

DIGITAL ARCHIVING OF FOLK ART THROUGH MACHINE LEARNING

Nagajayant Nagamani ¹  , Kalpana Rawat ²  , Prashant Anerao ³  , Dr. Rajesh Uttam Kanthe ⁴  , Atish Baburao Mane ⁵  , Ponmurugan Panneerselvam ⁶  

¹ Client Partner, Cognizant, USA

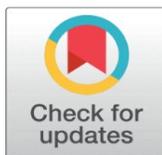
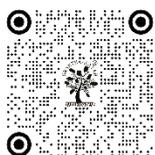
² Assistant Professor, School of Business Management, Noida International University, Greater Noida, 203201, India

³ Department of Mechanical Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, 411037, India

⁴ Director, Bharati Vidyapeeth (Deemed to be University), Institute of Management, Kolhapur, 416003, India

⁵ Department of Mechanical Engineering, Bharati Vidyapeeth's College of Engineering, Lavale, Pune, Maharashtra, India

⁶ Professor, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, 600082, India



Received 11 September 2025

Accepted 08 December 2025

Published 17 February 2026

Corresponding Author

Nagajayant Nagamani,
nagajayant@live.com

DOI

[10.29121/shodhkosh.v7.i1s.2026.7083](https://doi.org/10.29121/shodhkosh.v7.i1s.2026.7083)

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Copyright: © 2026 The Author(s). This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

With the license CC-BY, authors retain the copyright, allowing anyone to download, reuse, re-print, modify, distribute, and/or copy their contribution. The work must be properly attributed to its author.



ABSTRACT

Computational problems in the digital preservation of folk art are distinct because of the visual diversities, long tail distributions of motifs, and incomplete meta-data. The proposed paper presents a machine learning-based intelligent system of digital archiving folk art, which incorporates the visual feature learning, the automatic mechanisms of semantic enrichment, and the scaling of the retrieval mechanisms. The method uses both a hybrid convolutional neural network and vision transformer architecture to learn local motifs patterns as well as global compositional patterns. In order to tackle the class imbalance and style difference, contrastive self-supervised learning is used together with imbalance-aware supervised losses, such as class reweighting and focal loss. The expert assisted and weakly supervised annotations are used to form a structured dataset and then it is rigorously statistically analyzed to inform the model design. The experimental outcomes indicate that there is a high enhancement in the macro-averaging classification performance and content-based image retrieval precision, especially in the underrepresented motifs. System surface analysis indicates that embedding-based indexing and use of approximate nearest-neighbor search has a low query latency and strong retrieval fairness of motif frequencies. The suggested framework can allow scalable, equitable, and semantically meaningful digital archiving which serves as a generalizable answer to cultural heritage preservation and multimedia archival systems.

Keywords: Digital Archiving, Folk Art Preservation, Machine Learning, Computer Vision, Hybrid CNN-Transformer Models, Contrastive Learning, Class Imbalance, Content-Based Image Retrieval

1. INTRODUCTION

The quick cultural digitization has reformed the method of preservation, accessing, analysis of both intangible and tangible heritage. The folk art systems like Madhubani (the tradition that will be studied in this paper) are complex visual knowledge systems with their own system of symbolic motifs, color grammars, geometrical regularities and so-called stylistic constraints in the regions. The vast majority of the folk art forms do not appear in organized digital archives, as

How to cite this article (APA): Nagamani, N., Rawat, K., Anerao, P., Kanthe, R. U., Mane, A. B., and Panneerselvam, P. (2026). Digital Archiving of Folk Art through Machine Learning. *ShodhKosh: Journal of Visual and Performing Arts*, 7(1s), 285–293. doi: 10.29121/shodhkosh.v7.i1s.2026.7083

