International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT228418

# Content-Based Image Retrieval Using Deep Learning

**Dr. J C Karur*[1], Asma Hebbal[2], Dr. Jagadeesh Pujari[3]**

*[1] Professor, Department of Computer Science and Engineering, SDM College of Engineering and Technology, Dharwad, Karnataka, India

[2]Mtech Student Department of Computer Science and Engineering, SDM College of Engineering and Technology, Dharwad, Karnataka, India

[3] HOD Department of Information Science and Engineering, SDM College of Engineering and Technology, Dharwad, Karnataka, India

**Article Info**

**Publication Issue :**
Volume 8, Issue 4
July-August-2022

**Page Number :** 122-128

**Article History**
Accepted: 10 July 2022
Published: 22 July 2022

## ABSTRACT

The most prevalent and well-used method for obtaining images from huge, unlabelled image datasets is content-based image retrieval. Convolutional Neural Networks are pre-trained deep neural networks which can generate and extract accurate features from image databases. These CNN models have been trained using large databases with thousands of classes that include a huge number of images, making it simple to use their information. Based on characteristics retrieved using the pre-trained CNN models, we created CBIR systems in the work. These pre-trained CNN models VGG16, and MobileNet have been employed in this instance to extract sets of features that are afterward saved independently and used for image retrieval.

**Keywords :** Convolutional Neural Networks, VGG16, MobileNet, Cosine Similarity.

## I. INTRODUCTION

Digital photography and other multimedia elements have dramatically increased as a result of the quick development of digital computers and various smart gadgets, as well as the enormous and ongoing growth of various media channels. Multimedia, especially digital photography, is widely used in many fields, including medicine, satellite measurements, remote sensing, forensic analysis, and even digital evidence. To use these images in a certain area or application, it is necessary to retrieve them from the numerous sources from where they were originally stored. Another contemporary and successful method for returning images to various image storage, as well as on the internet, is Content-Based Image Retrieval (CBIR). CBIR is the process of recovering an image by extracting useful information by using low-level characteristics or content such as color, texture, shape, or any other level of features. Because the extracted features will be used as numerical values in comparing similarities between the query image and other images in the image database, they are critical to the success and use of any Content-Based Image Retrieval system (CBIR). Machine learning techniques and methodology have improved tremendously during the last few decades. These approaches have been used effectively in many disciplines, including classification, clustering, information retrieval, and image retrieval. The availability of massive amounts of pre-classified data, as well as the enhanced processing capability of modern computers, are the key motivators for the development of machine learning and deep learning approaches. A convolutional neural network is a set of nonlinear transformation methods that may learn from incoming data (CNN). It splits the input data into square cells before performing convolution operations on the pixel values. Deep convolutional neural networks are used in many digital images processing applications, including object identification, picture clustering, and image classification. Convolutional neural networks, on the other hand, need a large quantity of data, computing power, and processing time. Pre-trained models are ones that have recently been constructed, have high accuracy, and have achieved excellent outcomes in a variety of research domains. Using the data set, they were trained for a specific purpose. VGG16, ResNet-50, MobileNet, GoogleNet, and other pre-trained CNN models have been utilized to handle a wide range of problems, including computer vision, medical image classification, and natural language processing. In this paper, we present the Content-Based Image Retrieval approach, which is based on the three most prevalent and popular types of networks for extracting features and then applying cosine similarity to recover images through their content. Section II presents and explains our suggested technique in detail. Section III presents the project's outcomes and findings. Section IV concludes and discusses our study.

