Contents lists available at GrowingScience

International Journal of Data and Network Science

homepage: www.GrowingScience.com/ijds

Systematic literature review on optimization and exploration of retrieval methods digital image of ancient manuscript as an attempt conservation of cultural heritage

Ino Suryana[a,c], Diah Chaerani[b*], Khoirunnisa Rohadatul Aisy Muslihin[b], Athaya Zahrani Irmansyah[b], Setiawan Hadi[c], Anton Satria Prabuwono[d]

[a]*Former Student at Doctoral Mathematics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Indonesia*
[b]*Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Indonesia*
[c]*Department of Computer Sciences, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Indonesia*
[d]*Faculty of Computing and Information Technology Rabigh King Abdulaziz University, Saudi Arabia*

C H R O N I C L E

Article history:
Received: July 6, 2023
Received in revised format: August 4, 2023
Accepted: September 10, 2023
Available online: September 10, 2023

Keywords:
Retrieval method
Digital image
Optimization
Ancient manuscript
Cultural heritage

A B S T R A C T

Digitization technology has developed in the preservation of cultural heritage, especially ancient manuscripts. In this context, image retrieval methods allow for efficient access to the information contained in ancient manuscripts. Optimization techniques play a role in building an effective image retrieval method. This paper presents a systematic literature review that focuses on the role of optimization in digital image retrieval methods for preserving ancient manuscripts as part of cultural heritage preservation. This paper is organized based on the following research questions: (1) What are the research objectives regarding optimization in digital image retrieval of ancient manuscripts?; (2) What is the role of optimization in image retrieval methods for the preservation of ancient manuscripts?; and (3) How have existing studies determined the formulation of novel and new research? This study involved searching articles through Scopus, Dimensions, Science Direct, and Google Scholar databases and using bibliometric analysis to visualize research themes and trends. The results of the literature review show that research related to the role of optimization in the digital image retrieval method of ancient manuscripts is still open and this paper can be used as a reference for further research on this topic.

1. Introduction

Cultural heritage is an integral part of human civilization's history. It can be divided into two categories, tangible cultural heritage, which includes buildings, museums, monuments, manuscripts, books, and artifacts, as well as intangible cultural heritage, such as folklore, language, knowledge, and traditions (Foster & Kreinin, 2020; López et al., 2018; Skublewska-Paszkowska et al., 2022). Over time, much of this cultural heritage has suffered damage, especially tangible cultural heritage like ancient scripts. Therefore, technology has been developed to preserve cultural heritage. For example, 3D modeling is used to document various cultural heritage sites (Condorelli & Morena, 2023; Ishar et al., 2022; Remondino & Rizzi, 2010). The development of digitalization technology has also progressed in the preservation of ancient manuscripts. Converting ancient manuscripts into digital format has proven to be a potent tool in preserving critical information contained in the manuscripts (Jayanthi et al., 2017). Moreover, digital preservation can be utilized to manage analog information through the digitization of manuscripts (Ahmad & Sharma, 2020). Image retrieval methods can be used to retrieve or search for documents that have been digitized. By using these methods, visually similar images from a dataset can be extracted, which are similar to the queried image (Bagasi & Elrefaei, 2018). This opens up greater access for researchers to obtain in-formation contained

* Corresponding author.
E-mail address: d.chaerani@unpad.ac.id (D. Chaerani)

ISSN 2561-8156 (Online) - ISSN 2561-8148 (Print)

doi: 10.5267/j.ijdns.2023.9.010

in ancient manuscripts, while also revealing hidden details within the manuscripts (Sahoo & Mohanty, 2015; Tonazzini et al., 2019). Thus, digital technology offers significant opportunities for preserving and enhancing the understanding of the cultural heritage embodied in ancient manuscripts.

The challenges of using digital technology for the preservation of ancient manuscripts involve several technical complexities, including issues with restoration accuracy and color reproduction. Optimization plays a significant role in addressing these challenges. For instance, Shukla et al. (2015) employed the Swarm Optimization Richardson-Lucy algorithm for restoring images damaged by paper-and-salt noise, and Valova et al. (2006) studied optimization in image retrieval using color features. Therefore, the application of optimization can enhance classification and retrieval results (Mohammed et al., 2015).

In this paper, a systematic literature review (SLR) is conducted to obtain a discussion of existing studies and trends related to the use of optimization in the digital image retrieval of ancient manuscripts for the purpose of cultural heritage preservation. The aim of this literature review was to obtain an objective and comprehensive summary, as well as critical analysis through the identification, evaluation, and interpretation of relevant previous research on the studied topic (Firdaniza et al., 2021; Kitchenham, 2004). In this paper, the bibliometric analysis software is employed, which will be further explained in Section 3. To support this, the following are the research questions (RQ) in this SLR:

1. What are the research objectives regarding optimization in digital image retrieval of ancient manuscripts?
2. What is the role of optimization in image retrieval methods for the preservation of ancient manuscripts?
3. How have existing studies determined the formulation of novel and new research?