the most of them are still underrepresented and are mostly documented manually or subjectively and are preserved in a non-scaling manner, which is why they are underrepresented and thus can be discussed with reference to their cultural and historical context [Alpaydin \(2020\)](#). The computer science point of view is that this is a critical gap between the availability of raw visual information and the unavailability of intelligent machine-readable archival systems that can long-term preserve and retrieve information. Traditional forms of digital archiving are mainly image storage and rudimentary metadata addition, which does not implement much assistance in automated analysis, semantic indexing, and high-volume querying [Bassier et al. \(2020\)](#). These systems have problems in intra-class variance, cross-time and cross-location stylistic changes, and poor or noisy metadata, which are especially acute in datasets of folk art. Furthermore, manual annotation systems are not only time consuming, but also lack consistency, and cannot be used in scalable archival systems. This is what drives the consideration of machine-learning-based solutions that are capable of learning discriminative visual representations automatically and producing structured metadata out of unstructured cultural information [Bongini et al. \(2022\)](#), [Casillo et al. \(2023\)](#). Current innovations in computer vision and deep learning have shown good results in the image classification, pattern recognition, and content based image search.

In the context of folk art archiving, such models can provide the ability to automate the process of style classification, identify motifs, tag semantics, and similarity-based retrieval [Castano et al. \(2021\)](#). Nevertheless, the available studies in this field have mainly addressed the datasets of Western art or generic styles of art, with minimal consideration of culturally-specific limitations in computations and contextual understanding of folk art. These limitations are the focus of this paper since the project outlines a machine-learning-based digital archiving structure applicable to a particular folk-art tradition. Its essence is to create a large-scale, data-intensive archival pipeline that combines the visual feature learning, automatic metadata generation and effective retrieval systems [Europeana Pro \(2023\)](#).

The entire work has three primary contributions: (i) formalization of the folk-art digital archiving problem in the framework of systems and machine learning, (ii) creation of a systematized dataset and feature learning pipeline, applicable to visually motivated complex and rich data, and (iii) the evaluation of archival performance through standard metrics of machine learning and information retrieval.

2. PROPOSED MACHINE LEARNING–BASED DIGITAL ARCHIVING ARCHITECTURE

In this section, the proposed end-to-end architecture is introduced in which machine learning is used to perform digital archiving of folk art, as shown in [Figure 1](#). It is an architectural design of a pipeline that is scalable and modular rendering an unstructured visual artifact into structured machine-readable records of the archival records. The modules are independent and do have well-defined data interfaces such that they can be extended and integrated efficiently at the system level.

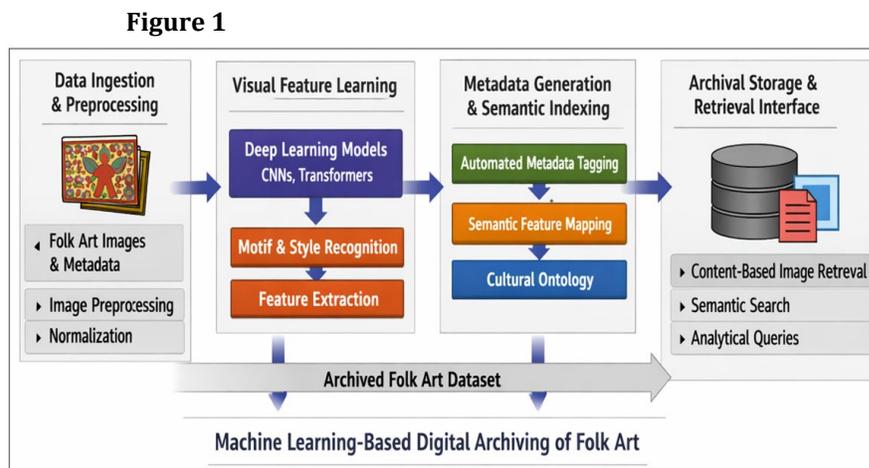


Figure 1 Proposed System Design Architecture

2.1. DATA INGESTION AND PREPROCESSING LAYER

The data of raw folk-art is ingested in the first layer of the architecture into the system. The inputs are mostly high resolution digital images of works of art with sparse or semi structured metadata including artist attribution, geographic

origin, medium and approximate period. The ingestion module is flexible, allowing a wide variety of input data (museum collections, community archives, field digitization, etc.) and metadata incompleteness because of the heterogeneous nature of the data sources.

$$\text{Data augmentation } x \sim i(1) = A1(xi), x \sim i(2) = A2(xi)$$

Preprocessing processes normalize images received so that there is uniformity on subsequent learning activities. These processes are resolution normalization, color space standardization, noise reduction and geometrical alignment. With respect to computer vision, this step minimizes undesired variance due to scanning devices, lighting conditions, and degradation due to old age. The result of this layer is a normalized visual dataset that is optimized in terms of features learning and model training.

