IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125233

Clustering based Indexing of Cartoon Images for Retrieval

Amruth V[1*], Santhosh Kumar M[1], Preethi B[2], Amrutha B[3], Prathiksha MS[4]

Assistant Professor, Department of ISE, Maharaja Institution of Technology Mysore, Mandya, India [1*]

Student, Department of ISE, Maharaja Institution of Technology Mysore, Mandya, India[1]

Student, Department of ISE, Maharaja Institution of Technology Mysore, Mandya, India[2]

Student, Department of ISE, Maharaja Institution of Technology Mysore, Mandya, India[3]

Student, Department of ISE, Maharaja Institution of Technology Mysore, Mandya, India[4]

Abstract: IEEE Content-based cartoon image retrieval systems face challenges due to intra-class variability, shape invariance, and scalability. This paper proposes a novel framework combining Scale-Invariant Feature Transform (SIFT) and Histogram of Oriented Gradients (HOG) for feature extraction, enhanced by Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) for dimensionality reduction. A Kd-tree indexing mechanism ensures efficient retrieval. Experiments on a custom dataset of 600 images (30 classes) demonstrate that fused SIFT+HOG achieves 84% higher precision than standalone methods. The system addresses pose variation, background clutter, and scalability, making it suitable for animation studios, advertising, and educational tools.

Keywords: Cartoon retrieval, feature fusion, SIFT, HOG, dimensionality reduction, Kd-tree indexing

I. INTRODUCTION

This Cartoons are a universal medium for storytelling and communication, with applications in entertainment, education, and advertising. The proliferation of digital cartoon databases has created a demand for efficient retrieval systems. Traditional text-based methods rely on manual annotations, which are impractical for large datasets. Content-based image retrieval (CBIR) systems analyze visual features like shape, color, and texture but face challenges unique to cartoons:

1. **Intra-class variability**: Characters appear in diverse poses, backgrounds, and artistic styles.
2. **Invariance requirements**: Retrieval must be robust to scaling, rotation, and translation.
3. **Scalability**: Efficient indexing for real-time queries on large datasets.

[1]This work introduces a hybrid approach combining SIFT (local keypoints) and HOG (global edges) for robust feature representation. Dimensionality reduction via PCA and LDA optimizes storage and computational efficiency, while Kd-tree indexing enables logarithmic-time retrieval in main time to take cartoons on it.

The cartoon based image retrieval systems can be defined as: For a given query cartoon image, finding similar cartoon images stored in database.

[2] The search process relies on how faster the images can be retrieved from the large database. To achieve a fast retrieval speed and to make the retrieval system truly scalable to the large size of the image collections, an effective indexing structure is a paramount part of the whole system.

[3]The cartoon based image retrieval systems can be defined as: For a given query cartoon image, finding similar cartoon images stored in database. The search process relies on how faster the images can be retrieved from the large database. To achieve a fast retrieval speed and to make the retrieval system truly scalable to the large size of the image collections, an effective indexing structure is a paramount part of the whole system.

There are mainly two problems lying on the cartoon image retrieval.
[4] Firstly, the appearance of the cartoon image changes from time to time, when used in different applications, such as advertisement, poster and operating system themes and what is main content in the cartoon image is still not clear. Secondly, most of the global features used in the Content Based Image Retrieval (CBIR) It is well known that cartoon was colourless in history and nowadays there are also lots of colourless cartoon, besides a vast number of colourful

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125233

cartoons. The shape features plays a critical role in searching for similar images.[5] In image retrieval, the shape description needs to be invariant to scaling, rotation, and translation of the image. Shape features are less developed than their colour and texture counterparts as they inherent complexity of representing shapes.

Here to retrieve the images, the shape information for each image is extracted and the special relationship between the images is considered this is used for the correspondence between the query image and the database image. The images contain the edge information, texture information and region information. The similarity between the relationship among the query image and the database image can be used to retrieve the images based on the indexing mechanism.

Fig. 1.1. Images of Narendra Modishowing inter class similarity.