The organization of this paper is as follows. Section 2 discusses related literature review papers that were published previously. In Section 3, the methods used are presented, particularly in how to collect the articles for analysis. Section 4 presents the results of the bibliometric analysis and literature review related to the topic raised. The discussion is provided in Section 5. Finally, the conclusion is presented in Section 6.

2. Relevant Literature Review

In the recent year, some papers performed literature reviews regarding the use of digitization in preserving cultural heritage in general. Rodrigues et al. (2023) discussed digitization in the field of tourism, focusing on literature search related to SDGs concerning digitalization in the tourism industry on the Scopus database. Another article that addressed a similar topic is Madzík et al. (2023), which studied digital transformation. The two papers mentioned that the topic of cultural heritage is a reasonably consistent topic but has experienced a decline in recent years. Unfortunately, both studies did not specifically address the topic of digital image for ancient manuscripts.

Furthermore, Cotella (2023) discussed a literature review on the topic of application of artificial intelligence in the preservation of cultural heritage. This article utilized three databases, namely Web of Science, Scopus, and Google Scholar. However, the paper did not specifically address digital images but rather focused on the use of 3D point clouds technology for the recognition and reconstruction of cultural heritage buildings. Tsaftaris et al. (2014) discussed algorithms used in the colorization of black and white historical photo-graphs. Vijendran & Deepa (2014) studied literature on the methodology used in retrieving text and images but did not address its application to ancient manuscripts or cultural heritage.

Based on the literature review articles that we found, there is currently no research that specifically discusses the use of optimization in the retrieval of digital images, particularly for ancient manuscripts. Therefore, in this article, a systematic literature review of research on this topic is complemented with bibliometric analysis. The focus that is covered in this article compared to the other review articles are summarized in Table 1.

Table 1
Summary of the main focus that is covered in this article compared to the other review articles

Author	SLR	Search Strategy	Data Extraction	Optimization	Digital Image Retrieval	Cultural Heritage	Ancient Manuscript
Rodrigues et al. (2023)	✓	✓	✓	✓	-	-	-
Madzík et al. (2023)	✓	✓	✓	✓	-	-	-
Cotella (2023)	✓	✓	✓	✓	-	-	-
Tsaftaris et al. (2014)	-	-	-	✓	✓	-	✓
Vijendran & Deepa (2014)	-	-	-	✓	✓	-	-
Our Article	✓	✓	✓	✓	✓	✓	✓

3. Method

3.1 Article collection

This research was conducted by searching for publications indexed in four data-bases: Science Direct, Scopus, Dimensions, and Google Scholar (accessed using the Publish or Perish application) using specific keywords. We limited the publication time from 2019 to 2023. The search was also restricted to only consider articles published in open-access journals in the English language, with a subject area in mathematics. Data retrieval in the Scopus, Science Direct, and Dimensions digital libraries was applied to the "title, abstract, and keywords" of the publications. However, in the Google Scholar database, the keywords were only applied to the title, as the Google Scholar search engine does not provide a search process using the abstract.

The first step was to search for articles using the initial keyword “RETRIEVAL METHODS DIGITAL IMAGE”, resulting a total of 811 articles, consisting of 583 from Scopus, 1 from Science Direct, 226 from Dimensions, and 1 from Google Scholar. Next, the keyword was expanded to include “OPTIMIZATION MODEL” to search for articles that link optimization models with retrieval methods for digital images. However, through this additional keyword, no articles were obtained from the four databases. This indicates that there are currently no articles that contain the combination of these two keywords.

Furthermore, we searched for articles that contained the keywords “RETRIEVAL METHODS DIGITAL IMAGE” AND “ANCIENT MANUSCRIPT” and “OPTIMIZATION MODEL” AND “ANCIENT MANUSCRIPT”, as well as the combination of both. However, the search results also indicated that there were no articles that contained the combination of these keywords. In Table 2, the summary of the searching and filtering on the four databases is presented. The “Type” column in Table 2 is representing the following keywords:

A. RETRIEVAL METHODS DIGITAL IMAGE
B. OPTIMIZATION MODEL
C. ANCIENT MANUSCRIPT

Table 2
Total number of publications from four databases with five keyword combinations

Keywords	Type	Scopus	Science Direct	Dimensions	Google Scholar	Total
Keyword 1	A	583	1	226	1	811
Keyword 2	A AND B	0	0	0	0	0
Keyword 3	A AND C	0	0	0	0	0
Keyword 4	B AND C	0	0	0	0	0
Keyword 5	A AND B AND C	0	0	0	0	0
Total		583	1	226	1	811

3.2 Selection Method

The selection method used in this research is divided into several stages as shown in Fig. 1. The first stage is identification. This stage is a data search process that has been described in the previous subsection and 811 articles were obtained.