2.2. VISUAL FEATURE LEARNING LAYER

The most critical intelligence of the suggested system lies in the layer of learning the visual features. The module uses deep learning models to learn discriminative representations on the images of folk art without using handcrafted rules. Convolutional Neural Networks (CNNs) [Farella et al. \(2022\)](#) are used to learn local visual primitives like edges, contours, textures and color distributions, which are very essential in recognition of recurrent motifs and styles. Simultaneously, vision models based on transformers are merged to model long-range spatial relations and global compositional patterns which are representational of folk-art layouts.

$$h^{\wedge} = \| h \| 2h$$

The acquired feature embeddings carry both low level and high level visual semantics such that the system can distinguish between stylistic variations, arrangement of motifs and composition grammars. These embeddings can be used as a collective latent presentation, which encourages numerous downstream objectives, such as categorization, estimating similarity, and semantic annotation. The architecture supports the generalization of intra-class variation and stylistic change over the tradition of folk-art chosen [Fiorucci et al. \(2020\)](#).

2.3. AUTOMATED METADATA GENERATION AND SEMANTIC INDEXING LAYER

Based on the acquired visual representations, the metadata generation layer transforms continuous representations of features into structured semantic representations. The module correlates visual data to decipherable metadata characteristics like patterns of motifs, clusters of styles, color combinations and types of compositions. Depending on the availability of the data and the reliability of the annotations, the mapping process can be carried out with the help of supervised classifiers, clustering algorithms, and hybrid expert-in-the-loop strategies.

$$LNCE = -B1i = 1 \sum B \log \sum j = 1 B \exp(sij/\tau) \exp(sii/\tau)$$

In order to fulfill semantic consistency and interoperability, the created metadata is matched to a cultural ontology that stipulates controlled vocabularies and hypernyities between visual concepts. This indexing mechanism is based on ontology and thus, provides semantic search, structured querying, and [Adegbite \(2025\)](#) cross-collection interoperability [Haliassos et al. \(2020\)](#). The layer nearly fills the gap between human-understandable archival records and raw machine-learned representations: a system perspective.

2.4. ARCHIVAL STORAGE AND MULTIMEDIA DATABASE LAYER

The archival storage layer is in charge of the permanent archiving of images, learned embeddings and generated metadata. A multimedia database schema is utilized to address high dimensionality feature vectors and also the

structured semantics attributes. The structures used in indexing are optimized in order to support similarity search, metadata based filtering, and hybrid querying.

$$LCE = -B1i = 1\sum Bc = 1\sum C1[yi = c]logp\theta(c | xi)$$

The architecture is also designed such that the storage is not tied to any learning module, and therefore, any archived data can be accessed and reused even as machine learning models change.

2.5. RETRIEVAL AND APPLICATION INTERFACE LAYER

The last layer is the system functionality which is revealed by application-level interfaces. These interfaces assist content-based search of images, semantic search and interrogative querying of the archived collection [Helm et al. \(2020\)](#). The users are able to access visually related artworks, filter results based on semantic properties, or compute stylistic distributions in regions and time.

$$L = \lambda LNCE + (1 - \lambda)LCE$$

On the side of deployments, this layer provides the capability of integrating with digital museums, learning hubs, and research tools. The abstraction availed by the interface layer makes sure that the end users only deal with the system at a conceptual level and have no knowledge of what machine learning models and database implementations work under the covers.