Fig. 1.2. Images of RajiniKanth with artistic style variations

II. LITERATURE SURVEY

A methodology for content-based image retrieval on cartoon images was proposed [1]. The similarity between the query image and database images were computed based on Scale Invariant Feature Transforms (SIFT) and Histogram of Gradient (HOG) and similar images were retrieved. This system was modelled on a dataset consisting of 30 classes of cartoon images without caricatures and sketches. In [2], a cartoon image retrieval system was proposed in which partial features, regions and aspects were used as keys to identify cartoon character images. The limitation of this model was that the system resulted in character design reuse rather than cartoon sequence synthesis. A model called Graph based cartoon clip synthesis (GCCS) that combines similar characters into user-directed sequence based on dissimilarity measure in edges was proposed [3]. This system performed well for simple cartoon characters but it failed for complex pattern GCCS.

Retrieval based Cartoon Clip Synthesis (RCCS) was proposed in [4] which used an unsupervised bi distance metric learning (UBDML) algorithm. This algorithm was based on the dissimilarity measure in motion and edges of cartoon characters. The limitation of this model was that the model failed to retrieve dissimilarity in color and gesture. In [6], a method based on Neighborhood-Exact Nearest Neighbor Search (NENNS) was proposed. This method guaranteed for efficient search and search exactness within a specified neighborhood around the query. A novel system called the contrastive weight aggregation histogram was modelled [7] to distinguish the slightly distinguishable and highly distinguishable features in deep convolutional feature maps for ima retrieval. Further work is needed to generate multi-dimensional contrastive weighting to provide more detailed distinguishable information.

A simple and efficient framework for zero-shot sketch- based image retrieval was examined in [8]. In this model three types of losses were proposed: domain-balanced quadruplet loss, semantic classification loss and semantic knowledge preservation loss. The domain-balanced quadruplet loss addressed the issue of domain-balanced loss which was used to reduce the domain gap. The semantic classification loss was introduced to learn semantic features. The semantic knowledge preservation loss was presented to prevent the rich knowledge learned from dataset from being forgotten during fine-tuning of the pre-trained model. In [9], a fuzzy clustering algorithm for clustering feature vectors of the face database for constructing a retrieval pedigree map was proposed.

The threshold value was set for face image retrieval. Only those images within the threshold range were retrieved thus reducing the number of comparisons. In summary, though there are a few attempts on cartoon image retrieval, in general to the best of our knowledge, there has been no work on celebrity cartoon face image retrieval so far. Further, the existing works have been reported on a small corpus of maximum 30 classes and in addition they have certain limitations. This too motivated us to build an efficient model for the celebrity cartoon face image retrieval. Here in our work, we propose a model by experimenting with IIIT- CFW dataset having 100 classes

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125233

III. METHODOLOGY

A. System Architecture

The above chapter defines about the cartoon image and its retrieval. It describes the challenges and applications of our proposed method and also briefly explains about how different authors proposed their methods in retrieving the cartoon images and it also summarizes the drawbacks of the all the proposed methods and the motivation that leads to carry this work.[2] Chapter 2 describes the cartoon images that are used in the proposed work. In Chapter 3 we discuss about the feature extraction and the dimensionality reduction methodologies that are used in the cartoon image retrieval process. Chapter 4 discusses about the indexing mechanism used for the retrieval process and the results of the experiments on cartoon based image retrieval. The chapter 5 contains the conclusion from this work and future enhancement that can be done.

.