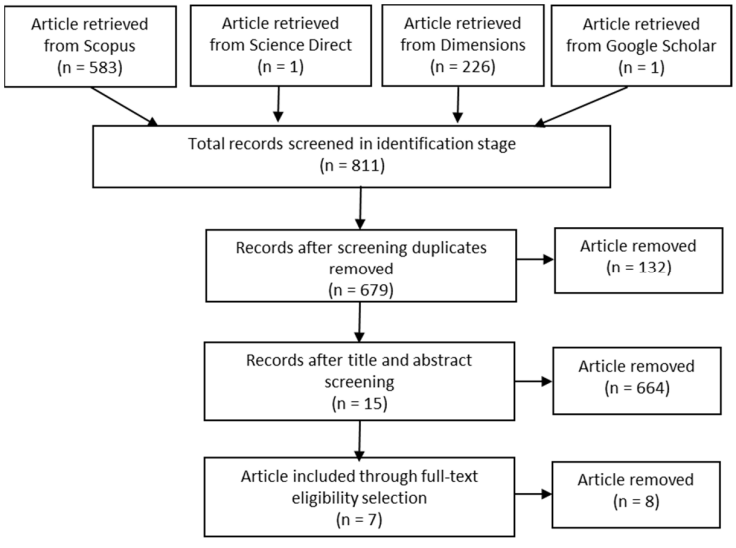

Fig. 1. The data selection process diagram

Furthermore, the selection proceeds to the screening stage. At this stage the article first goes through the duplication selection stage. The duplicate article data is then removed and generates a new total. At this stage, 132 duplicate papers are removed. The data is then selected based on the relevance of the title and abstract. In this screening stage, there are 15 articles obtained, which are referred to as Dataset 1. It should be noted that the data obtained from digital libraries that is used initially had different formats. Therefore, the data format of the article from Science Direct, Dimensions, and Google Scholar is adjusted, so that it matches the data from Scopus. After that, bibliometric analysis was then carried out for Dataset 1. The next stage is eligibility selection. At this stage, we conducted manual filtering by reading the full text of the selected articles to further evaluate the relevance of each article. From this selection, we obtained 7 articles that met our specific criteria. This filtered dataset is referred to as Dataset 2, which will be subjected to further analysis. The selection process results are shown in Table 3.

Table 3
Number of publications from four databases with five keyword combinations.

Keywords	Total	Duplication		Abstract and Title		Full Text	
		Included	Excluded	Included	Excluded	Included	Excluded
Keyword 1	811	679	132	15	551	7	8
Keyword 2	0	0	0	0	0	0	0
Keyword 3	0	0	0	0	0	0	0
Keyword 4	0	0	0	0	0	0	0
Keyword 5	0	0	0	0	0	0	0
Total	811	679	132	15*	551	7**	8

*Dataset 1 for bibliometric analysis, **Dataset 2 for literature review.

3.3 Analysis of Bibliometric

In this section, a bibliometric analysis for Dataset 1 is performed using aprogram called R-bibliometrix. Bibliometrix is a package for bibliometricanalysis written in R, which is an open-source software (Derviş, 2019). The R-bibliometrix package comesequipped with the “biblioshiny” command, which allows for thefunctional combination of bibliometrix in a web-based interface (Rusliana et al., 2022). This was used to obtainscientific data mapping and conduct a comprehensive analysis of the availablebibliographic information.

4. Results

4.1 Results from Bibliometric Analysis

In this section, the results of bibliometric analysis of Dataset 1 are presented. The results are performed using R-bibliometric. The 15 articles in Dataset 1 have a timespan between 2019 to 2023 with 5 years of publication, 7.33 average citations per document, 3.6 co-authors per document, 41 author keywords, and 730 references. Country production is dominated by China with a total of 14 publications, followed by the USA with 6 publications.

4.1.1 Co-occurrence Network

The co-occurrence analysis of Dataset 1 is conducted by searching for the most frequent words that appear in all documents. Dataset 1 contains data obtained from titles, keywords, and abstracts of the articles. The co-occurrence network illustrates how terms present in the articles are related to each other (see Fig. 2).

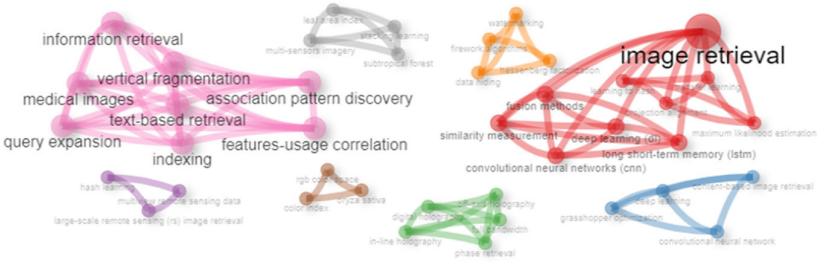

Fig. 2. Co-occurrence network of Dataset 1

Based on Fig. 2, we can observe that the terms appearing are connected to at least one other term. These connections indicate the interrelation between different topics. For example, “image retrieval” is linked to “fusion method”, and “information retrieval” is connected to “vertical fragmentation”. Additionally, the words appearing in Fig. 1 form Fig. 8 clusters of words that are separated and distinguished by different colors. The presence of these clusters of words indicates that This means that there are several topic clusters that share similarities and often appear together in articles, but they do not form strong