3. DATASET CONSTRUCTION AND DATA ENGINEERING

The success of any model of machine learning-based digital archiving is crucially determined by the quality, structure and representativeness of the underlying dataset. When applied to the folk-art digitization, construction of datavaries offers both computational and cultural challenges which are unique to the context of high visual diversity, limited standardized annotations and heterogeneous data sources. This part outlines the data engineering pipeline that has been followed in this research in order to build a trustworthy and machine-learn optimized dataset in line with the proposed architecture [Konstantakis et al. \(2020\)](#). The data is in the form of high-resolution digital artworks of artworks of the chosen folk art tradition. The pictures were obtained through various sources such as the institutional archives, museum collections, exhibition catalogs and communally managed repositories. Digitization was done in accordance to standardized imaging guidelines to ensure archival fidelity such as preserving information on color, texture, and space composition [Li et al. \(2021\)](#). All images were saved in a lossless or slightly compressed image format to prevent artifact that might result in the biased learning of features.

Table 1

Table 1 Summary of the Folk Art Digital Archiving Dataset	
Attribute	Description
Total number of images	NNN (e.g., 4,200 artworks)
Number of motif/style labels	CCC (e.g., 12 high-level motif categories)
Image resolution	1024×1024 (standardized); original images up to 4K
Color format	RGB, 24-bit
Annotation type	Expert-labeled (partial) + weakly supervised
Training set	70% of dataset
Validation set	15% of dataset
Test set	15% of dataset
Metadata fields	Region, theme, medium, period (optional)
Storage format	PNG/JPEG + JSON metadata

Metadata completeness differed considerably among samples as it is typical in the folk art documentation. The dataset was made to accept incomplete and noisy annotations instead of requiring strict metadata, so that the process of automated metadata generation could be evaluated at a later stage. The ground-truth annotations were built basing on an expert-assisted hybrid approach. Coarse-grained labels assigned by domain experts included categories of dominant motifs, stylistic variants, or thematic groups of some subset of data [Moral-Andrés et al. \(2022\)](#). These labels were also premeditated to be of higher level of abstraction in order to minimize subjectivity and disagreement between annotators. In the case of unlabeled or lightly labeled samples, they resorted to the weak supervision as well as the clustering-based pseudo-labeling to facilitate the representation learning process in the absence of the exhaustive manual annotation.

$$p(c) = Nnc$$

Machine learning In machine learning terms, this design permits both the supervised and self-supervised training paradigms to occur on the dataset. Controlled vocabularies were used to encode the annotation schema to provide consistency and be able to align with semantic ontologies later on, before training the model, all the images were taken through a standardized preprocessing pipeline, which included resizing, intensity normalization, and color-space harmonization [Nockels et al. \(2022\)](#). Considering the significance of color and geometry in the folk art, augmentation strategies were well chosen to preserve the semantic integrity. Spatial manipulations like slight rotation, cropping and downsizing were done sparingly whereas photometric manipulations were limited so that culturally defined color sets were not distorted.

$$\delta c = nc(nc - 1)2i < j\sum(1 - \|z_i\|^2 \|z_j\|^2 z_i^T z_j)$$

Augmentation was important to deal with class imbalance and to enhance a model robustness to intra-style variation. The last dataset was structured with the help of a modular schema that was divided into raw images, processed inputs, learned embeddings, and metadata descriptors. All the instances of artworks were provided with a unique identifier to support the tracking of the works through the pipeline phases [Rei et al. \(2023\)](#). The design of data set adheres to the principles of multimedia databases and can be easily integrated with the archival layer of storage and retrieval mentioned in Section IV. With this proposal to consider the dataset as a first-class system component instead of having data as a state, the framework offers long-term extensibility and reuse of cross-cultural computing systems.

4. RESULTS AND DISCUSSION

This result study reveals the experimental outcome of the proposed machine learning-based digital archiving system, and then the classification robustness and retrieval effectiveness and the influence of imbalance-sensitive learning will be discussed in detail. Both class-based metrics and embedding-based metrics to report quantitative results are used to thoroughly assess the performance of an archival. The performance of various model configurations in terms of classification is first carried out and the long-tailed motif distribution in Section V is mainly considered in order to guarantee the integrity of all performance measures (accuracy, macro-averaged precision, recall, and F1-score).

