International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT239024

An Efficient CNN-Based Method for Content-Based Image Retrieval

A Mukul Kumar Patro*, Dr. J Bhuvana
School of Computer Science and IT, Jain University, Bangalore, Karnataka, India

ARTICLEINFO

Article History:

Accepted: 01 March 2023
Published: 15 March 2023

Publication Issue
Volume 10, Issue 2
March-April-2023

Page Number
72-77

ABSTRACT

Image recovery has been one of the most fascinating and active study fields in the field of computer vision. The use of content-based image retrieval (CBIR) systems allows for the automatic indexing, searching, retrieval, and exploration of picture datasets. Important characteristics in content-based picture retrieval systems include colour - texture elements. As a result, content-based image retrieval (CBIR) is attractive as a source of precise and speedy retrieval in the modern era. The (CBIR) system uses a feature-based approach to retrieve images from image databases. Low grade characteristics and high grade characteristics are the two categories that image features fall under. Low level aspects of an image include colour, texture, and shape, whereas high level features define the image's semantic content. CBIR is a rapidly developing technology, and as datasets grow as a result of recent advancements in multimedia, it is crucial to enhance this technology to suit user needs.

Keywords: CBIR, Image Processing, Feature Extraction, Image Retrieval

I. INTRODUCTION

The visual components of the images specified in the approach of short level characteristics, such as colour, texture, shape, and spatial locations, are used by a CBIR system to represent the images in the databases. When a sample image or design is available for the system to use as input, the structure retrieves similar images. Low level characteristics in a typical CBIR system image, such as colour, texture, size, and spatial placements, are represented using a multidimensional component direction.

The restoration procedure begins when a user queries the system using an instance picture. To convert the input images into the internal representation of feature vector, the feature extraction method that was abandoned to structure the feature databank is employed. The distance between the feature routes of the query picture and those of the target picture in the feature database is determined using the comparison measure. Finally, an indexing system is used for retrieval, making it easier to search the image collection effectively. Recently, user relevance feedback has also been incorporated into the development of the retrieval process to provide retrieval effects that are more perceptually and semantically important. Here, we discuss these core methods for CBIR. [1]

APPLICATIONS

- Advertising and marketing
- Medical imaging
- Surveillance systems
- Multimedia Retrieval

Purpose of CBIR

The Content - based image retrieval system is dependent on low level visual properties like colour, texture, and shape. Many potential low level features that are mentioned in the literature. Low level features are subtracted from database image files and stored in a feature database. The query image's low level features are similarly extracted, and they are compared to the database image's low-level characteristics using the distance measure. [2]

Technology-driven methods outperform other approaches in the field of CBIR systems. According to user needs, extensive research is being done in the field of CBIR systems. It is a difficult task to model complicated human behaviour. [3]

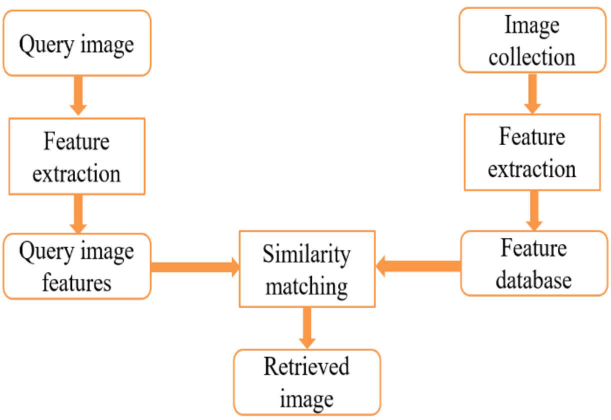

Figure 1: CBIR system

II. LITERATURE SURVEY

Deep learning Method for content-based image retrieval using Distance measurement.

Convolution neural networks (CNN) with deep learning were employed in the exploration work conducted in [4] by Sirisha Kopparthi, Dr. N. K. Kameswara Rao, and they performed exceptionally well in a variety of image processing procedures. It is possible to employ a single CNN structure and CNN-based algorithms to extract image features from the final layer and locating related photos. The CBIR consists of effective similarity comparison and learning feature extraction (CBIR). Both feature extraction and similarity measures are crucial in CBIR. The tests are run on two datasets, including the UC Merced Land Use Dataset. Utilising a model that has already been trained and customised for the reclamation process using millions of photos as training data. It uses pre-trained CNN models to produce feature descriptors of pictures to aid in the retrieval procedure. By employing transfer learning and retrieving feature vectors using multiple similarity measures, this technique deals with attribute extraction from the two completely connected layers included in the VGG-16 and VGG19 network.

Picture Retrieval Using a Mixture of Color Moment and Color Histogram

S. Mangijao Singh and K. Hemachandran used a cutting-edge method for CBIR in their research, which proposes the use of colour histograms and colour moments of images [5]. The rotation and translation invariance of the colour histogram is an advantage, but the lack of spatial information is a drawback. This study proposes a content-based picture retrieval strategy that combines colour histogram and colour moment feature vectors to increase retrieval accuracy. By dividing the image horizontally into three equal, nonoverlapping parts, a minimum quantity of spatial information is encoded in the colour index for colour instant in order to increase the discriminating ability of colour indexing approaches.