## II. METHODS AND MATERIAL

According to Figure 1, the strategy suggested in this study is divided into two phases: an offline process and an online process. For feature extraction, the offline procedure makes use of a pre-trained CNN model, while the online stage is in charge of modifying and receiving the user query. The pre-trained deep CNN model is made up of multiple layers that carry out numerous subsampling and convolutional operations as well as the learning process step by step. Feature extraction is performed using pre-trained deep CNN models VGG16, and MobileNet, and the feature vector is saved in the feature database for further use in similarity computation. The most important stage is the online one, in which the top comparable images for a user are obtained, and the extracted characteristics established in the previous stage are utilized to compute matching and similarity instead of the actual image

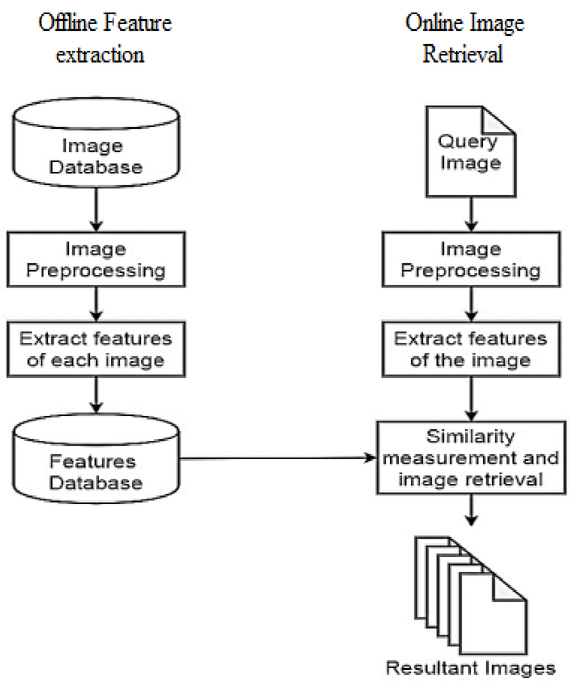



Fig1: Framework of the content-based image retrieval

A. *VGG16 Pre-Trained CNN Model for Feature Extraction*

The next pre-trained CNN model employed in the study is VGG16, a 16-layer convolutional neural network built by Karen Simonyan at the University of Oxford. This model has 16 layers and requires a 224*224*3 input picture. After each image is adjusted to the required size, the process in the VGG16 model runs over 13 convolutional processing block components with unique resampling and scaling techniques. Finally, 4096 features are retrieved from the dense layer and saved in a separate database for use in the next online similarity and rating procedure.

B. *MobileNet Pre-Trained CNN Model for Feature Extraction*

TensorFlow's first mobile computer vision model, MobileNet, is designed for use in mobile applications. Depthwise Separable Convolutions are used by MobileNet. It drastically decreases the number of parameters when compared to a network with traditional convolutions of the same depth in the nets. The result is lightweight deep neural networks. The MobileNet model, which accepts input pictures with dimensions of 224*224*3, has 27 convolution layers, including 13 depthwise convolution layers, one average pool layer, one fully connected layer, and one softmax layer. A total of 1000 characteristics are then retrieved and put into a separate database to be utilized by the following online similarity and rating procedure.

C. *Similarity Measure*

The group characteristics are recovered from deep CNN models that have previously been trained and stored in databases to conduct the similarity measures, which are considered a critical stage in the retrieval process. To compare the similarity of two strings, data points, probability distributions, or collections, similarity measurements or metrics are utilized. Cosine similarity was employed as the similarity metric in this study. In text analysis, the cosine similarity measure is frequently used to analyze document similarity. For computing cosine similarity, we use the following formula:

$$\text{Similarity} = (A.B) / (||A||.||B||)$$

where the vectors A and B are:

- A.B is the dot product of A and B: It is calculated as the sum of A and B's element-wise products.
- ||A|| is A's L2 norm: It is calculated as the square root of the sum of squares of the vector A's elements.

## III. RESULTS AND DISCUSSION

*A. Images Datasets*

The Corel-1000 database, which has 1000 photos divided into 10 classes with 100 images each, serves as the basis for this evaluation's dataset. Each image in this database has either 256 *384 or 384 *256-pixel dimensions. The graphic below displays image samples from each semantic class in the Corel-1000 image database. Samples from datasets are shown in figure 2.

