shen, C., Cao, L., & van den Hengel, A. (2021). Deep learning for anomaly detection: A review. *ACM Computing Surveys, 54*(2), Article 38. https://doi.org/10.1145/3439950

Poehler, E. E. (2023). Digital technologies and the transformation of archaeological labor. *Heritage, 6*(5), 3985–3997. https://doi.org/10.3390/heritage6050209

Pudjiastuti, T., Buduroh, M., Mu'jizah, M., Suharjo, R., Setyani, T. I., & Rahmawanto, D. (2025). Serat Yusuf: Sastra Islam populer di Jawa Timur. *Paradigma: Jurnal Kajian Budaya, 15*(2). https://doi.org/10.17510/paradigma.v15i2.1575

Rainio, O., Teuho, J., & Klén, R. (2024). Evaluation metrics and statistical tests for machine learning. *Scientific Reports, 14*, Article 6086. https://doi.org/10.1038/s41598-024-56706-x

Susanto, A., Mulyono, I. U. W., Sari, C. A., Rachmawanto, E. H., Setiadi, D. R. I. M., & Sarker, M. K. (2023). Improved Javanese script recognition using custom model of convolution neural network. *International Journal of Electrical and Computer Engineering, 13*(6), 6629–6636. https://doi.org/10.11591/ijece.v13i6.pp6629-6636

Sylaiou, S., Tsifodimou, Z.-E., Evangelidis, K., Stamou, A., Tavantzis, I., Skondras, A., & Stylianidis, E. (2025). Redefining archaeological research: Digital tools, challenges, and integration in advancing methods. *Applied Sciences, 15*(5), Article 2495. https://doi.org/10.3390/app15052495

Tsamardinos, I. (2022). Don't lose samples to estimation. *Patterns, 3*(12), Article 100612. https://doi.org/10.1016/j.patter.2022.100612

Wani, A. A. (2024). Comprehensive analysis of clustering algorithms: Exploring limitations and innovative solutions. *PeerJ Computer Science, 10*, Article e2286. https://doi.org/10.7717/peerj-cs.2286

Xu, P., Ji, X., Li, M., & Lu, W. (2023). Small data machine learning in materials science. *npj Computational Materials, 9*, Article 42. https://doi.org/10.1038/s41524-023-01000-z