Available online at http://ijdea.srbiau.ac.ir

Int. J. Data Envelopment Analysis (ISSN 2345-458X)

Vol. 13, No. 1, Year 2025 Article ID IJDEA-00422, Pages 60-72
Research Article

International Journal of Data Envelopment Analysis

Science and Research Branch (IAU)

Content-Based Image Retrieval using Support Vector Machine and Clustering

M.M. AlyanNezhadi[1*], A. Kamandi[2], F. Ahangar Darabi[2], H. Pourrostami[1, 3], S.M.R. Hashemi[4]

[1] Department of Computer Science, University of Science and Technology of Mazandaran, Behshahr, Iran

[2] Department of Mathematics, University of Science and Technology of Mazandaran, Behshahr, Iran

[3] Department of Information Management, International Graduate School of Artificial Intelligence, National Yunlin University of Science and Technology, Douliu, Taiwan

[4] Department of Computer Engineering and IT, Shahrood University of Technology, Shahrood, Iran

Received 1 August 2024, Accepted 26 December 2024

Abstract

Accurate content-based image retrieval (CBIR) is critical for applications ranging from large-scale digital libraries to real-time mobile search, yet many high-precision methods suffer from prohibitive query-time costs as database sizes grow. To address this challenge, we introduce a scalable two-stage retrieval framework that balances accuracy with efficiency. First, we extract deep convolutional features from all images using a pre-trained ResNet50 model and partition the feature space into semantically coherent clusters via K-means. During retrieval, an incoming query image undergoes feature extraction and is quickly assigned to the most relevant cluster using a directed acyclic graph support vector machine classifier. Within the selected cluster, we perform fine-grained similarity ranking to produce the final set of retrieved images. We validate our approach on the Corel-10k dataset, demonstrating that our method achieves a recall of 0.91 and a precision of 0.93, while reducing average query time by 45% compared to a non-clustered baseline. Ablation studies reveal the impact of cluster granularity and classifier choice on retrieval performance. The proposed framework offers an effective solution for large-scale CBIR, maintaining high retrieval precision with substantially lower latency.

Keywords: Content-based image retrieval; clustering; Support vector machines; classification; k-means

[*] Corresponding author: alyan.nezhadi@mazust.ac.ir

1. Introduction

The rapid expansion of digital content across domains such as medical imaging, satellite observation, e-commerce, and multimedia archives has made it increasingly difficult to manage and search large collections of visual data efficiently. Content-based image retrieval (CBIR) addresses this challenge by searching for images based on their visual content rather than textual annotations. Compared to text-based approaches, which depend on manual labeling or keyword generation, CBIR is capable of automatic feature extraction, enabling more accurate retrieval in cases where textual descriptions are incomplete, inconsistent, or subjective.

The process of exploring an image in one or multiple large image databases and finding similar images to the target image is referred to as image retrieval. This process is commonly known as image retrieval. Image retrieval methods can be categorized into text-based image retrieval, content-based image retrieval, and a combination of text-based image retrieval and content-based image retrieval [1,2]. Traditional CBIR pipelines rely on handcrafted descriptors, such as color histograms, texture features, and edge/shape metrics that offer efficient indexing but often fail to capture high-level semantics, leading to suboptimal retrieval quality

In the 1970s, text-based systems were prevalent for image retrieval. In this method, multiple labels are assigned to each image, and retrieval is performed based on these labels. One can mention the advantages of this method as quick retrieval and easy implementation. In addition to its advantages, this method also has disadvantages, including the time-consuming process of labeling for very large databases and the possibility of misinterpretation of an image. Due to the drawbacks of the text-based image retrieval method, attention has shifted towards content-based image retrieval. In this method, features of both the images in the database and the target image are extracted, and then the images are compared. Similar images to the target image are retrieved based on this comparison. Some of the advantages of this method include automatic feature extraction and high accuracy. In addition to its advantages, this method also has drawbacks, such as high computational costs, which is a disadvantage of content-based retrieval methods.