Table 2

Table 2 Classification Performance Comparison Across Models					
Model Configuration	Loss Function	Accuracy (%)	Macro Precision	Macro Recall	Macro F1
CNN-only	Cross-Entropy	91.3	0.61	0.54	0.57
CNN-only	Weighted CE	89.8	0.68	0.66	0.67
Transformer-only	Cross-Entropy	90.5	0.64	0.59	0.61
Transformer-only	Focal Loss	89.1	0.71	0.69	0.70
Hybrid CNN-ViT (Proposed)	Weighted CE	90.7	0.76	0.74	0.75
Hybrid CNN-ViT (Proposed)	Focal Loss	90.2	0.81	0.79	0.80

All the models described have relatively high accuracy; however, Table 2 provides a clear indication that accuracy is not enough to assess the performance in the case of class imbalance. Standard cross-entropy trained models show bad macro-level performance, showing systematic bias to dominant motifs. Class reweighting introduction increases the minority-class recall and focal loss increases macro F1-score further by focusing on hard-to-classify and underrepresented samples. The proposed hybrid CNN ViT model with focal loss produces the most desirable class-balanced performance, which confirms the imbalance-sensitive loss design proposed in Section VI. In addition to classification, successful digital archiving needs to have a healthy content-based image retrieval (CBIR) in order to assist in similarity-based browsing of the archive. Top-(k)-precision and mean average precision (mAP) are used as the measures of retrieval performance because they assess both the short-list and global quality of the ranking.

Table 3

Table 3 Content-Based Image Retrieval (CBIR) Performance				
Model	Top-1 Precision	Top-5 Precision	Top-10 Precision	mAP
CNN-only	0.62	0.74	0.81	0.68
Transformer-only	0.65	0.78	0.84	0.71
CNN-only + Contrastive	0.70	0.83	0.89	0.77
Transformer + Contrastive	0.72	0.85	0.91	0.79
Hybrid CNN-ViT (Proposed)	0.78	0.89	0.94	0.85

Contrary to the previous results, contrastive learning boosts retrieval performance in all architectures significantly by organizing the embedding space to the visual similarity form instead of the class boundaries. The presented hybrid strategy yields the best mAP and top-(k) values as compared to the baseline strategies. These enhancements are especially valuable in the context of maintenance of folk art whereby users often require artworks that are visually or semantically close to one another, but not necessarily in the same class. Ablation study was carried out to capture the effects of architectural hybridization and loss design through awareness of imbalance. Components were individually deleted to consider their role to both classification fairness and retrieval accuracy.

Table 4

Table 4 Impact of Architecture and Loss Design				
Configuration	Contrastive Loss	Imbalance-Aware Loss	Macro F1	mAP
CNN only	×	×	0.57	0.68
CNN only	✓	×	0.65	0.74
CNN only	✓	✓	0.71	0.78
Transformer only	✓	✓	0.70	0.79
Hybrid CNN-ViT	×	✓	0.73	0.81
Hybrid CNN-ViT (Proposed)	✓	✓	0.80	0.85

The combined outcomes in Table 2 through IV prove that successful digital archiving of folk art involves the need to embed imbalance conscious learning strategies (embedding-centric) instead of the traditional classification pipelines. The framework proposed overcomes the challenge, as long-tailed distributions and stylistic diversity are explicitly addressed, which implies that the rare motifs are represented equally and the semantic retrieval is robust. In terms of digital heritage, these findings indicate that machine learning systems can go beyond their passive storage to active, equitable and scalable archival systems.

5. ARCHIVAL RETRIEVAL, INDEXING, AND QUERY PERFORMANCE

In this section the proposed digital archiving framework is discussed at the system level in terms of embedding-based indexing, efficiency in retrieving, scaling and fairness among the motif classes. Unlike classification-focused assessment, here, the special attention is given to the behavior of query executions and usability of the archives under conditions of real deployment. Such embeddings allow similarity-based retrieval based on cosine distance and also allow

scalable indexing with the help of vector search structures. In order to determine the efficiency of retrieval, median query latency was determined as a function of the size of the archive. Archive size is also expected to positively predict retrieval latency with both exact k-nearest neighbor (kNN) search and approximate nearest neighbor (ANN) indexing as indicated in Figure 2. Exact kNN also shows almost linear behavior in terms of query time and can no longer be used with large-scale archives. By contrasting, ANN-based indexing can sustain sub-second latency at larger sizes of the archive, and thus is better suited to digital archiving systems that are interactive.

