

Implementation of Content-Based Image Retrieval Using Artificial Neural Networks [†]

Sarath Chandra Yenigalla ^{*}, Karumuri Srinivasa Rao and Phalguni Singh Ngangbam 

Multi-Core Architecture Computation (MAC) Lab, Department of Electronics and Communication Engineering, K L University (Deemed to be University), Vijayawada 522501, India

^{*} Correspondence: sarathcy2000@gmail.com

[†] Presented at the International Conference on “Holography meets Advanced Manufacturing”, Online, 20–22 February 2023.

Abstract: CBIR (Content Based Image Retrieval) has become a critical domain in the previous decade, owing to the rising demand for image retrieval from multimedia databases. Typically, we take low-level (colour, texture and shape) or high-level (when machine learning techniques are used) features out of the photos. In our research, we examine the CBIR system utilising three machine learning methods, namely SVM (Support Vector Machine), KNN (K Nearest Neighbours), and CNN (Convolution Neural Networks), using Corel 1K, 5K, and 10K databases, by splitting the data into 80% train data and 20% test data. Moreover, compare each algorithm’s accuracy and efficiency when a specific task of image retrieval is given to it. The final outcome of this project will provide us with a clear vision of how effective deep learning, KNN and CNN algorithms are to finish the task of image retrieval.

Keywords: Content Based Image Retrieval; Convolution Neural Networks; deep learning

1. Introduction

1.1. Image Retrieval

The explosive growth of digital images in recent years has led to the development of image retrieval systems. The viewing, searching, and retrieval of images from a sizable database of digital photographs are made possible by image retrieval systems. Adding metadata to images, such as captions, keywords, titles, or descriptions, is a common practise in traditional techniques of image retrieval. To address this challenge, extensive research has been conducted on automatic image annotation.

The design of web-based picture annotation tools has been influenced by both conventional approaches as well as the growth of social web apps and the semantic web. Image retrieval search techniques include content-based image retrieval (CBIR), image collection exploration, and image meta-search. A user-supplied query picture or user-specified image features are used by CBIR instead of written descriptions to determine how similar an image’s contents, such as textures, colours, and forms, are to the query image.

It is critical to establish the extent and nature of picture data in order to assess the degree of complexity of the image search system architecture. A search system’s predicted user traffic and the user base’s diversity are two more elements that affect design.

1.2. Content-Based Image Retrieval

The task of finding digital images in massive databases is known as the “image retrieval problem”, and content-based image retrieval (CBIR) is the practical application of machine learning algorithms to this problem. Query by image content (QBIC) is another name for CBIR [1]. Traditional concept-based methods that rely on metadata such as keywords, tags, or image descriptions are not supported by CBIR. The term “content” in



Citation: Yenigalla, S.C.; Rao, K.S.; Ngangbam, P.S. Implementation of Content-Based Image Retrieval Using Artificial Neural Networks. *Eng. Proc.* **2023**, *34*, 25. <https://doi.org/10.3390/HMAM2-14161>

Academic Editor: Vijayakumar Anand

Published: 13 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

this case can refer to colours, shapes, textures, or any additional data that can be inferred from the image itself. The search studies the contents of the image rather than its metadata.

CBIR is advantageous since searches that just employ metadata are reliant on the accuracy and completeness of the annotations, and because manually annotating photos by inserting keywords or other metadata in a big database can be difficult to execute and may not yield the required results. Similar to keyword image search, which is arbitrary and poorly defined, CBIR systems face difficulties in quantifying success. IBM built the first commercial CBIR system, QBIC (Query by Image Content), and newer network and graph-based systems have presented easy and appealing replacements to existing methods.

Due to the limits of metadata-based systems and the wide variety of applications for effective image retrieval, CBIR has increased interest. To address these needs, user-friendly interfaces and human-centred design have begun to be included in CBIR research. Many other features are now employed in CBIR systems, which were first designed to search databases using picture properties including colour, texture, and shape. However, scalability and miscategorisation problems persist with standards created to classify photos.

CBIR has been applied to a variety of applications, including satellite imagery [2], mapping, medical imaging [3], fingerprint scanning [4,5], and biodiversity information systems. The overall goal of this research article is to investigate content-based picture retrieval including its approaches, strategies, and uses. The paper also covers current research initiatives, difficulties, and CBIR's future directions.