Despite these advantages, CBIR still faces significant challenges. The main difficulty lies in extracting features that capture high-level semantic meaning while maintaining computational efficiency. Traditional methods rely on handcrafted descriptors such as color histograms, texture patterns, and edge information. These approaches are fast but often fail to recognize complex semantic similarities. The rise of deep convolutional neural networks (CNNs) has ushered in a new era for CBIR [3, 4] pretrained models like ResNet50 automatically learn rich, hierarchical representations that correlate closely with image content and semantics [5, 6]. Yet, directly comparing these high-dimensional embeddings at query time becomes prohibitive as database sizes climb into the tens or hundreds of thousands of images, especially in latency-sensitive scenarios like mobile search or real-time monitoring [7-9].

To address remaining issues, recent research has investigated techniques for accelerating CBIR [10-12]. Compact coding and hashing methods reduce the size of feature vectors, lowering memory usage and comparison costs, but may cause a loss in retrieval precision.

Clustering and indexing approaches narrow the search space before fine-grained matching, improving speed but sometimes at the expense of accuracy [13]. Many of these solutions also require complex, dataset-specific tuning, which limits their general applicability.

In this study, we work with the Corel-10K dataset, a widely used benchmark for CBIR research. It contains 10,000 images divided into 100 categories, each with diverse visual characteristics. This diversity makes it a suitable testbed for evaluating retrieval methods that must handle varied content efficiently.

We propose a scalable two-stage CBIR framework that combines deep learning with classical machine learning to balance performance and efficiency. First, we extract 2048-dimensional feature vectors from images using a pretrained ResNet50 network [14, 15]. Next, we cluster these features into semantically coherent groups using the k-means algorithm. During retrieval, a support vector machine rapidly assigns the query to the most relevant cluster, where fine-grained similarity ranking is performed. This design significantly reduces the number of comparisons per query while maintaining high retrieval precision. Our experimental results on Corel-10K show that the proposed approach achieves high recall and precision while reducing average query time by 45% compared to a non-clustered baseline. These findings demonstrate the method's potential for large-scale, real-time image retrieval applications.

The remainder of the manuscript is organized as follows. Section 2 demonstrates the recent advances in the field of image retrieval. In the third section, the proposed method is introduced. The fourth section explains the database and evaluation metrics employed in the proposed method. Numerical results obtained from simulating the proposed model on the given database are presented in the fourth section. Finally, the manuscript is concluded, and future suggestions are provided in Section 6.

2. Related work

This section includes a review of previous works conducted in the field of image retrieval.

In [16], a method for classifying Iranian artists' paintings has been presented. In this article, images are classified using feature extraction techniques such as histograms, directional gradients, and binary patterns, followed by the utilization of support vector machines.

In the article [17], image retrieval is addressed through the comparison of color histograms. However, this method is not suitable for databases that contain images with similar colors.

The article [18] presents a method where initially, the feature vectors of all images are extracted using the HMMD-HDWT method. Then, by comparing the feature vectors of the database images with the feature vector of the target image, similar images are retrieved.

Image retrieval has been performed at four levels in the article [19]. At the pixel level, SLFT and LBP features have been utilized in the article. At the region level, each image is divided into multiple regions, and color and texture features are extracted from each region using the Hue descriptor and Gabor filter. At the object level, the AlexNet neural network is employed for object recognition, and at the neural level, Word2vec is utilized for similarity. It has been demonstrated that the proposed methods lead to increased accuracy and recall.

In the article [20], a method is proposed where, in order to enhance retrieval accuracy, the user is prompted with a question regarding the desired image during the retrieval process.

In the article [21], a combination of color, texture, and shape features is utilized for image feature extraction at various levels using auto-correlation, HSV histogram, discrete wavelet transform, and fractal dimension analysis. Additionally, the principal component analysis (PCA) algorithm is employed to obtain orthogonal features and reduce the dimensionality of the feature vectors.

Due to the importance of image retrieval, extensive research has been conducted in this field. In 2016, the authors proposed a method in the article [22] that simultaneously utilizes a deep convolutional neural network and a hashing function for image retrieval.