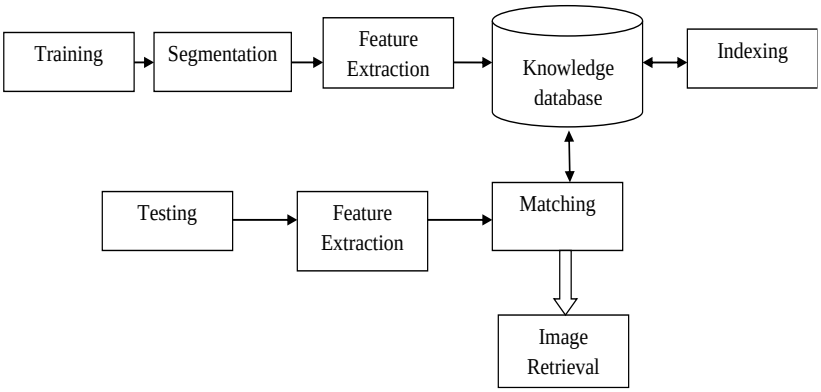


Fig 3.1: Basic architecture of the proposed method

Feature extraction is an important component of a pattern recognition system. A well defined feature extraction algorithm will help in the accurate retrieval of images. The main aim of the feature extraction is to preserve most of the information and remove the redundant information in order to improve the efficiency of the classification without much of degradation.[6] For this it performs two tasks: Firstly, transforming input parameter vector into a feature vector secondly, reducing its dimensionality. In this work, we extracted features for the input query using SIFT (Scale Invariant Feature Transformation), HOG (Histogram of Oriented Gradients) and fusion of this features. The dimensions of the extracted features are reduced using Linear Discriminant Analysis (LDA). The reduced features are indexed for the retrieval process.

A general approach to effectively retrieve the images from the large databases has not developed.[4] Most of the current approaches are designed to handle a large dataset.[8] In our approach size of the database is taken to consideration. The Kd-tree (K- dimensional) indexing schema lies its efficiency from the point of search time.Given a set of n.[3] data points X={X1,.....,Xn } with Xi ε Rk being a K- dimensional point, the aim is to build a tree structure to index these data point so that the nearest neighbours of a query vector Xq can be quickly found.A Kd-tree (K-dimensional tree) is a special case of the binary space portioning trees, which is constructed in a recursive fashion. At the root, the data points are split into two halves by a partition hyper plane. Each half is then assigned to one child node and is recursively splitted to create a balanced binary tree. The leaf node may contain a single point or more than one point in different implementations. Each node in the constructed Kd-tree corresponds to a cell in Rk , bounded by a set of partition hyper planes. A partition hyper plane is perpendicular to a partition axis and it is decided by a partition value. The partition axis in conventional Kd-tree is the coordinate axis which has the greatest variance and partition value is the median of the projections of the data points along the partition axis.The image have some set of features, once the feature set is obtained, it is stored in the database. However storing in the database in an efficient manner is required such that the possible candidate list is selected the matching process should be effective.[9] Hence a backend tool of indexing mechanism is used, which stores the data in some pre-defined manner so that during the matching phase only a few potential candidates are selected. Hence we use Kd-tree based approach for indexing the obtained features. The following subsection provides the overview of Kd-tree indexing method Kd-tree is a space- partitioning data structure for organizing points in a k-dimensional space.[10] It is useful data structure for several applications that involves a multidimensional search key. Kd-trees are a special case of a binary space portioning (BSP) trees which uses only splitting planes that are perpendicular to one of the coordinate system axes whereas in BSP arbitrary splitting planes can be used. Every node of a Kd-tree, from the root to the leaves, whereas in BSP, leaves are typically the only nodes that contain stores points.In this work we have created our own

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125233

database having different images in it to create the database we collected the images from different locations. These are taken to study the effect of our proposed method.[7] The database contains thirty different classes of cartoon images and each class having multiple images in it. Here we have taken SIFT, HOG and fusion of SIFT and HOG descriptors for extracting the features. The experimental results for each class is shown in the below graph. In this chapter we described a proposed method for the cartoon based image retrieval.

IV. ALGORITHM

B. Feature Extraction

SIFT Descriptor:

1. Scale-space extrema detection: Identify keypoints using Difference of Gaussians (DoG).
2. Keypoint localization: Reject low-contrast points via Hessian matrix analysis.
3. Orientation assignment: Compute gradient magnitude m(x,y)$m(x,y)$ and direction θ(x,y)$\theta(x,y)$:
4. m(x,y)=(L(x+1,y)−L(x−1,y))2+(L(x,y+1)−L(x,y−1))2$m(x,y)=(L(x+1,y)-L(x-1,y))2+(L(x,y+1)-L(x,y-1))2$ θ(x,y)=tan−1(L(x,y+1)−L(x,y−1)L(x+1,y)−L(x−1,y))$\theta(x,y)=\tan-1(L(x,y+1)-L(x,y-1)L(x+1,y)-L(x-1,y))$θ(x,y)=tan−1(L(x+1,y)−L(x−1,y)L(x,y+1)−L(x,y−1))
5. Descriptor generation: 128-dimensional vector per keypoint.