*B. Performance Evaluation*

Two performance metrics are employed in this context: precision and accuracy. These two measures might be used to evaluate any retrieval model, particularly the CBIR model. The recovered yield provides a wealth of data that helps assess the framework's success. Images should be divided into relevant and irrelevant categories, where relevant categories include images that satisfy the needs of the user and irrelevant categories include images that don't. What percentage of identification is genuinely correct is how precision is defined.

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

Where True positive is abbreviated as TP and False Positive is abbreviated as FP.

A True Positive result occurs when the model correctly predicts the positive class, whereas a True Negative outcome occurs when the model correctly predicts the negative class. Accuracy is determined by dividing the number of correct predictions by the total number of forecasts.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FN+FP}$$

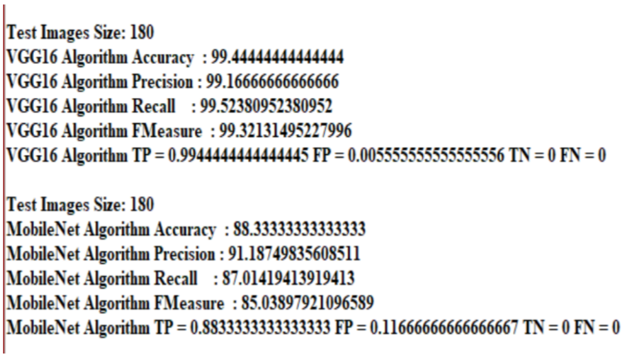
Precision is sometimes referred to as positive predictive value or specificity, whereas recall is referred to as sensitivity. Both are measured and utilized to determine the predictive model's performance.

## IV. RESULTS AND DISCUSSION

Image retrieval algorithms are developed in this work, and the results and conclusions are examined. The findings of each experiment are described, and the figures that reflect the results are provided below. The article employs the pre-trained models VGG16, and MobileNet. The corel1k dataset has 1000 images organized into10 classes each with 100 images.


Fig2: Samples Images from Corel-1K Dataset

Figure 3 shows the accuracy precision, and recall of all three models VGG16 and MobileNet. Figure 4 shows the performance graph of all three models. From figure 3 and figure 4, we conclude that the accuracy of VGG16 is higher than that of MobileNet.



Test Images Size: 180
VGG16 Algorithm Accuracy : 99.44444444444444
VGG16 Algorithm Precision : 99.16666666666666
VGG16 Algorithm Recall : 99.52380952380952
VGG16 Algorithm FMeasure : 99.32131495227996
VGG16 Algorithm TP = 0.9944444444444445 FP = 0.005555555555555556 TN = 0 FN = 0

Test Images Size: 180
MobileNet Algorithm Accuracy : 88.33333333333333
MobileNet Algorithm Precision : 91.18749835608511
MobileNet Algorithm Recall : 87.01419413919413
MobileNet Algorithm FMeasure : 85.03897921096589
MobileNet Algorithm TP = 0.8833333333333333 FP = 0.11666666666666667 TN = 0 FN = 0

Fig 3 : Accuracy, precision, recall of the VGG16, MobileNet.

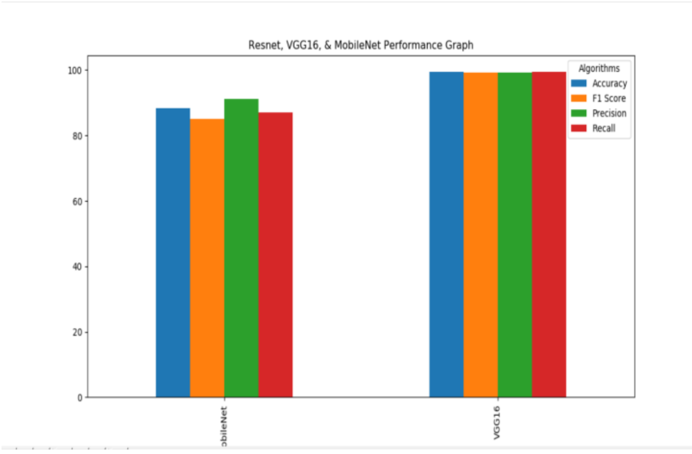



Fig4: Performance Graph of, VGG16, MobileNet

Figures 6 and 7 show the top 20 retrieved images for figure 5, where 19 of the 20 returned images are for the VGG16 model. This is due to the clarity and simplicity of the color and texture features in the images of the specified class.