connections among each other. Each cluster of words may have different topic coverage and is discussed more extensively in specific contexts. Furthermore, in this co-occurrence network, terms related to retrieval methods dominate, and there are no terms related to optimization, ancient manuscripts, or cultural heritage. This could indicate that the use of these three topics is still relatively limited in re-search on image retrieval methods.

4.1.2 Thematic Evolution

The overview of the evolution of themes is acquired using R-Bibliometrix. This analysis can determine important information related to the development of research topics over time. In this section, the overall number of published articles in Dataset 1 is analyzed. In Fig. 3, the number of issues per year from Dataset 1 for the four databases is presented.

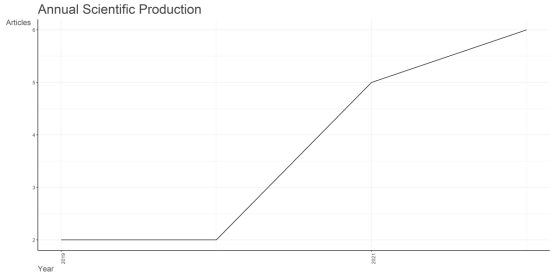

Fig. 3. Yearly number of publications in Dataset 1

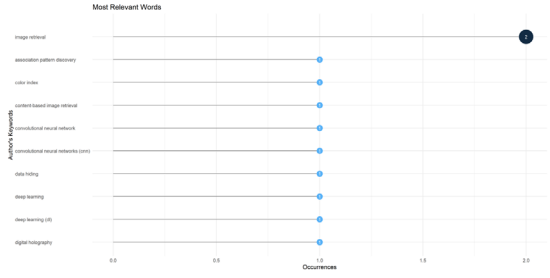

Fig. 4. Most relevant words in the keywords field

As seen in Fig. 3, the number of publications from the year 2019 to 2021 continued to increase. In 2019 and 2020 there are two articles each. This value increases in 2021 to 5 articles and 6 articles in 2022. Fig. 4 shows the 10 most relevant words that appear in the articles in Dataset 1. We can see that apart from “image retrieval”, which occurs twice, the other terms appear only once. The higher occurrence of “image retrieval” indicates that this term is the main focus of the conducted research. Additionally, the presence of other terms that only appear once demonstrates the diversity of topics or variations in the studied research. Furthermore, through R-bibliometrix, we can also obtain a Three Field Map, as seen in Fig. 5. This map depicts the interrelation of research topics with authors and their affiliations. The left column indicates the countries involved in the publication of the articles, the middle column displays the names of authors, and the right column shows the keywords that appear in the articles. From Fig. 5, we observe that not all authors conduct research on the same topics. For instance, the term “full bandwidth” is connected to three different authors, while "image retrieval" is only linked to two authors. The variation in authors' involvement with re-search topics indicates that within the studied collection of articles, there is diversity in authors' research interests.

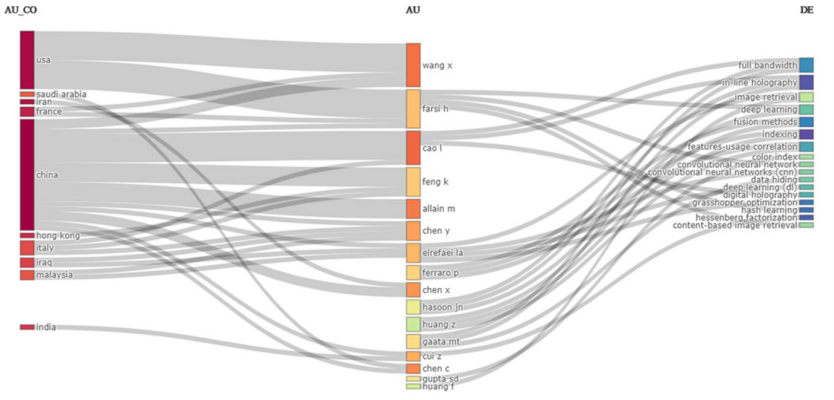

Fig. 5. Three Field Map for Dataset 1

4.2 Result from Systematic Literature Review

The results of the study from Dataset 2 are presented, in the form of 7 selected articles. In Dataset 2, the articles were published within the 2019-2023 time frame.

4.2.1 The Purpose of Research on Optimization for Retrieval Methods Digital Image of Ancient Manuscript

In this section, we conducted an analysis based on RQ1: What are the research objectives regarding optimization in digital image retrieval of ancient manuscripts?. After carefully reading the seven articles in Dataset 2, a summary of the objectives is given in Table 4.