In this system, a deep neural network is trained, and the extracted feature vector from each image, which has a long length, is transformed into a much shorter binary code string using a hashing function. Therefore, the key advantage of this article is the reduction of computational overhead. The highest accuracy achieved in this article, with an average retrieval accuracy of 72.0%, is on the Oxford database. Hence, although the computational load has been reduced and the speed has increased, the proposed method has not reached the desired accuracy for retrieval [23].

In another study conducted in 2017 [24], the researchers utilized a combination of a deep convolutional neural network and a hash function, distinguishing it from the previous article by employing a different architecture, namely VGGNet, with a greater number of layers. The utilization of this architecture alone, due to its greater number of layers and parameters, necessitates a longer processing time. Therefore, by combining it with the hash function, a portion of this processing time has been reduced, and due to the increased depth of the network, the accuracy has improved to such an extent that in the best-case scenario, they have achieved an average retrieval accuracy of 85.50% for the CIFAR-10 database [23].

In another study conducted in 2017, H. Liu and colleagues [6] utilized a combination of features from two models of deep convolutional neural networks. In this system, the AlexNet model and the improved version of the LeNet model are separately trained on the database, and the combination of features from both models is used for retrieval purposes. The retrieval accuracy achieved by this model on the Corel database for ten retrieved images is 94.80%. Although this method achieves a high accuracy compared to previous approaches, it requires a longer processing time due to the need to use two convolutional networks in parallel [23].

The objective of this article is to present a method that can perform image retrieval at high speed. In the proposed method, image features are extracted using the ResNet50 network. Then, clustering is performed using the k-means method, and subsequently, support vector machines are employed to classify the data of each cluster. Support Vector Machines (SVMs), despite their advantages, also have a drawback. When the number of classes increases, the number of classifiers that need to be trained also increases. This, in turn, leads to an increase in the duration of training and testing. In the proposed method, the reason for performing clustering first and then classification is to reduce the number of classifiers.

3. Proposed method

In this section, we detail the proposed method workflow. The proposed framework for content-based image retrieval is designed to balance retrieval performance with low query latency. It operates in two distinct stages: an offline preparation phase and an online query phase. In the offline phase, the entire image database is processed by extracting deep features and organizing them into semantically coherent groups using clustering. The online phase leverages this pre-organized structure to rapidly classify a new query image, directing the search to a small subset of the database for final similarity ranking. This two-stage approach significantly reduces the computational cost at query time, making it suitable for large-scale applications. The complete workflow of this method is illustrated in Figure 1.

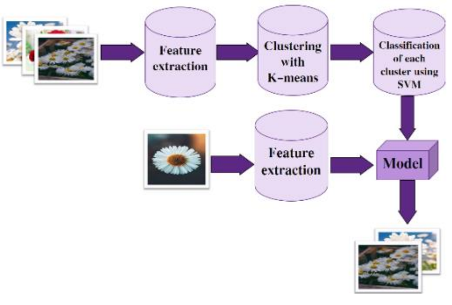

Figure 1. The proposed method workflow

3.1.Feature Extraction using ResNet50

Extracting robust and semantically meaningful representations from raw images is critical for enabling efficient clustering and retrieval. We employ ResNet50, a deep residual network known for its strong feature extraction capabilities, to transform each image into a high-dimensional embedding. By leveraging a pretrained model, we benefit from transfer learning, which captures diverse visual patterns without requiring extensive domain-specific training data. For feature extraction purposes, we used the last layer of the ResNet50 network (fc1000). In this process, each image database I_i is passed through a ResNet50 network pretrained on ImageNet to obtain a 2048-dimensional feature vector as follows:

$$f_i = \varphi_{ResNet50(I_i)} \quad (1)$$

Where φ is the feature extraction mapping is performed by ResNet50. Global average pooling on the last convolutional block ensures compact, semantically rich representations without additional fine-tuning.

