proceedings

Type of the Paper (Proceedings, Abstract, Extended Abstract, Editorial, etc.)

IMPLEMENTATION OF CONTENT-BASED IMAGE RETRIEVAL USING ARTIFICIAL NEURAL NETWORKS †

Sarath Chandra Yenigalla [1], Srinivas Rao K [2] and Phalguni Singh Ngangbam[2]

[1] Multi-Core Architecture Computation (MAC) Lab, Department of ECE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, 522501, India

[2] Professor, Department of Electronics and Communication Engineering, KL University (Deemed to be University), Green Fields, Vaddeswaram, Andhra Pradesh, India

† Presented at the Holography meets advanced manufacturing, India, 15 February 2023.

Abstract: CBIR (Content Based Image Retrieval) is a crucial domain, mainly in the last decade, due to the increased need for image retrieval from the multimedia database. In general, we extract low-level (colour, texture, and shape) or high-level features (when we include machine learning techniques) from the images. In our work, we compare the CBIR system using three algorithms based on machine learning, i.e., SVM (Support Vector Machine), KNN (K Nearest Neighbors) and CNN (Convolution Neural Networks) algorithms using Corel 1K,5K,10K databases, by dividing the data into 80% train data and 20 % test data. Also, 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.

Citation: To be added by editorial staff during production.

Academic Editor: Firstname Lastname

Published: date

1. Introduction

1.1. Image Retrieval

The explosive growth of digital images in recent years has led to the development of image retrieval systems. Image retrieval systems enable the browsing, searching, and retrieval of images from a large database of digital images. Traditional methods of image retrieval include the addition of metadata such as captioning, keywords, titles, or descriptions to images, which can be used for retrieval. To address this challenge, extensive research has been done on automatic image annotation.

In addition to traditional methods, the development of social web applications and the semantic web has inspired the creation of web-based image annotation tools. There are several search methods for image retrieval, including image meta search, content-based image retrieval (CBIR), and image collection exploration. CBIR utilizes computer vision to retrieve images based on similarities in their contents, such as textures, colors, and shapes, to a user-supplied query image or user-specified image features, thereby avoiding the use of textual descriptions. It is important to understand the scope and nature of image data to determine the complexity of the image search system design. The design is also influenced by factors such as the diversity of the user base and expected user traffic for a search system. Along this dimension, search data can be classified such as archives, domain-specific collections, enterprise collections, personal collections, and web collections.

1.2. Content-Based Image Retrieval

Content-based image retrieval (CBIR) is the application of computer vision techniques to the image retrieval problem, which is the challenge of searching for digital images in large databases. CBIR is also known as query by image content (QBIC)[1]. CBIR is opposed to traditional concept-based approaches that rely on metadata such as keywords, tags, or descriptions associated with the image. The search analyses the contents of the image rather than its metadata, and the term "content" in this context may refer to colours, shapes, textures, or any other information that can be derived from the image itself.

CBIR is desirable because searches that rely purely on metadata are dependent on annotation quality and completeness, and humans manually annotating images by entering keywords or metadata in a large database can be time-consuming and may not capture the desired keywords to describe the image. However, CBIR systems have challenges in defining success, just like keyword image search, which is subjective and not well-defined. QBIC (Query by Image Content) was the earliest commercial CBIR system developed by IBM, and recent network and graph-based approaches have presented simple and attractive alternatives to existing methods. As the interest in CBIR has grown due to the limitations inherent in metadata-based systems and the large range of possible uses for efficient image retrieval, efforts in the CBIR field have started to include human-centred design that meets the needs of users, including user-friendly interfaces. The first CBIR systems were developed to search databases based on image colour, texture, and shape properties, but now many other attributes are used. Nonetheless, standards developed to categorize images still face scaling and miscategorization issues.

CBIR has been used in several fields, such as satellite images[2], remote sensing, medical imaging[3], fingerprint scanning ([4], [5]) and biodiversity information systems. Overall, this research paper aims to explore content-based image retrieval, including its methods, techniques, and applications. The paper will also discuss current research efforts, challenges, and future directions for CBIR.