2. Methodology

The proposed CBIR (Content-Based Image Retrieval) system with machine learning consists of an offline and online phase. In the offline phase, the system extracts feature vectors using Local Patterns methods for all images in the database, labels 60–70% of images from each class [6], and trains a machine learning classifier (e.g., SVM, KNN, and CNN) [7] to predict class names for each feature vector. In the online phase, the user inputs a query image, its feature vector is calculated using LNP (Local Neighbour Pattern), and the machine learning classifier predicts the class name. The system retrieves images from the same class in the offline phase using Euclidean distance calculations and presents the top K results to the user. Three datasets were used to test the system: Corel 1K (1000 images with 10 classes of 100 images each), Vistex (640 images of size 512×512), and Faces (40 classes, each with 10 images of size 112×92 pixels, showing variations in lighting, facial details, and expressions). Figure 1 shows the architecture of Content-Based Image Retrieval.

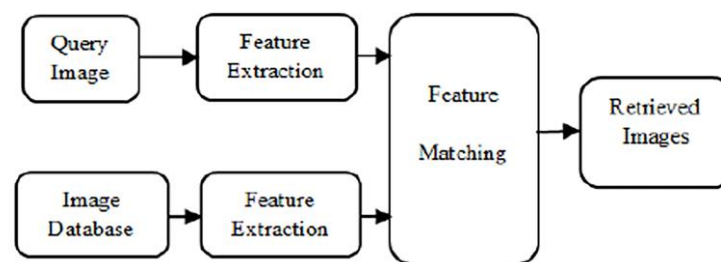


Figure 1. CBIR architecture.

2.1. Co-Occurrence Matrix Calculation

Suppose the input image has N_c and N_r pixels in the horizontal and vertical directions, respectively. Assume $Z_c = \{1, 2, \dots, N_c\}$ is a horizontal space domain and $Z_r = \{1, 2, \dots, N_r\}$ is a vertical space domain. When the direction θ and distance d are given, the matrix element $P(i, j/d, \theta)$ can be expressed by calculating the pixel logarithm of co-occurrence grey levels i and j . Assume the distance is 1, θ equals 0° , 45° , 90° , and 135° , respectively, the formulae are:

$$P(i, j/1, 0) = \#\{[(k, l), (m, n)] \in (Z_r \times Z_c) \mid k - m = 0, |l - n| = 1, f(k, l) = i, f(m, n) = j\}, \quad (1)$$

$$P(i,j/1,90) = \#\{[(k,l),(m,n)] \in (Z_r \times Z_c) \mid k - m = 1, |l - n| = 0, f(k,l) = i, f(m,n) = j\}, \quad (2)$$

$$P(i,j/1,45) = \#\{[(k,l),(m,n)] \in (Z_r \times Z_c) \mid (k - m) = 1, (l - n) = -1 \text{ or } (k - m) = -1, (l - n) = 1, f(k,l) = i, f(m,n) = j\}, \quad (3)$$

$$P(i,j/1,135) = \#\{[(k,l),(m,n)] \in (Z_r \times Z_c) \mid (k - m) = 1, (l - n) = 1 \text{ or } (k - m) = -1, (l - n) = -1, f(k,l) = i, f(m,n) = j\}, \quad (4)$$

where # represents the pixel logarithm, which generates brackets, and k,m and l,n reflect modifications of selected calculation windows.

2.2. Texture Features Extraction

Formula (5) will convert a colour image into a 256-level greyscale image.

$$Y = 0.114 \times B + 0.587 \times G + 0.29 \times R, \quad (5)$$

where Y is the grey-scale value. R, G, and B represent red, green, and blue component values, respectively. Because the grey scale is 256, the corresponding co-occurrence matrix is 256×256 . The grey scale of the initial image will be compressed to reduce calculations before the co-occurrence matrix is formed. A total of 16 compression levels were chosen in the paper to improve the texture feature extracting speed. Four co-occurrence matrices are formed according to Formula (3) to Formula (6) in four directions. The four texture parameters are calculated: capacity, entropy, moment of inertia, and relevance.

For an image li and its corresponding feature vector $Hi = [hi, 1, hi, 2, \dots, hi, N]$, assume the feature component value satisfies a Gaussian distribution. The Gaussian normalization approach is used to implement internal normalization in order to make each feature of the same weight.

$$hi,j' = hi,j - mj\sigma_j, \quad (6)$$

where m_j is mean and σ_j is the standard deviation. hi,j will be unitized on a range $[-1,1]$. The texture feature of each image is calculated according to the above steps. The texture values are compared by Euclidean distance, the closer the distance the higher the similarity.