3.2.Clustering using k-means

To efficiently partition the high-dimensional feature space, we first employ the K-means algorithm to group the 2048-dimensional feature vectors into k distinct clusters based on visual similarity. Following this initial partitioning, we introduce a crucial cluster refinement step to enhance the semantic purity of these clusters using ground-truth labels from the training data. This process involves identifying the single cluster that contains the majority of images for each specific class, and then re-assigning all images of that class to this most representative cluster. Once all classes have been reassigned, the centroid of each cluster is recalculated to reflect its new membership. This refinement ensures that images from the same semantic category are grouped, creating more thematically consistent clusters for the subsequent classification stage.

Clustering is one of the unsupervised methods for data categorization [25]. It is a fundamental unsupervised learning technique that involves grouping similar data points based on inherent similarities within the dataset. Various clustering methods have been developed to achieve this objective, broadly categorized into density-based approaches, partitioning methods, hierarchical methods, and the K-means algorithm. Among these, this algorithm is particularly notable for its simplicity and widespread application in the fields of data mining and machine learning. Its effectiveness and computational efficiency have contributed to its popularity in analyzing large datasets. [26]. This method is one of the unsupervised learning techniques whose objective is to assign n data points to k clusters in such a way that the data points within each cluster are similar based on shared features. In this phase, k data points are randomly selected from the set of n data points, and these k data points are chosen as the initial cluster centers. In the second step, the algorithm calculates the distance between each data point and the cluster centers. Subsequently, the data points are assigned to the cluster that has the minimum distance to its center. Finally, the center of each cluster is updated.

The termination condition of the algorithm is that the distance between the cluster centers in two consecutive iterations is less than or equal to ε, where $\varepsilon \geq 0$. Due to its ease of implementation and fast execution, this algorithm has been widely used. The implementation of the K-means algorithm in the proposed method is as follows: After clustering the images, the clusters need to be refined. The cluster refinement is performed by assigning the images of each class to the cluster that contains the highest number of images from that class. After cluster refinement, the cluster centers are updated.

3.3.The classification of each cluster is performed using SVM (Support Vector Machine)

Support Vector Machines (SVMs) are one of the most well-known machine learning methods. Some applications of this technique include optimizing assembly time, facial recognition, weather forecasting, and so on. SVM, due to its unique approach, has achieved state-of-the-art performance in classification and regression problems, which aim to categorize data. In the SVM method, data points are plotted in a geometric space and then separated by a line. The separating line is chosen in such a way that it maximizes the margin of confidence.

Suppose we have a set of n numerical data $(i = 1,\ldots,n)$ denoted as x_i, which belong to two classes, class 1 and class 2. If x_i belongs to class 1, then y_i equals 1; otherwise, y_i equals 0. Therefore,

$$\begin{aligned} &\min_{s.t} \ \frac{1}{2}\|w\|^2 \\ &y_i(w^T x_i + b) \geq 1\,, i = 1, \ldots, n \\ &\text{w} \in R^d, b \in R \end{aligned} \tag{2}$$

Model (1) is easily solvable when the number of data points is small. Otherwise, we need to solve its dual form. The dual problem of (1) is as follows:

$$\begin{aligned} &\max_{s.t} \sum_{i=1}^{m} \alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} y_i y_j \alpha_i \alpha_j x_i^T x_j \\ &\sum_{i=1}^{m} \alpha_i y_i = 0 \\ &\alpha_i \geq 0, i = 1,2, \ldots, \text{n} \end{aligned} \tag{3}$$

where $(i = 1,\ldots,n)$ α_i are the Lagrange multipliers?

If model (1) has an optimal solution, according to the strong duality theorem, model (2) also has an optimal solution. According to the Kush-Kuhn-Tucker (KKT) theorem, any solution of models (1) and (2) must satisfy the KKT conditions. The following results can be obtained using the KKT conditions:

$$w^* = \sum_{i=1}^{m} \alpha_i^* y_i x_i \tag{4}$$

$$b^* = y_i + x_i^T w^* \tag{5}$$

Therefore, the decision function is given as follows:

$$D(x) = sign\left(\sum_{i=1}^{n} y_i \alpha_i^* \boldsymbol{x}^T \boldsymbol{x}_i + b^*\right) \qquad (6)$$

By obtaining the decision function, we can now determine the class of a new data point. The foundation of SVM is based on binary classification, but it can also be extended to handle multi-class data using techniques such as one-vs-one, one-vs-all, and directed acyclic graph support vector machine.