Picture Retrieval Using Deep Learning for Content-Based Search

The author of this thesis was motivated to work on this project because Anshuman Vikram Singh's exploration work in [6] employed deep learning techniques, particularly Convolutional Neural Networks (CNN), in working computer vision operations. This thesis aims to solve the CBIR problem by using a data source of annotated images. Only 3000 photos from 41 orders and 8 classes were used in his work. Further dataset expansion and the addition of additional classes that are similar to man, person, aeroplane, etc. will help the system become more universal and effective. It can be observed that the confirmation error rate and testing error rate for each marker was relatively low. The stylish test and confirmation error rates are achieved on replication 3, 6 or 9. After every 100 duplicates, training continues for 500 to verify that the failure rate has not changed. Once it hits 500 at a steady rate, it pauses and provides the stylish failure rates.

Fast content based image retrieval using Convolutional Neural Network and Hash function

The research done by Domonkos Varga and Tamas Sziranyi in [7] has inspired them to investigate how Convolution Neural Network (CNN)-like methods of deep learning operate in their own environments. The increasing growth of internet photos has drawn a lot of interest to content-based image retrieval. Their key contribution is a novel end-to-end supervised learning framework that simultaneously learns probabilistically based semantic position similarity and point position similarity. The primary benefit of the new mincing strategy is that it can greatly lower the computational cost of reclamation at the most cutting-edge effectiveness position.

III.RESEARCH METHODOLOGY

There are numerous methods for achieving convolutional neural network-based content-based image retrieval (CBIR). Transfer learning, deep metric learning, multi-scale, attention-based, and adversarial training are some examples of methodologies that have been used in previous studies.

A pre-trained CNN model is adjusted via transfer learning on a smaller dataset. This methodology consists of selecting a pre-trained CNN model, freezing its layers, adding new layers, and training it on the new dataset. [8]

Deep metric learning is based on learning a distance metric that maps images into a high-dimensional feature space where the distance between images is a measure of their similarity. This methodology involves training a CNN model to learn a metric space that maps images into a feature space, where the distance between images reflects their semantic similarity. [9]

Multi-scale methodology extracts features from images at different resolutions. According to this approach, a CNN model is created that employs many convolutional layers with various kernel sizes to extract information at various scales. [10]

Attention-based methodology involves an attention mechanism to weight the importance of different regions of an image. Using an attention mechanism, this process entails creating a CNN model that extracts features according to the relevance of various sections of the image. [11]

IV.PROPOSED METHODOLOGY

There are multiple phases involved in Content-Based Image Retrieval utilising Convolutional Neural Networks, which are frequently employed in various

research investigations. These steps include data collection, data pre-processing, feature extraction, similarity measurement, user interface design, evaluation, and optimization.
The first step involves collecting a large dataset of images for training and testing the CNN model. The dataset must be diverse and representative of the target domain. The images are then pre-processed by resizing them to a standard size and normalizing the pixel values.

The following stage is to extract high-level features from the images using a previously trained CNN model, such as VGG, ResNet, or Inception. The output of the CNN model can be flattened and used as a feature vector. This step is known as feature extraction.

After feature extraction, the similarity between images can be computed using a distance metric such as Euclidean distance, cosine similarity, or correlation coefficient. The vectors of features of the query picture and the pictures in the database can be compared to accomplish this.

Additionally, the system needs to have an intuitive interface that lets users submit a query image and shows the database's most comparable pictures. The user interface design ought to be simple and straightforward. Metrics like recall, F1-score, and precision can be used to assess the system's success.

Finally, the whole thing can be made more accurate and effective by changing the CNN model and changing the system settings. This step is crucial for ensuring that the system performs well on new and unseen data.

V. RESULTS AND DISCUSSION

In the results and discussion section we have described some of the outcomes of implemented project which include some output figures and their respective explanation for the same.

Figure 2: Loading image

The above figure represent use of custom function load_image () to load an image file from a specified directory and preprocess it for use with a CNN model. The loaded image is stored in the variable img and the preprocessed image data is stored in x. The function returns the image and the preprocessed data as a tuple.

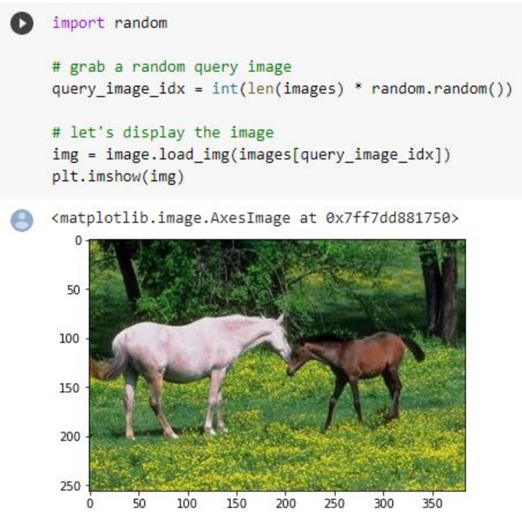

Figure 3: Importing random picture

This code randomly selects an image from a dataset of images and displays it using Matplotlib. The purpose of this code is to provide a visual representation of the query image that will be used for image retrieval. The query image will be selected randomly in order to ensure that the retrieval system is able to retrieve images from the entire dataset rather than just a specific subset.