Fig5: Original Image of Beach

Fig6: Top 1-10 Retrieved Images from the original image of the beach

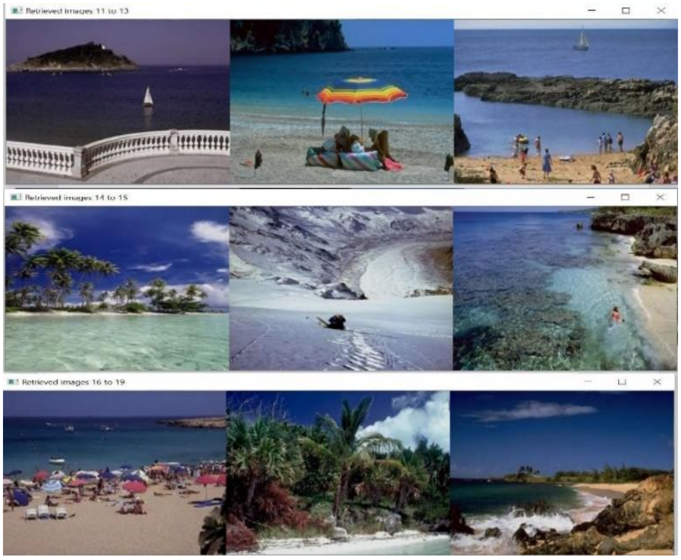


Fig7: Top 11-20 retrieved images from the original image beach

Figures 9 and 10 show the top 20 recovered images for figure 8, where 20 of the 20 retrieved images are for the VGG16 model.


Fig 8: Original Image of Bus

Fig9: Top 1-10 Retrieved Images from the original image of the bus

Fig10: Top 11-20 Retrieved Images from the original image of the bus

Visual examination demonstrates the pre-trained CNN models' good retrieval capabilities for these challenging Corel1K dataset images. Similar best results in terms of top retrieved images were obtained for Corel-1K elephants, flowers, and other classes.

## V. CONCLUSION

The content similarity is a key factor in matching images to a particular query image in the CBIR approach. The most significant image attributes exploited in CBIR are those related to color, texture, and shape. Convolution Neural Networks (CNNs) can extract important expressive features from a set of image data, making them ideal for image processing tasks including classification, object recognition, and grouping, among others. In this article, we developed a method to recover images using CNNs. The Corel1000k dataset was used as the subject of experiments. The trials' findings demonstrate the effectiveness and efficiency of the suggested approach for picture matching and retrieval, with an accuracy of 98.91666 for VGG16, and 91.80 for MobileNet. As the Accuracy of VGG16 is the highest among the models we have made use of VGG16 for Image Retrieval.

## VI. REFERENCES


[1]. Lin, X., J.-H. Li, S.-L. Wang, F. Cheng, and X.-S. Huang, Recent advances in passive digital image security forensics: A brief review. Engineering, 2018. 4(1): p. 29-39.

[2]. Fonseca, L.M.G., L.M. Namikawa, and E.F. Castejon., Digital image processing in remote sensing. in 2009 Tutorials of the XXII Brazilian Symposium on Computer Graphics and Image Processing. 2009. IEEE.

[3]. Liu, Y., D. Zhang, G. Lu and W.-Y. Ma, A survey of content-based image retrieval with high-level semantics. Pattern recognition, 2007. 40(1): p. 262-282.