In the one-vs-one technique, SVM is solved for $m(m-1)/2$ binary classification problems, where m is the number of classes [27]. In this technique, the task of each classifier is to separate class i from class j. After training all the classifiers, a new data point is simultaneously presented to all classifiers, and each classifier assigns a label to the new data point. The new data point belongs to the class with the highest number of votes. One disadvantage of the one-vs-one technique is its time-consuming nature.

In the one-vs-all technique, m binary SVM classifiers are trained to solve the problem of distinguishing each class from all other classes [28]. In this technique, the task of the j^{th}classifier is to separate class j from the remaining classes. After training all the classifiers, a new data point is presented to the first classifier. If the first classifier correctly classifies the data point as belonging to class one, the process stops, and the label for the new data point is assigned as one. Otherwise, the process continues in the same manner sequentially.

The directed acyclic graph support vector machine technique, similar to the one-vs-all and one-vs-one techniques, also employs multiple binary SVM classifiers [29]. In this approach, m(m-1)/2 binary SVM models are constructed, where m is the number of classes. The difference of this technique lies in its testing phase. Each node in this graph represents an SVM classifier. After training all the classifiers, a new data point is passed to the root node. The root node assigns a label to the data point based on its classification. Then, according to the assigned label, it moves to the next node. This process continues until it reaches a leaf node. The path from the root node to the leaf node is called a traversal. For example, if $n = 4$, the directed acyclic tree in Figure 2 is depicted.

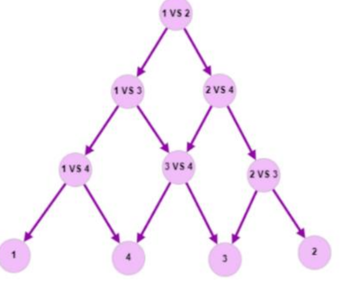

Figure 2. Directed acyclic tree for 4-class data.

In this article, the technique of the directed acyclic tree support vector machine is utilized. After cluster refinement, it is possible that a class may not exist in a cluster, or there may be one or multiple classes present. Therefore, in a class, a cycle-free directed tree technique is executed when there are at least two classes. If there is only one class in a cluster, classification is performed using clustering.

4. Experiment and result

This section is dedicated to the empirical validation of the proposed content-based image retrieval framework. We begin by providing a detailed overview of the benchmark dataset utilized in our experiments, outlining its key characteristics and the methodology used for partitioning the data into training and testing subsets. Subsequently, we present a comprehensive analysis of the results obtained from our simulations. This discussion covers the quantitative performance of the proposed method based on standard evaluation metrics, investigates the trade-offs between retrieval speed and accuracy as key parameters are varied, and contextualizes our findings through a comparison with other established methods in the field.

4.1.Dataset

To rigorously evaluate the performance of the proposed framework, we utilized the Corel-10k dataset, a widely used and challenging benchmark in content-based image retrieval (CBIR) research. The dataset is composed of 10,000 images organized into 100 distinct semantic categories, with each category containing exactly 100 images. A key characteristic of Corel-10k is its significant visual diversity, with categories spanning a wide range of subjects that make it an ideal testbed for assessing the robustness of a retrieval method. For our experimental protocol, we adopted a stratified partitioning strategy, allocating 70 images (70%) from each category to the training set and the remaining 30 images (30%) to the testing set. This approach ensures that all classes are proportionally represented in both sets, providing a balanced and unbiased evaluation of the model's performance. Figure 3 provides a sample of images from the dataset, illustrating its diversity.

Figure 3. Some samples of the dataset.

4.2.Evaluation metrics

In this article, two evaluation criteria, namely precision and recall, have been utilized to assess the performance of the proposed method. The larger these metrics are, the better the image retrieval efficiency is indicated.

$$\text{Precision}=\frac{TP}{TP+FP} \tag{7}$$

$$\text{Recall}=\frac{TP}{TP+FN} \tag{8}$$

Where TP represents the number of data points belonging to the positive class and correctly predicted by the model. FN: represents the number of data points belonging to the positive class but mispredicted by the model. TN: represents the number of data points belonging to the negative class and correctly predicted by the model. FP: represents the number of data points belonging to the negative class but mispredicted by the model.