The use of a random query image is a common practice in content-based image retrieval research, as it allows for the evaluation of the retrieval system's performance on a diverse set of images. By selecting a random image, the evaluation is not biased towards any specific image or category, which can lead to more robust and accurate results. Additionally, displaying the query image allows for a better understanding of the image retrieval system's behavior and can aid in identifying any potential issues or limitations.

```
thumbs = []
for idx in idx_closest:
    img = image.load_img(images[idx])
    img = img.resize((int(img.width * 100 / img.height), 100))
    thumbs.append(img)

# concatenate the images into a single image
concat_image = np.concatenate([np.asarray(t) for t in thumbs], axis=1)

# show the image
plt.figure(figsize = (16,12))
plt.imshow(concat_image)
```

<matplotlib.image.AxesImage at 0x7f7f385f9c50>

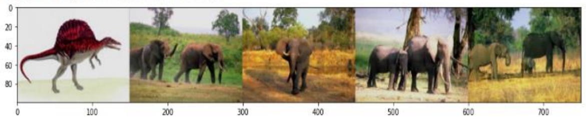

Figure 4: Creating a list of thumbnail

The following code generates a list of thumbnail pictures for the images that have the greatest similarities identical to the query image. It begins by looping through the indices of the images which are most similar to the query image, then for each index it loads the corresponding image and resizes it to a fixed height while maintaining the aspect ratio. These thumbnail images are then appended to the list "thumbs".

Next, the code concatenates the images in the "thumbs" list horizontally to create a single image of all the closest images. This concatenated image is stored in the variable "concat_image". Finally, the concatenated image is displayed using the Matplotlib library. The figure size is set to 16x12 and the image is displayed using the imshow() function.

Figure 5: query image with output

The preceding illustration demonstrates how the code takes the index of a query image, uses the suggested approach to find the images that are the most similar to the query image, and then combines the query image with the related pictures into one image for viewing purposes. The resulting images are then displayed using the **matplotlib** library.

The intent of the code is to demonstrate how the suggested method obtains images that are comparable to a given search query, allowing researchers to evaluate the method's performance using specific evaluation metrics. The performance evaluation can be based on metrics such as precision, recall, or mean average precision (mAP), depending on the research objectives.

VI. CONCLUSION

In conclusion, significant improvements in the area of image retrieval have been shown as a result of the incorporation of convolutional neural networks (CNNs) into content-based image retrieval (CBIR). With a growing number of digital images, CBIR has become a critical task and the utilization of deep learning methods has enhanced the accuracy and speed of image retrieval systems. The use of CNNs in CBIR has been demonstrated to effectively identify and analyse high-level image features, leading to improved image retrieval performance. It is expected that future developments, including the incorporation of other advanced techniques and further improvement of CNN models, will further advance the CBIR process. Overall, the use of CNNs in CBIR offers a promising future and potential for further growth in the field.

VII.REFERENCES

[1] Schaefer, Gerald. "An introduction to content-based image retrieval." Eighth International Conference on Digital Information Management (ICDIM 2013).

[2] Dharani, T., and I. Laurence Aroquiaraj. "A survey on content based image retrieval." Pattern Recognition, Informatics and Mobile Engineering (PRIME), 2013 International Conference on. IEEE, 2013.

[3] Kaur, Sukhman, and Rupinder Kaur. "Survey of Content Based Image Retrieval Architecture, Advantages and Disadvantages." International Journal of Research in Electronics and Computer Engineering, vol. 4, no. 3, July 2016, pp. 108–110, https://doi.org/10.13140/RG.2.2.22093.03043.

[4] Sirisha Kopparthi, Dr. N. K. Kameswara Rao, "Content based Image Retrieval using Deep Learning Technique with Distance Measures", Science, Technology & Human Values 9(12):251-261, 2020.

[5] S. Mangijao Singh, K. Hemachandran, "Image Retrieval based on the Combination of Color Histogram and Color Moment", International Journal of Computer Applications, Volume 58– No.3, November 2012.

[6] Anshuman Vikram Singh, "Content-Based Image Retrieval Using Deep Learning", Thesis. Rochester Institute of Technology, RIT Scholar Works, 2015.

[7] Domonkos Varga, Tamas Sziranyi, "Fast Content-Based image retrieval using Convolutional Neural Network and Hash Function", IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2016.

[8] 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.

[9] S. Chopra, R. Hadsell, and Y. LeCun, "Learning a similarity metric discriminatively, with application to face verification," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2005.

[10] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015.

[11] J. Xu, R. Zhao, and Y. Wang, "Attentional neural networks for image captioning," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015.

Cite this article as :

A Mukul Kumar Patro, Dr. J Bhuvana, "An Efficient CNN-Based Method for Content-Based Image Retrieval ", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 9, Issue 2, pp.72-77, March-April-2023. Available at doi : https://doi.org/10.32628/CSEIT239024

Journal URL : https://ijsrcseit.com/CSEIT239024