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, 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 512x512), and Faces (40 classes, each with 10 images of size 112x92 pixels, showing variations in lighting, facial details, and expressions).

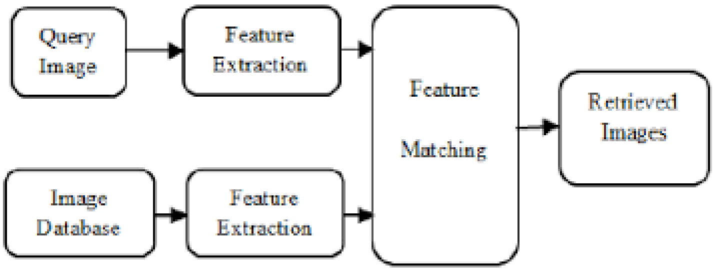

Figure 1. CBIR Architecture

2.1 Calculating the Co-Occurrence Matrix

Suppose the input image has Nc and Nr pixels in the horizontal and vertical directions respectively. Assume Zc={1,2,…, Nc} is a horizontal space domain and Zr={1,2,…, Nr} is a vertical space domain. When the direction θ and distance d are given, the matrix element P(i,j/d,θ) 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°, 135° respectively, the formulae are:

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

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

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

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

where, k,m and l,n represent changes of selected calculating windows, and # represents pixel logarithm which establishes brackets.

2.2 Extracting Texture Features

The colour image will be converted to a grey-scale image by formula 7, the number of the grey scale is 256.

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

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. 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,…,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.

$$h_{i,j}'=h_{i,j}-m_j\sigma_j \quad (6)$$

where mj is mean and σj 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

The performance of the proposed retrieval system is evaluated on each query using average precision. The area under precision for each query can be calculated by Equation (2), where the precision is the ratio of the number of relevant images retrieved to the total number of images retrieved.

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

In this Proposed method, three different types of databases area used and three types of techniques are used. Here the query images (k) is fixed at 8. Figure 2,3,4 shows 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.