Table 4
Objectives and relevant topics of the articles in Dataset 2.

Author	Article Objective	Image Retrieval	Involving Optimization	Ancient Manuscript
(C. Chen et al., 2022)	Construct a novel approach to online hashing for supervised image retrieval that balances the similarity between multiple samples while assigning code words	✓	✓	-
(Zhang et al., 2020)	Propose a novel unsupervised hashing method for image retrieval that can handle real-world scenarios where there may not be sufficient training images or labeled data	✓	✓	-
Khayyat & Elrefaei (2020)	Develop a method for image retrieval using deep learning that incorporates a variety of fusion levels	✓	✓	✓
Kandel et al. (2021)	Discusses the use of the LM and PLM algorithms for optimizing the Gaussian error metric for the ptychography problem	✓	✓	-
Gao & Cao (2021)	Propose a generalized optimization framework for high-fidelity and high-resolution holographic imaging by combining phase retrieval and pixel super-resolution techniques as a unified optimization problem	-	✓	-
(Mohan & Gupta, 2019)	Develop a new imaging method using a smartphone to acquire rice leaf images under field conditions and modeling approaches to retrieve the leaf chlorophyll content from digitized images	-	✓	-
Sezavar et al. (2019)	Propose a new method for content-based image retrieval that combines a Convolutional Neural Network (CNN) with a Modified Grasshopper Optimization Algorithm (MGOA) to efficiently search a database and retrieve images that are similar to a query image	✓	✓	-

4.2.2 The Role of Optimization in Image Retrieval Method of Ancient Manuscript

In this section, an analysis of the role of optimization in image retrieval methods for the preservation of ancient manuscripts is conducted. Based on the found articles, several methods used in the retrieval of digital images are related to optimization. The summary is described in Table 5.

Table 5
Studies of method and optimization role of Dataset 2.

Author	Method	Optimization Role
Chen et al., 2022)	Discriminative Similarity-Balanced Online Hashing (DSBOH)	Optimization is used in hashing techniques to enhance the efficiency and performance of the hash function, thereby speeding up the data search and access process.
(Zhang et al., 2020)	Optimal Projection Guided Transfer Hashing (GTH).	Optimization involves hashing methods and is used to derive two projections of target and source domains.
Khayyat & Elrefaei (2020)	Optimized bidirectional LSTM deep learning model with attention and batch normalization layers.	The process of training and using BiLSTM involves optimization steps to find suitable model parameters and optimize their performance in specific tasks.
Kandel et al. (2021)	LM and PLM algorithm	Optimization involves solving a problem to minimize the difference between the measured diffraction patterns and the predicted diffraction patterns calculated from the object and probe variables by finding the appropriate object and probe variables
Gao & Cao (2021)	Iterative projection algorithms and gradient descent algorithms	The optimization process supports in reconstructing high-precision and high-resolution holographic images by improving the accuracy and pixel value resolution
(Mohan & Gupta, 2019)	ANN model	Optimization is involved in the development of a feed-forward backpropagation type network to predict the chlorophyll content of rice leaves from digital images
Sezavar et al. (2019)	Grasshopper optimization algorithm (MGOA)	Optimization formulated to efficiently search the database and retrieve similar images by minimizing the distance between the query image features and the features of the image in the database.

5. Discussion

5.1 The State-of-the-Art of Optimization of Retrieval Methods Digital Image of Ancient Manuscript

In Table 4 and Table 5, a review of Dataset 2, which includes the state-of-the-art of our research, is presented. These two tables summarize the objectives, methods, and role of optimization discussed in each of the articles from Datasets 2.

5.2 Research Gaps

Based on the results of the bibliometric analysis, in Fig. 2 no keywords related to optimization are found. In addition, keywords related to cultural heritage, especially ancient manuscripts, have not yet appeared on the co-occurrence network. Therefore, research on optimization and ancient manuscripts in image retrieval has not become the focus of research related to image retrieval. Based on Tables 4 and 5, image retrieval is discussed by five articles. The two articles that do not discuss image retrieval are Gao and Cao (2021) and Mohan and Gupta (2019). Gao and Cao (2021) discussed phase retrieval and pixel super-resolution imaging using a generalized optimization framework. Meanwhile, Mohan and Gupta (2019) did not specifically discuss image retrieval. However, this article discusses the retrieval of Chl content in rice leaves using various RGB color indices through the use of image analysis. All articles in Dataset 2 involve an optimization process. The methods used in the articles including hashing techniques (C. Chen et al., 2022; Zhang et al., 2020), LSTM deep learning models (Khayyat & Elrefaei, 2020), LM and PLM algorithms (Kandel et al., 2021), and Grasshopper Optimization Algorithm