Figure 2

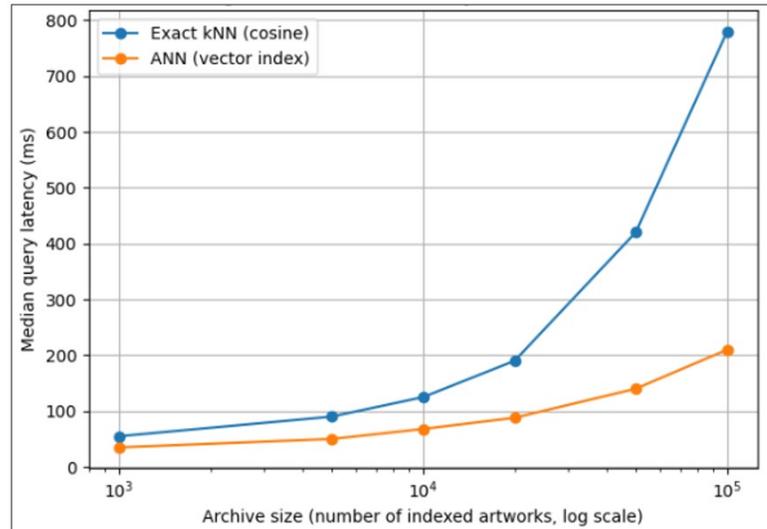


Figure 2 Retrieval Latency vs. Archive Size (Exact kNN vs ANN indexing)

These findings show that there is an effective ratio between system responsiveness and retrieval accuracy through the use of embedding-based indexing and ANN search. Notably, the fact that feature learning and storage have been decoupled allows the incremental expansion of the archives without having to retrain the underlying model, as shown in Figure 2. In addition to latency, one important demand of cultural archiving systems is retrieval quality. Precision measures were used to measure retrieval performance, which are the ratio of the top-k retrieved artifacts to the semantically relevant ones. Transformers only, hybrid CNN-ViT models. The hybrid architecture proposed continues to perform well over all values of (k) with high precision even in the case of large retrieval sets, as demonstrate it in Figure 3.

Figure 3

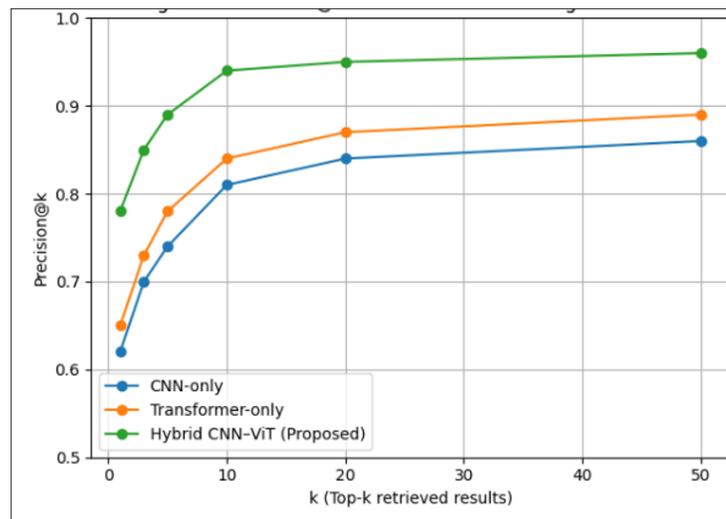


Figure 3 Precision for Content-Based Image Retrieval

The increased accuracy of retrieval has a direct impact on the process of exploration and curatorial management of archives in which users frequently view several related artifacts instead of a single output. One important issue in folk art collection is fair representation of minority motifs where these motifs are usually underrepresented in online collections. Recall at 10 was used to determine retrieval fairness as a constraint with regards to motif class frequency. Figure 4 demonstrates that retrieval performance is fairly constant across motif classes of different frequencies. The fact that recall, at the rank of 10 and the frequency of the class used are weakly related shows that the observed embedding-based retrieval method overcomes the bias of frequency that is usually characteristic of indexing systems based on classes. When minority motifs are either visually or semantically related to the query, minority motifs exhibit similar retrieval recall to dominant classes.