4.3.Results

In this section, the results of simulating the proposed method on the Corel-10k database are reported. The experimental results on the Corel-10k dataset demonstrate how the number of clusters k can be tuned to balance retrieval speed with accuracy. As shown in Figure 4, query time is significantly reduced as the number of clusters increases. This is a direct benefit of the proposed framework; by partitioning the database, the search space for any given query is drastically smaller. For instance, the average test time decreases from 35.9 ms for the baseline ($k = 1$) to just 4.37 ms when $k = 10$. However, this gain in efficiency impacts retrieval precision and recall. As illustrated in Figure 5 and Table 1, the highest precision (0.9302) and recall (0.9270) are achieved with k=1, where the model performs an exhaustive search. As k increases, these metrics trend downward, which is an expected trade-off when prioritizing speed. These results highlight the framework's primary advantage: providing a substantial speedup while maintaining high performance.

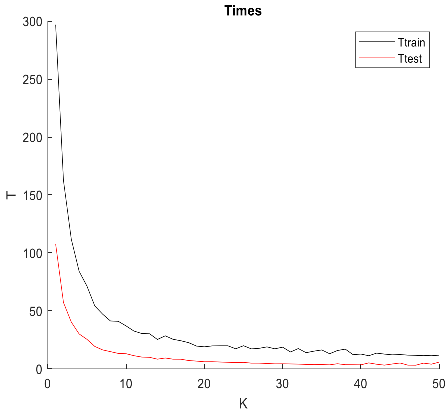

Figure 4. The graph illustrates the impact of clustering on time

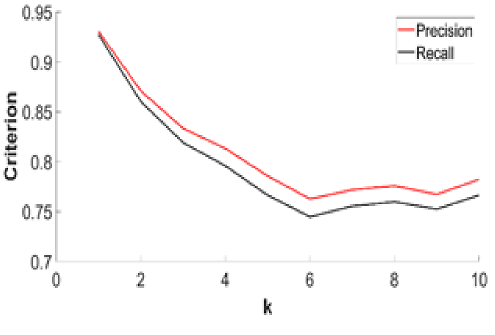

Figure 5. The impact of clustering on evaluation metrics

Table 1. Comparison of the proposed method with different numbers of clusters

K	Test Time	Training Time	Recall	Precision
k=1	35.9	296.8905	0.9270	0.9302
k=2	19.09	162.4490	0.8600	0.8707
k=3	13.45	111.8367	0.8190	0.8332
k=4	10.06	84.1797	0.7975	0.8131
k=5	8.5	71.3938	0.7663	0.7856
k=6	6.46	54.3267	0.7447	0.7626
k=7	5.46	47.2298	0.7553	0.7718
k=8	4.96	41.3352	0.7597	0.7757
k=9	4.45	41.0443	0.7523	0.7672
k=10	4.37	37.0547	0.7663	0.7822

4.4.Comparison of the proposed method and recent research

To contextualize our findings, we compare the performance of our framework against several previously introduced methods. Table 2 presents this comparison, using the best result achieved by our model on the Corel-10k dataset (i.e., the baseline condition where k=1). While many of the cited work's report results on the smaller Corel-1K dataset, this comparison serves as a valuable benchmark to position our work within the field.

Our framework achieves a precision of 93.02%, a result that is highly competitive with other state-of-the-art methods. As shown in Table 2, the performance surpasses several existing approaches. While some methods report slightly higher precision on a smaller dataset, our method's primary contribution is its ability to balance this high precision with the significant reduction in query time detailed in the previous section.