(MGOA) (Sezavar et al., 2019). The model proposed by Khayyat and Elrefaei (2020) is not directly related to the optimization model, but in practice, the use of LSTM involves optimization steps to find the appropriate model parameters and optimize their performance in specific tasks. But there are still several other optimization methods that can be used in image retrieval, such as Support Vector Machines (SVM) (Chen et al., 2001; Renita & Christopher, 2020), Genetic Algorithms (Lai & Chen, 2011; Magliani et al., 2019; Syam & Rao, 2013), Ant Colony Optimization (Ye et al., 2021), or combination of two or more methods. Moreover, the article by Khayyat and Elrefaei (2020) is the only publication that focuses on its implementation for ancient manuscripts. In this instance, the article applies the approach to images of historical Arabic manuscripts. Although the other articles do not explicitly mention the application to ancient manuscripts, their methods could potentially be adapted for various image retrieval tasks, including digital image retrieval for ancient manuscripts.

5.3 The description of the most cited articles in this paper

In this paper, there are 15 articles in the list of references that is highly cited in Section 1 to Section 5. To give more description, in this subsection a discussion for these 15 most cited articles in this paper is presented. Citations of the articles are calculated by counting its appearance in every section in this paper, i.e., form Section 1 to Section 5. As it is presented in Table 6, the discussion on content analysis of optimization and exploration of retrieval methods digital image of ancient manuscript is complete the discussion of this review paper. The use of optimization methods is checked to these 15 cited articles. The content on exploration of retrieval methods digital image of ancient manuscript is also checked. Furthermore, it can be seen in Table 6, the most cited articles in this paper are the papers as listed in Dataset 2 (see Table 4 and Table 5). This means that the PRISMA results in Section 3.2 is relevant with the most cited articles in this paper.

Table 6
The description of the most cited articles in this paper

No	Authors	Citation	Description	OT[1]	DI[2]
1	Khayyat & Elrefaei (2020).	5	A novel approach that utilizes a fusion model for classifying and retrieving historical Arabic manuscripts' images.	√	√
2	Gao & Cao (2021)	4	Combine the phase retrieval and pixel super-resolution techniques as a unified optimization problem and propose a generalized algorithmic framework for pixel-super-resolved phase retrieval. In particular, the iterative projection algorithms and gradient descent algorithms for solving this problem is introduced.	√	√
3	Mohan, & Gupta (2019)	4	A feed-forward backpropagation-type network is developed for the optimization of hidden neurons, training, and transfer functions.	√	√
4	Zhang et al. (2020)	4	A minimization on the error matrix between two hashing projections of target and source domains is discussed. The aim is to seek for the maximum likelihood estimation (MLE) solution of the error matrix between the two hashing projections due to the domain gap.	√	√
5	Chen et al., 2022	3	A novel discriminative similarity-balanced online hashing (DSBOH) framework for Supervised Image Retrieval is presented	√	√
6	Kandel et al. (2021)	3	Specify the physics-based forward model for a specific imaging application; the first and second-order derivative terms are calculated automatically through matrix-vector products, without explicitly forming the large Jacobian or Gauss-Newton matrices typically required for the LM method	√	√
7	Rodrigues et al. (2023)	3	a systematic literature review to examine the state-of-the-art about the implications of digital transformation in tourism as a catalyst for sustainable development, identifying gaps and providing directions for future research.	×	×
8	Sezavar et al. (2019)	3	Modified grasshopper optimization algorithm (MGOA) is proposed to solve modeled problem and to retrieve similar images efficiently.	√	√
9	Vijendran, & Deepa (2014)	3	The central contention is handled by a semantic search and by providing re-ranking procedure. In this survey, various approaches for retrieving text and images using various retrieval systems, semantic search and ranking based techniques are discussed.	×	√
10	Ahmad & Sharma, 2020	2	This paper reviews scholarly literature and analyses it to understand the application of standards for digitisation, access, and preservation in the digital process by Indian institutions.	×	√
11	Bagasi & Elrefaei (2018)	2	The Content-Based Image Retrieval (CBIR) system is proposed to retrieve the Arabic manuscript images. The system has three stages: Preprocessing, feature extraction, and feature similarity matching.	×	√
12	Condorelli & Morena, (2023)	2	The recovery of past architecture through 3D modelling is an important challenge today to the preservation of heritage is proposed.	×	√
13	Cotella (2023)	2	Offering a constructive synthesis that will provide as a springboard for the advancement of innovative strategies in the field of BIM and AI.	×	√
14	Madzík et al. (2023)	2	City and urban planning, Social media, Data analytics, Sustainable and economic development, Technology-based experience and interaction, Cultural heritage, Digital destination marketing and Smart tourism management are the identified eight topics related to DT in the tourism industry.	×	√
15	Tsaftaris et al. (2014)	2	The formulation of a set of equations for missing color is discussed. The localized color information and the grayscale intensities of the black and white historical photographs formulate a set of equations for the missing color values of the remaining pixels.	×	√