3. Results and Discussion

Using average precision, the proposed retrieval system's effectiveness is assessed for each query. Equation (7) can be used to obtain the area beneath each query's precision, where the precision is the proportion of relevant images to all images retrieved.

$$\text{Precision} = \frac{\text{relevant images} - \text{retrieved images}}{\text{retrieved images}} \quad (7)$$

In this Proposed method, three different types of databases area used, and three types of techniques are used. Here, the query images (k) are fixed at 8. Figures 2–4 show the retrieved images in an animal database and the accuracy of the image with respect to the searched image is shown below the respective image.

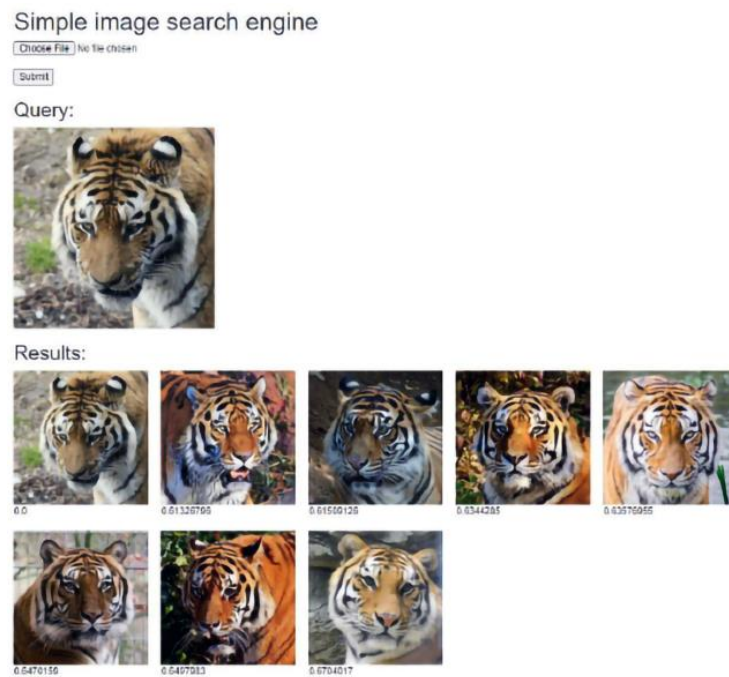


Figure 2. Retrieved animal images.



Figure 3. Retrieved bus images.

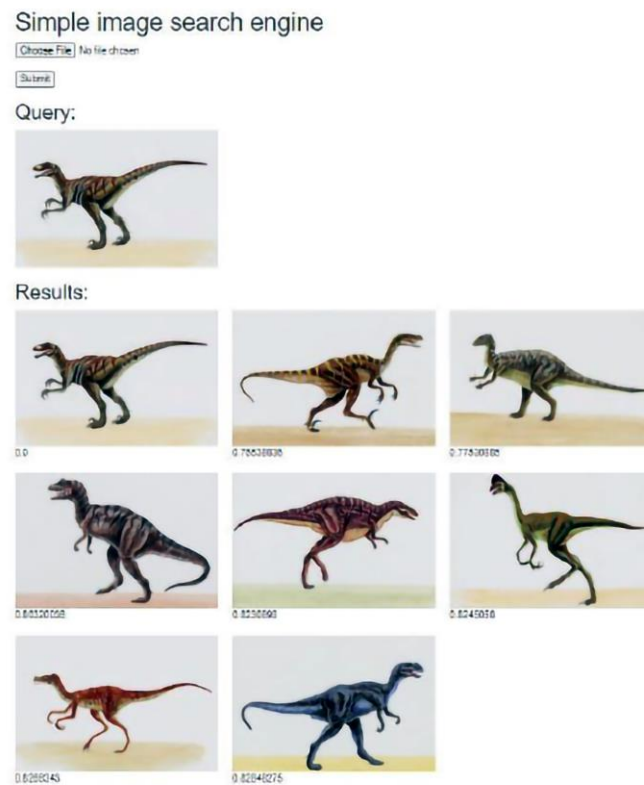


Figure 4. Retrieved dinosaur images.