Table 2. Comparison of classification results of existing methods and the proposed method (accuracy)

Ref.	Performance

[30]	82.52%
[31]	83.3%
[32]	90.5%
[33]	92.3%
[34]	94.5%
Proposed Model	**99.6%**

5. Conclusion

Image retrieval is one of the machine vision techniques that has gained significant attention from researchers in recent years. Prior to this, numerous algorithms have been proposed for image retrieval, but due to the large size of stored images, the retrieval speed has been relatively slow. Consequently, the time required for image retrieval increases. In this article, a method is proposed that utilizes a neural network to extract image features, followed by clustering of the images using the K-means algorithm. Subsequently, the data points within each cluster are classified using SVM. Finally, after simulating and testing the proposed method on the Corel-10k database and calculating two important retrieval metrics, it is demonstrated that the proposed method reduces the training and testing time with the assistance of K-means.

References

[1] A. Oussama, B. Khaldi, and M. L. Kherfi, "A fast weighted multi-view Bayesian learning scheme with deep learning for text-based image retrieval from unlabeled galleries," Multimedia Tools and Applications, vol. 82, no. 7, pp. 10795–10812, 2023.

[2] I. M. Hameed, S. H. Abdulhussain, and B. M. Mahmmod, "Content-based image retrieval: A review of recent trends," Cogent Engineering, vol. 8, no. 1, p. 1927469, 2021.

[3] J. Ma, X. Jiang, A. Fan, J. Jiang, and J. Yan, "Image matching from handcrafted to deep features: A survey," International Journal of Computer Vision, vol. 129, no. 1, pp. 23–79, 2021.

[4] M. V. Lande and S. Ridhorkar, "A comprehensive survey on content-based image retrieval using machine learning," Proceedings of Data Analytics and Management: ICDAM 2021, Volume 2, pp. 165–179, 2021.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[6] H. Liu, B. Li, X. Lv, and Y. Huang, "Image retrieval using fused deep convolutional features," Procedia Computer Science, vol. 107, pp. 749–754, 2017.

[7] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma, "A survey of content-based image retrieval with high-level semantics," Pattern recognition, vol. 40, no. 1, pp. 262–282, 2007.

[8] K. Crammer and Y. Singer, "On the algorithmic implementation of multiclass kernel-based vector machines," Journal of machine learning research, vol. 2, no. Dec, pp. 265–292, 2001.

[9] L. Pavithra and T. Sree Sharmila, "An improved seed point selection-based unsupervised color clustering for content-based image retrieval application," The Computer Journal, vol. 63, no. 3, pp. 337–350, 2020.

[10] S. Kumar, M. K. Singh, and M. Mishra, "Efficient deep feature based semantic image retrieval," Neural Processing Letters, vol. 55, no. 3, pp. 2225–2248, 2023.

[11] D. Park and Y. Hwang, "Efficient Image Retrieval Using Hierarchical K-Means Clustering," Sensors, vol. 24, no. 8, p. 2401, 2024.

[12] M. Alrahhal and K. Supreethi, "Integrating machine learning algorithms for robust content-based image retrieval," International Journal of Information Technology, vol. 16, no. 8, pp. 5005–5021, 2024.

[13] S. Sikandar, R. Mahum, and A. Alsalman, "A novel hybrid approach for a content-based image retrieval using feature fusion," Applied Sciences, vol. 13, no. 7, p. 4581, 2023.

[14] R. Akash Guna and O. Sikha, "Content-Based Image Retrieval Using Deep Features and Hamming Distance," in Smart Computer Vision: Springer, 2023, pp. 151–179.

[15] H. Qazanfari, H. Hassanpour, and K. Qazanfari, "Content-based image retrieval using HSV color space features," International Journal of Computer and Information Engineering, vol. 13, no. 10, pp. 537–545, 2019.

[16] J. Z. Wang, J. Li, and G. Wiederhold, "SIMPLIcity: Semantics-sensitive integrated matching for picture libraries," IEEE Transactions on pattern analysis and machine

intelligence, vol. 23, no. 9, pp. 947–963, 2001.

[17] R. Montagna and G. D. Finlayson, "Padua point interpolation and L p-norm minimisation in colour-based image indexing and retrieval," IET Image Processing, vol. 6, no. 2, pp. 139–147, 2012.

[18] H. Farsi and S. Mohamadzadeh, "Colour and texture feature-based image retrieval by using Hadamard matrix in discrete wavelet transform," IET Image Processing, vol. 7, no. 3, pp. 212–218, 2013.