[1]OT: Paper uses Optimization Techniques
[2]DI: The topic is on the exploration of retrieval methods digital image of ancient manuscript

6. Conclusion

In this review article, a systematic literature review regarding the role of optimization in the image retrieval method for ancient manuscripts is discussed. The results show that there are 811 articles obtained through Google Scholar, Scopus, Science Direct, and Dimensions. All these articles were selected through the identification, screening, and eligibility stages and produced 15 articles after the results of the abstract screening (Dataset 1), and the results of the full text screening of 7 articles (Dataset 2). Dataset 1 is used for bibliometric analysis and obtaining co-occurrence networks and thematic evolution with the help of R-bibliometrix software. Furthermore, according to the results of the systematic literature review, a conclusion converges to the fact that the research on the use of optimization in digital image retrieval methods for ancient manuscripts is still limited. Although there have been several studies that examine the topic of digital image retrieval and cultural heritage preservation separately, the combination of these two topics has not been widely explored. In searching related articles, it was also found that several optimization techniques have been used in image retrieval in general, but the focus on ancient manuscripts is still relatively under-studied.

Acknowledgements

This research was funded by Universitas Padjadjaran Doctoral Dissertation Research Grant or Riset Disertasi Doktor Universitas Padjadjaran (RDDU) entitled Optimisasi dan Eksplorasi Metode Temu Kembali Citra Digital Manuskrip Kuno sebagai Upaya Pelestarian Warisan Budaya for the year of 2020 with contract number 1427/UN6.3.1/LT/2020

References

Ahmad, A., & Sharma, S. (2020). Sustainable digital preservation and access of heritage knowledge in India: A review. *DESIDOC Journal of Library & Information Technology*, *40*(5), 321–325.

Bagasi, B., & Elrefaei, L. A. (2018). Arabic manuscript content based image retrieval: A comparison between SURF and BRISK local features. *International Journal of Computing and Digital Systems*, *7*(06), 355–364.

Chen, C., Wang, X., Chen, X., Lan, R., Liu, Z., & Luo, X. (2022). Discriminative Similarity-Balanced Online Hashing for Supervised Image Retrieval. *Scientific Programming*, *2022*. https://doi.org/10.1155/2022/2809222

Chen, Y., Zhou, X. S., & Huang, T. S. (2001). One-class SVM for learning in image retrieval. *Proceedings 2001 International Conference on Image Processing (Cat. No. 01CH37205)*, *1*, 34–37.

Condorelli, F., & Morena, S. (2023). Integration of 3D modelling with photogrammetry applied on historical images for cultural heritage. *VITRUVIO-International Journal of Architectural Technology and Sustainability*, *8*, 58–69.

Cotella, V. A. (2023). From 3D point clouds to HBIM: Application of Artificial Intelligence in Cultural Heritage. *Automation in Construction*, *152*, 104936.

Derviş, H. (2019). Bibliometric analysis using bibliometrix an R package. *Journal of Scientometric Research*, *8*(3), 156–160.

Firdaniza, F., Ruchjana, B. N., Chaerani, D., & Radianti, J. (2021). Information diffusion model in twitter: A systematic literature review. *Information*, *13*(1), 13.

Foster, G., & Kreinin, H. (2020). A review of environmental impact indicators of cultural heritage buildings: A circular economy perspective. *Environmental Research Letters*, *15*(4), 43003.

Gao, Y., & Cao, L. (2021). Generalized optimization framework for pixel super-resolution imaging in digital holography. *Optics Express*, *29*(18), 28805. https://doi.org/10.1364/oe.434449

Ishar, S. I., Zlatanova, S., & Roberts, J. L. (2022). 3D GAMING FOR YOUNG GENERATIONS IN HERITAGE PROTECTION: A REVIEW. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *48*, 53–60.

Jayanthi, N., Indu, S., Hasija, S., & Tripathi, P. (2017). Digitization of ancient manuscripts and inscriptions-a review. *Advances in Computing and Data Sciences: First International Conference, ICACDS 2016, Ghaziabad, India, November 11-12, 2016, Revised Selected Papers 1*, 605–612.

Kandel, S., Maddali, S., Nashed, Y. S. G., Hruszkewycz, S. O., Jacobsen, C., & Allain, M. (2021). Efficient ptychographic phase retrieval via a matrix-free Levenberg-Marquardt algorithm. *Optics Express*, *29*(15), 23019. https://doi.org/10.1364/oe.422768

Khayyat, M. M., & Elrefaei, L. A. (2020). Manuscripts Image Retrieval Using Deep Learning Incorporating a Variety of Fusion Levels. *IEEE Access*, *8*, 136460–136486. https://doi.org/10.1109/ACCESS.2020.3010882

Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele, UK, Keele University*, *33*(2004), 1–26.

Lai, C.-C., & Chen, Y.-C. (2011). A user-oriented image retrieval system based on interactive genetic algorithm. *IEEE Transactions on Instrumentation and Measurement*, *60*(10), 3318–3325.