HOG Descriptor:

1. Divide image into 16×16 cells.
2. Compute gradient histograms (8 bins) for each cell.
3. Normalize blocks (2×2 cells) to reduce lighting variations.

C. Feature Fusion

SIFT (local keypoints) and HOG (global edges) are concatenated into a 256-dimensional vector. Fusion leverages SIFT's invariance and HOG's edge sensitivity.

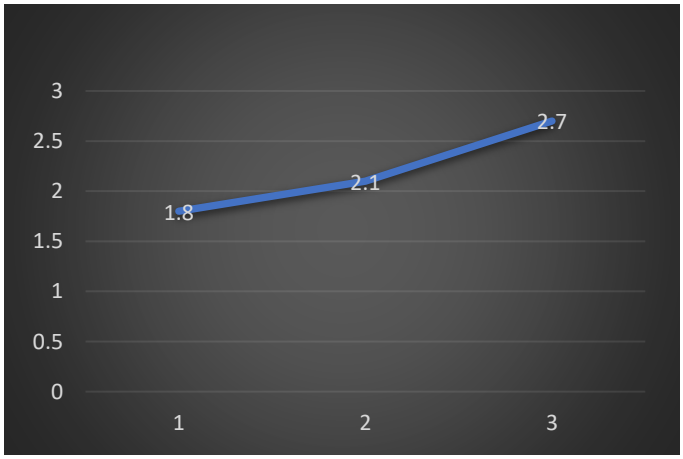
D. Dimensionality Reduction

1. PCA:
   a. Retains 95% variance by selecting top eigenvectors.
   b. Reduces feature space from 256D to 50D.
2. LDA:
   a. Maximizes Fisher's criterion for class separability:J(W)=WTSbWWTSwWJ(W)=WTSwWWTSbW$J(W)=WTSwWWTSbW$
   b. SbSb: Between-class scatter matrix.
   c. SwSw: Within-class scatter matrix.

E. Kd-Tree Indexing

1. Construction: Recursively partition feature space using median splits.
2. Search: Priority-based backtracking for approximate nearest neighbors.

V. RESULT AND DISCUSSION



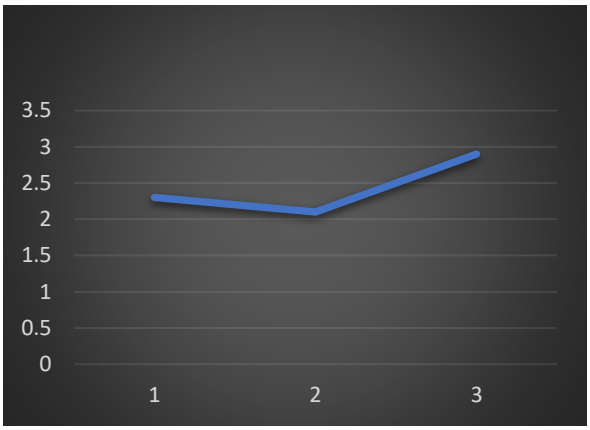
(a)

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125233



(b)

Fig 4.1: Exprimental results of (a) Narendra Modi (b)Rajinikanth

Dims	HOG	SIFT	SIFT+ HOG
d = 5	2.43	2.54	3.12
d= 10	1.89	1.79	2.93
d = 15	1.26	1.35	1.82
d = 20	0.72	0.62	0.94
d = 50	0.53	0.61	0.85