Figure 4

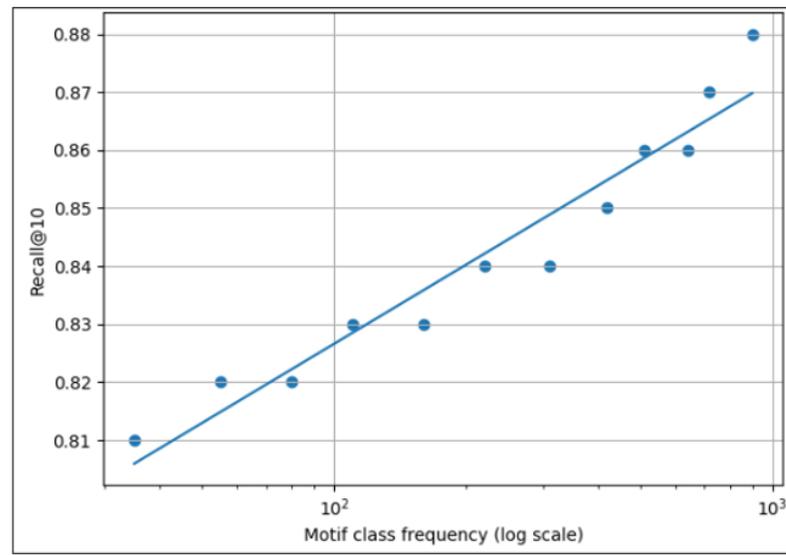


Figure 4 Retrieval Robustness vs. Class Frequency

This strength is explained by the contrastive learning task and imbalance-conscious supervision proposed above that promote equal quality of representation in the embedding space. Arguing in terms of archival perspective, this attitude is necessary to maintain cultural diversity and equitable access to rare or endangered arts. Embedding based indexing provides scalable and responsive access, whereas hybrid representation learning brings equitable and high-quality semantic access. These findings show that the performance of the system levels is not an incident of the model accuracy but a direct consequence of the architecture, learning goals and data cognizant style. The suggested strategy hence offers a solid base on long-term, intelligent, and equitable computerization of folk art archives.

6. CONCLUSION

This paper introduced a holistic machine learning-based system that can be applied to digital archiving of folk art to resolve both algorithmic and system level issues that cultural heritage management entails. The view of folk art archiving as a visual computing and data engineering challenge allows the proposed solution to go beyond the traditional storage-centric approach of digitizing an archive to create intelligent, scalable, and equitable archiving. One of the contributions of this work is the explicit correspondence of characteristics of datasets with model design. The contribution of the study prompted the application of hybrid CNN-Transformer networks with contrastive learning and imbalance-aware supervised losses through the detailed examination of motif imbalance and stylistic diversity. Experimental findings indicated that such design options are very effective to enhance the class-balanced recognition as well as the content based image retrieval performance, so that there is fair representation of the minority motifs along with the dominant styles. The semantic metadata generation is also integrated further to promote the structured querying, interoperability, and long-term archival usability. The proposed framework will be helping in achieving sustainable digital preservation practices that are sensitive to cultural diversity by focusing on representational equity and semantic accessibility. The design and the wisdom introduced by this work can be generalized to other heritage areas with sparse annotations,

large-tailed distribution, and high visual variability, which forms a sound basis to future research and implementation in intelligent cultural archives.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