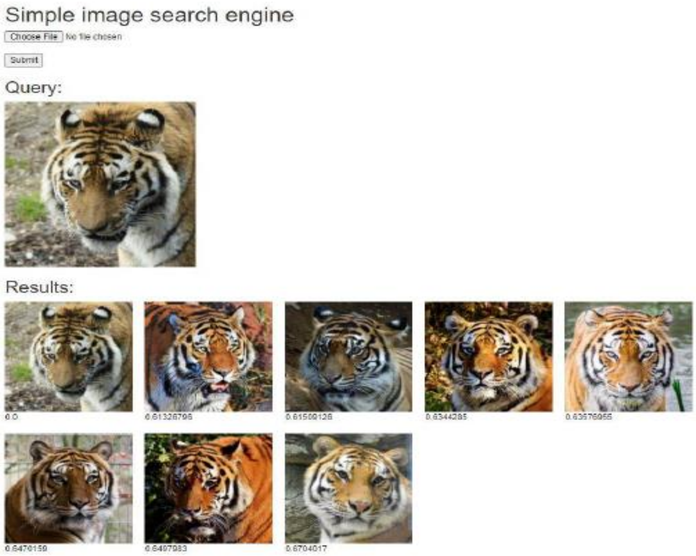

Figure 2. Retrieved animal images

Figure 3. Retrieved bus images

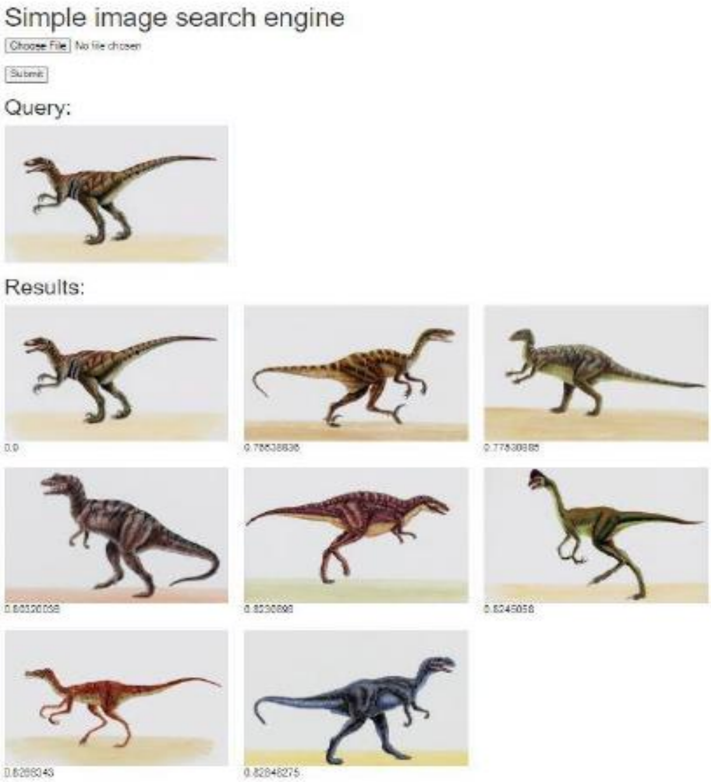

Figure 4. Retrieved Dinosaur images

```
[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

Figure 5,6,7 shows the average accuracy results of pixel and histogram accuracy both with respect to the algorithm used.

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%

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%.

4. Conclusion

After a survey of the previous CBIR works, the paper explored the low-level features of colour and texture extraction for CBIR. After comparing the three types of algorithms (KNN, CNN, SVM) with 3 different databases (Corel 1k, Animal, dinosaur), the paper implemented a CBIR system using colour and texture fused features. Similar images can be retrieved quickly and accurately by inputting a query image. Also, the above works discuss the calculation of the accuracy of each algorithm with all the 3 databases and conclude that the CNN algorithm has 50% accuracy with all three databases, the KNN algorithm has 60% accuracy with all three databases, 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%.

More low-level features such as shape and spatial location features etc. will be fused to make the system more robust in the future. The image feature matching method and semantic-based image retrieval are the other two important aspects of the CBIR system.

References

1. S. Sharma, "Use of Artificial Intelligence Algorithm for Content-Based Image Retrieval System", International Journal of Advance Research, Ideas and Innovations in Technology, vol. 4, no., pp. 680-684, 2018.
2. A. Ferran, S. Bernabe, P. Rodriguez and A. Plaza, "A Web-Based System for Classification of Remote Sensing Data", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 6, no. 4, pp. 1934-1948, 2013.
3. J. Ramos, T. Kockelkorn, I. Ramos, R. Ramos, J. Grutters, M.Viergever, B. van Ginneken and A. Campilho, "Content-Based Image Retrieval by Metric Learning From Radiology Reports: Application to Interstitial Lung Diseases", IEEE Journal of Biomedical and Health Informatics, vol. 20, no. 1, pp. 281-292,2016.
4. J. Montoya Zegarra, N. Leite and R. da Silva Torres, "Wavelet-based fingerprint image retrieval", Journal of Computational and Applied Mathematics, vol. 227, no. 2, pp. 294-307, 2009.
5. M. Gavrielides, E. Sikudova and I. Pitas, "Color-based descriptors for image fingerprinting", IEEE Transactions on Multimedia, vol. 8, no. 4, pp. 740-748, 2006.
6. M. Alrahhal and K. P. Supreethi, "Content-Based Image Retrieval using Local Patterns and Supervised Machine Learning Techniques," *2019 Amity International Conference on Artificial Intelligence (AICAI)*, Dubai, United Arab Emirates, 2019, pp. 118-124, doi: 10.1109/AICAI.2019.8701255.
7. K. L. Wiggers, A. S. Britto, L. Heutte, A. L. Koerich and L. E. S. Oliveira, "Document Image Retrieval Using Deep Features," *2018 International Joint Conference on Neural Networks (IJCNN)*, Rio de Janeiro, Brazil, 2018, pp. 1-8, doi: 10.1109/IJCNN.2018.8489722.
8. https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/
9. https://medium.com/analytics-vidhya/convolutional-neural-networks-cnn-explained-step-by-step-69137a54e5e7
10. https://www.tutorialspoint.com/machine_learning_with_python/machine_learning_with_python_knn_algorithm_finding_nearest_neighbors.htm