Figures 5–7 show the average accuracy results of pixel and histogram accuracy both with respect to the algorithm used.

```
[INFO] describing images...
[INFO] evaluating raw pixel accuracy...
[INFO] raw pixel accuracy: 50.00%
[INFO] evaluating histogram accuracy...
[INFO] histogram accuracy: 80.00%
>>> |
```

Figure 5. CNN Accuracy Output.

```
[INFO] describing images...
[INFO] evaluating raw pixel accuracy...
[INFO] raw pixel accuracy: 60.00%
[INFO] evaluating histogram accuracy...
[INFO] histogram accuracy: 80.00%
```

Figure 6. KNN Accuracy Output.

```
[INFO] describing images...
[INFO] evaluating raw pixel accuracy...
[INFO] raw pixel accuracy: 80.00%
[INFO] evaluating histogram accuracy...
[INFO] histogram accuracy: 80.00%
```

Figure 7. SVM Accuracy Output.

Table 1 shows the experimental results of images retrieved depending upon the accuracy of the algorithm in which the CNN technique has the least 50% and KNN has 60% and SVM has 80%.

Table 1. Accuracy of Algorithms using 3 Databases.

Algorithm	Corel 1K	Animal	Dinosaur
KNN	60%	60%	60%
CNN	50%	50%	50%
SVM	80%	80%	80%

4. Conclusions

Following a review of prior CBIR efforts, the paper investigated the low-level aspects of CBIR colour and texture extraction. The paper created a CBIR system using fused characteristics of colour and texture after testing the three distinct types of algorithms (KNN, CNN, and SVM) with various databases (Corel 1K, Animal, and Dinosaur). By entering a query image, similar photographs can be correctly and quickly retrieved. Moreover, the above works discuss the calculation of the accuracy of each algorithm with all the three databases and conclude that the CNN algorithm has 50% accuracy with all three databases, the KNN algorithm has 60% accuracy with all three databases, and the SVM algorithm has 80% accuracy with all the three databases, and so finally we came to a conclusion that SVM is the best suit algorithm among the three algorithms with an accuracy of 80%.

In the future, more low-level features will be combined to strengthen the system, such as spatial position and shape features. The other two key components of the CBIR system are the image feature matching approach and semantic-based image retrieval.

Author Contributions: S.C.Y.: conceptualization, methodology, software, investigation, writing—original draft preparation, writing—review and editing, formal analysis. K.S.R.: project administration, resources, investigation, and methodology. P.S.N.: supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No data was used for the research described in the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Sharma, S. Use of Artificial Intelligence Algorithm for Content-Based Image Retrieval System. *Int. J. Adv. Res. Ideas Innov. Technol.* **2018**, *4*, 680–684.
- Ferran, A.; Bernabe, S.; Rodriguez, P.; Plaza, A. A Web-Based System for Classification of Remote Sensing Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 1934–1948. [[CrossRef](#)]
- Ramos, J.; Kockelkorn, T.; Ramos, I.; Ramos, R.; Grutters, J.; Viergever, M.; van Ginneken, B.; Campilho, A. Content-Based Image Retrieval by Metric Learning From Radiology Reports: Application to Interstitial Lung Diseases. *IEEE J. Biomed. Health Inform.* **2016**, *20*, 281–292. [[CrossRef](#)] [[PubMed](#)]
- Zegarra, J.M.; Leite, N.; da Silva Torres, R. Wavelet-based fingerprint image retrieval. *J. Comput. Appl. Math.* **2009**, *227*, 294–307. [[CrossRef](#)]
- Gavrielides, M.; Sikudova, E.; Pitas, I. Color-based descriptors for image fingerprinting. *IEEE Trans. Multimed.* **2006**, *8*, 740–748. [[CrossRef](#)]
- Alrahhhal, M.; Supreethi, K.P. Content-Based Image Retrieval using Local Patterns and Supervised Machine Learning Techniques. In Proceedings of the 2019 Amity International Conference on Artificial Intelligence (AICAI), Dubai, United Arab Emirates, 4–6 February 2019; pp. 118–124. [[CrossRef](#)]
- Wiggers, K.L.; Britto, A.S.; Heutte, L.; Koerich, A.L.; Oliveira, L.E.S. Document Image Retrieval Using Deep Features. In Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 8–13 July 2018; pp. 1–8. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.