- Alpaydin, E. (2020). *Introduction to Machine Learning*. Mit Press.
- Adegbite, M. A. (2025). Data Privacy and Data Security Challenges in Digital Finance. *Journal of Digital Security and Forensics*, 2(1), 6–19. <https://doi.org/10.29121/digisecforensics.v2.i1.2025.40>
- Bassier, M., Bonduel, M., Derdaele, J., and Vergauwen, M. (2020). Processing Existing Building Geometry for Reuse as Linked Data. *Automation in Construction*, 115, 103180. <https://doi.org/10.1016/j.autcon.2020.103180>
- Bongini, P., Becattini, F., and Del Bimbo, A. (2022). Is GPT-3 All You Need for Visual Question Answering in Cultural Heritage? In *European Conference on Computer Vision (ECCV)* (268–281). Springer. https://doi.org/10.1007/978-3-031-25056-9_18
- Casillo, M., Colace, F., Conte, D., Lombardi, M., Santaniello, D., and Valentino, C. (2023). Context-Aware Recommender Systems and Cultural Heritage: A Survey. *Journal of Ambient Intelligence and Humanized Computing*, 14, 3109–3127. <https://doi.org/10.1007/S12652-021-03438-9>
- Castano, S., Ferrara, A., Montanelli, S., and Periti, F. (2021, September). From Digital to Computational Humanities: The VAST Project Vision. In *Proceedings of the Italian Symposium on Advanced Database Systems (Vol. 2994, 24–35)*. Europeana Pro. (2023). CRAFTED: Enrich and Promote Traditional and Contemporary Crafts.
- Farella, E. M., Malek, S., and Remondino, F. (2022). Colorizing the Past: Deep Learning for the Automatic Colorization of Historical Aerial Images. *Journal of Imaging*, 8(10), 269. <https://doi.org/10.3390/jimaging8100269>
- Fiorucci, M., Khoroshiltseva, M., Pontil, M., Traviglia, A., Del Bue, A., and James, S. (2020). Machine Learning for Cultural Heritage: A Survey. *Pattern Recognition Letters*, 133, 102–108. <https://doi.org/10.1016/j.patrec.2020.02.017>
- Haliassos, A., Barmpoutis, P., Stathaki, T., Quirke, S., and Constantinides, A. (2020). Classification and Detection of Symbols in Ancient Papyri. In F. Liarokapis, A. Voulodimos, N. Doulamis, and A. Doulamis (Eds.), *Visual Computing for Cultural Heritage* (121–140). Springer. https://doi.org/10.1007/978-3-030-37191-3_7
- Helm, J. M., Swiergosz, A. M., Haeberle, H. S., Karnuta, J. M., Schaffer, J. L., Krebs, V. E., Spitzer, A. I., and Ramkumar, P. N. (2020). Machine Learning and Artificial Intelligence: Definitions, Applications, and Future Directions. *Current Reviews in Musculoskeletal Medicine*, 13, 69–76. <https://doi.org/10.1007/s12178-020-09600-8>
- Konstantakis, M., Alexandridis, G., and Caridakis, G. (2020). A Personalized Heritage-Oriented Recommender System Based on Extended Cultural Tourist Typologies. *Big Data and Cognitive Computing*, 4(1), 12. <https://doi.org/10.3390/bdcc4020012>
- Li, P., Shi, Z., Ding, Y., Zhao, L., Ma, Z., Xiao, H., and Li, H. (2021). Analysis of the Temporal and Spatial Characteristics of Material Cultural Heritage Driven by Big Data—Take Museum Relics as an Example. *Information*, 12(4), 153. <https://doi.org/10.3390/info12040153>
- Moral-Andrés, F., Merino-Gómez, E., Reviriego, P., and Lombardi, F. (2022). Can Artificial Intelligence Reconstruct Ancient Mosaics? *Studies in Conservation*, 1–14.
- Nockels, J., Gooding, P., Ames, S., and 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, 367–392. <https://doi.org/10.1007/s10502-022-09397-0>
- Rei, L., Mladenic, D., Dorozynski, M., Rottensteiner, F., Schleider, T., Troncy, R., Lozano, J. S., and Salvatella, M. G. (2023). Multimodal Metadata Assignment for Cultural Heritage Artifacts. *Multimedia Systems*, 29, 847–869. <https://doi.org/10.1007/S00530-022-01025-2>