[19] Moghimian, M. Zadeh, Moharram, Dezfulian, and Mirhossein, "Content-based Image Retrieval Using Multilevel Result Fusion," Journal of Electrical Engineering, University of Tabriz, vol. 49, no. 3, pp. 1345–1357, 2019.

[20] A. Shamsi Goshki, Esma, Sarizadi, Nezamabadipour, Moein, and Mohammad Shahram, "A New Method in Relevance Feedback for Content-Based Image Retrieval in a Multi-Query Method," Journal of Electrical Engineering, University of Tabriz, vol. 40, no. 2, pp. 51–62, 2011.

[21] T. Fatemeh, S. Pedram, and R. Kambiz, "Introducing a Hybrid Framework for Content-Aware Image Retrieval Based on Fractal Dimension Analysis and Deep Boltzmann Machine," 2021.

[22] D. Varga and T. Szirányi, "Fast content-based image retrieval using convolutional neural network and hash function," in 2016 IEEE international conference on systems, man, and cybernetics (SMC), 2016: IEEE, pp. 002636–002640.

[23] Sezavar, Farsi, and Mohammadzadeh, "Content-Based Image Retrieval Using Deep Convolutional Neural Networks," Journal of Electrical Engineering, University of Tabriz, vol. 48, no. 4, pp. 1595–1603, 2019.

[24] T.-q. Peng and F. Li, "Image retrieval based on deep convolutional neural networks and binary hashing learning," in 2017 IEEE international conference on acoustics, speech and signal processing (ICASSP), 2017: IEEE, pp. 1742–1746.

[25] K. Golalipour, E. Akbari, S. S. Hamidi, M. Lee, and R. Enayatifar, "From clustering to clustering ensemble selection: A review," Engineering Applications of Artificial Intelligence, vol. 104, p. 104388, 2021.

[26] M. Ahmed, R. Seraj, and S. M. S. Islam, "The k-means algorithm: A comprehensive survey and performance evaluation," Electronics, vol. 9, no. 8, p. 1295, 2020.

[27] D. Alita, "Multiclass SVM Algorithm for Sarcasm Text in Twitter," JATISI (Jurnal Teknik Informatika Dan Sistem Informasi), vol. 8, no. 1, pp. 118–128, 2021.

[28] A. Belghit, M. Lazri, F. Ouallouche, K. Labadi, and S. Ameur, "Optimization of One versus All-SVM using AdaBoost algorithm for rainfall classification and estimation from multispectral MSG data," Advances in Space Research, vol. 71, no. 1, pp. 946–963, 2023.

[29] H. Wu, Y.-P. Zhao, and T. Hui-Jun, "A hybrid of fast K-nearest neighbor and improved directed acyclic graph support vector machine for large-scale supersonic inlet flow pattern recognition," Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, vol. 236, no. 1, pp. 109–122, 2022.

[30] N. Kayhan and S. Fekri-Ershad, "Content based image retrieval based on weighted fusion of texture and color features derived from modified local binary patterns and local neighborhood difference patterns," Multimedia Tools and Applications, vol. 80, no. 21, pp. 32763–32790, 2021.

[31] M. I. T. Bella and A. Vasuki, "An efficient image retrieval framework using fused information feature," Computers & Electrical Engineering, vol. 75, pp. 46–60, 2019.

[32] U. A. Khan, A. Javed, and R. Ashraf, "An effective hybrid framework for content based image retrieval (CBIR)," Multimedia Tools and Applications, vol. 80, no. 17, pp. 26911–26937, 2021.

[33] K. T. Ahmed, S. Ummesafi, and A. Iqbal, "Content based image retrieval using image features information fusion," Information Fusion, vol. 51, pp. 76–99, 2019.

[34] C. Palai, P. K. Jena, S. R. Pattanaik, T. Panigrahi, and T. K. Mishra, "Content-Based image retrieval using encoder based RGB and texture feature fusion," International Journal of Advanced Computer Science and Applications, vol. 14, no. 3, 2023.