López, F. J., Lerones, P. M., Llamas, J., Gómez-García-Bermejo, J., & Zalama, E. (2018). A review of heritage building information modeling (H-BIM). *Multimodal Technologies and Interaction*, *2*(2), 21.

Madzík, P., Falát, L., Copuš, L., & Valeri, M. (2023). Digital transformation in tourism: bibliometric literature review based on machine learning approach. *European Journal of Innovation Management*, *26*(7), 177–205.

Magliani, F., Sani, L., Cagnoni, S., & Prati, A. (2019). Genetic algorithms for the optimization of diffusion parameters in content-based image retrieval. *Proceedings of the 13th International Conference on Distributed Smart Cameras*, 1–6.

Mohammed, M. M., Badr, A., & Abdelhalim, M. B. (2015). Image classification and retrieval using optimized pulse-coupled neural network. *Expert Systems with Applications*, *42*(11), 4927–4936.

Mohan, P. J., & Gupta, S. D. (2019). Intelligent image analysis for retrieval of leaf chlorophyll content of rice from digital images of smartphone under natural light. *Photosynthetica*, *57*(2), 388–398. https://doi.org/10.32615/ps.2019.046

Remondino, F., & Rizzi, A. (2010). Reality-based 3D documentation of natural and cultural heritage sites—techniques, problems, and examples. *Applied Geomatics*, *2*, 85–100.

Renita, D. B., & Christopher, C. S. (2020). Novel real time content based medical image retrieval scheme with GWO-SVM. *Multimedia Tools and Applications*, *79*(23–24), 17227–17243.

Rodrigues, V., Eusébio, C., & Breda, Z. (2023). Enhancing sustainable development through tourism digitalisation: a systematic literature review. *Information Technology & Tourism*, *25*(1), 13–45.

Rusliana, N., Komaludin, A., & Firmansyah, M. F. (2022). A Scientometric Analysis of Urban Economic Development: R Bibliometrix Biblioshiny Application. *Jurnal Ekonomi Pembangunan*, *11*(2), 80–94.

Sahoo, J., & Mohanty, B. (2015). Digitization of Indian manuscripts heritage: Role of the National Mission for Manuscripts. *IFLA Journal*, *41*(3), 237–250.

Sezavar, A., Farsi, H., & Mohamadzadeh, S. (2019). A modified grasshopper optimization algorithm combined with CNN for content based image retrieval. *International Journal of Engineering, Transactions A: Basics*, *32*(7), 924–930. https://doi.org/10.5829/ije.2019.32.07a.04

Shukla, H. S., Kumar, N., & Tripathi, R. P. (2015). Image restoration using modified binary particle Swarm Optimization Richardson-Lucy (MBSO-RL) algorithm. *Int. J. Appl. Eng. Res.*, *10*(22), 43077–43081.

Skublewska-Paszkowska, M., Milosz, M., Powroznik, P., & Lukasik, E. (2022). 3D technologies for intangible cultural heritage preservation—literature review for selected databases. *Heritage Science*, *10*(1), 1–24.

Syam, B., & Rao, Y. (2013). An effective similarity measure via genetic algorithm for content based image retrieval with extensive features. *Int. Arab J. Inf. Technol.*, *10*(2), 143–151.

Tonazzini, A., Salerno, E., Abdel-Salam, Z. A., Harith, M. A., Marras, L., Botto, A., Campanella, B., Legnaioli, S., Pagnotta, S., Poggialini, F., & others. (2019). Analytical and mathematical methods for revealing hidden details in ancient manuscripts and paintings: A review. *Journal of Advanced Research*, *17*, 31–42.

Tsaftaris, S. A., Casadio, F., Andral, J.-L., & Katsaggelos, A. K. (2014). A novel visualization tool for art history and conservation: Automated colorization of black and white archival photographs of works of art. *Studies in Conservation*, *59*(3), 125–135.

Valova, I., Rachev, B., & Vassilakopoulos, M. (2006). Optimization of the algorithm for image retrieval by color features. *International Conference on Computer Systems and Technologies-CompSysTech*, 1–4.

Vijendran, A. S., & Deepa, C. (2014). A survey on various optimization techniques to retrieve text and images. *Journal of Emerging Technologies in Web Intelligence*, *6*(2).

Ye, F., Meng, X., Chen, S., & Xin, J. (2021). Content-based remote sensing image retrieval based on ant colony optimization and convolutional neural networks. *2021 14th International Symposium on Computational Intelligence and Design (ISCID)*, 382–386.

Zhang, L., Liu, J., Yang, Y., Huang, F., Nie, F., & Zhang, D. (2020). Optimal projection guided transfer hashing for image retrieval. *IEEE Transactions on Circuits and Systems for Video Technology*, *30*(10), 3788–3802. https://doi.org/10.1109/TCSVT.2019.2943902