Table 4.2: Dimensionality reduction using Principal Component Analysis

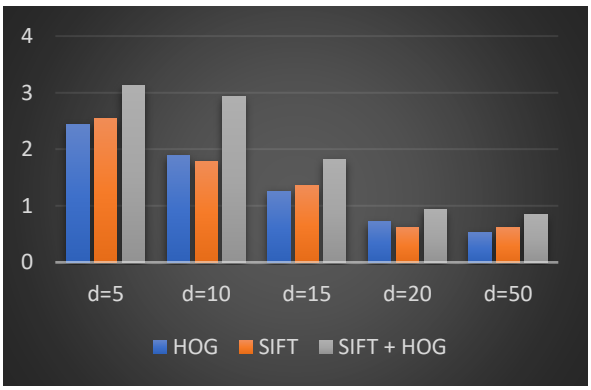


Fig 4.3: cartoon character retrieval with different dimensionalities using Linear Discriminant Analysis.

HOG:- Histogram of Oriented Gradients

SIFT:- Scale Invariant and Feature Transform

In this work we have created our own database having different images in it to create the database we collected the images from different locations. These are taken to study the effect of our proposed method. The database contains thirty different classes of cartoon images and each class having multiple images in it. Here we have taken SIFT, HOG and fusion of SIFT and HOG descriptors for extracting the features. The experimental results for each class is shown in the below graph.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125233

VI. CONCLUSION

The proposed method works on cartoon image retrieval by the user query given to the system using SIFT and HOG descriptors. We have conducted the experimentation on our own dataset. To obtain the efficiency of the proposed method we pre-processed the image then we extracted the shape feature by SIFT and HOG from the query image and the pre-processed image. Later we matched both the images by using extracted feature and retrieve the images from large database based on the kd-tree indexing.

We have observed that the cartoon based image retrieval is still an open problem; so far we have worked on thirty classes , in future the database can be extended to more classes and manually draw the images in an interface and to retrieve the image from large database.

REFERENCES

[1] Suman Muralidhar and Sharath Kumar Y H, "Cartoon based image retrieval: An indexing approach," in Machine Learning and Applications: An International Journal (MLAIJ) Vol.2, No.3/4, December 2015.

[2] M. Haseyama and A. Matsumura, "A trainable retrieval system for cartoon character images," in Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP'03). 2003 IEEE International Conference on, vol. 3. IEEE, pp. III–673, 2003.

[3] J. Yu, D. Liu, D. Tao, and H. S. Seah, "On combining multiple features for cartoon character retrieval and clip synthesis," Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, vol. 42, no. 5, pp. 1413–1427, 2012.

[4] Y. Yang, Y. Zhuang, D. Xu, Y. Pan, D. Tao, and S. Maybank, "Retrieval based interactive cartoon synthesis via unsupervised bi- distance metric learning," in Proceedings of the 17th ACM international conference on Multimedia. ACM, pp. 311–320, 2009.

[5] Mishra, A., Rai, S.N., Mishra, A., Jawahar, C.V. (2016). "IIIT-CFW: A Benchmark Database of Cartoon Faces in the Wild". In: Hua, G., Jégou, H. (eds) Computer Vision – ECCV 2016 Workshops. ECCV 2016. Lecture Notes in Computer Science, vol 9913. Springer, 2016.

[6] Fanglin Chen, Wenjie Pei, Guangming Lu, "Neighborhood-Exact Nearest Neighbor Search for face retrieval," Knowledge-Based Systems, Volume 248, 108757, 2022.

[7] Fen Lu, Guang-Hai Liu, "Image retrieval using contrastive weight aggregation histograms," Digital Signal Processing, Volume 123, 103457, 2022.

[8] Tursun, Osman, et al. "An efficient framework for zero-shot sketch- based image retrieval." Pattern Recognition 126, 108528, 2022

[9] Dong, Xiajun, Bin Huang, and Yuncai Zhou. "Research on Fast Face Retrieval Optimization Algorithm Clustering." Scientific Programming 2022. Based on Fuzzy

[10] Muhammad Irfan Khan, Muhammad Kashif Hanif and Ramzan Talib, "Caricature Face Photo Facial Attribute Similarity Generator", Hindawi Complexity, Vol. 2022.

[11] Qiang Ma and Qingshan Liu, "Facial Components-based Representation for Caricature Face Recognition", International Journal of Performability Engineering, Vol. 15, No. 3, 2019.