[4]. Zare, M.R., R.N. Ainon, and W.C. Seng., Content-based image retrieval for blood cells. in 2009 Third Asia International Conference on Modelling & Simulation. 2009. IEEE.

[5]. Sreedevi, S. and S. Sebastian., Content-based image retrieval based on Database revision. in 2012 International Conference on Machine Vision and Image Processing (MVIP). 2012. IEEE.

[6]. Latif, A., A. Rasheed, U. Sajid, J. Ahmed, N. Ali, N. I. Ratyal, B. Zafar, S. H. Dar, M. Sajid, and T. Khalil ., Content-based image retrieval and feature extraction: a comprehensive review. Mathematical Problems in Engineering, 2019. 2019.

[7]. Wang, S., K. Han, and J. Jin., Review of image low-level feature extraction methods for content-based image retrieval. Sensor Review, 2019.

[8]. Thilagam, M. and K. Arunish., Content-Based Image Retrieval Techniques: A Review. in 2018 International Conference on Intelligent Computing and Communication for Smart World (I2C2SW). 2018. IEEE.

[9]. Chen, Y., The image retrieval algorithm based on the color feature. in 2016 7th IEEE International Conference on Software Engineering and Service Science (ACCESS). 2016. IEEE.

[10]. Zhang, Z., W. Li, and B. Li., An improving technique of color histogram in segmentation-based image retrieval. in 2009 Fifth International Conference on Information Assurance and Security. 2009. IEEE.

[11]. Chmelar, P. and A. Benkrid. Efficiency of HSV over RGB Gaussian Mixture Model for fire detection. in 2014 24th International Conference Radioelektronika. 2014. IEEE.

[12]. Nazir, A., R. Ashraf, T. Hamdani, and N. Ali., Content-based image retrieval system by using HSV color histogram, discrete wavelet transform, and edge histogram descriptor. in 2018 international conference on computing, mathematics, and engineering technologies (iCoMET). 2018. IEEE.

[13]. Weszka, J.S., C.R. Dyer, and A. Rosenfeld., A comparative study of texture measures for terrain classification. IEEE transactions on Systems, Man, and Cybernetics, 1976(4): p. 269-285.

[14]. Ahmed, A. and S. Mohamed, Implementation of early and late fusion methods for content-based image retrieval. International Journal of Advanced and Applied Sciences, 8(7) 2021, Pages: 97-105.

[15]. Rezazadeh, N., Initialization of weights in deep belief neural network based on a standard deviation of feature values in training data vectors. Vol (6), 2017(6): p. 708-715.

[16]. Han, D., Q. Liu, and W. Fan, A new image classification method using CNN transfer learning and web data augmentation. Expert Systems with Applications, 2018. 95: p. 43-56.

[17]. ImageNet. http://www.image-net.org.

[18]. Lou, Y., G. Fu, Z. Jiang, A. Men and Y. Zhou, PT-NET: Improve object and face detection via a pre-trained CNN model. in 2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP). 2017. IEEE.

[19]. Rajesh, K.B., Cybernetic microbial detection system using transfer learning. Multimedia Tools and Applications, 2018: p. 1-18.

[20]. Marmanis, D., M. Datcu, T. Esch and U. Stilla ., Deep learning earth observation classification using ImageNet pretrained networks. IEEE Geoscience and Remote Sensing Letters, 2015. 13(1): p. 105-109.

[21]. Hussain, M., J.J. Bird, and D.R. Faria. A study on CNN transfer learning for image classification. in UK Workshop on computational Intelligence. 2018. Springer.

[22]. Akcay, S., M. E. Kundegorski, C. G. Willcocks and T. P. Breckon ., Using deep convolutional neural network architectures for object classification and detection within x-ray baggage security imagery. IEEE transactions on information forensics and security, 2018. 13(9): p. 2203-2215.

[23]. Gopalakrishnan, K., S. K. Khaitan, A. Choudhary, and A. Agrawal ., Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection. Construction and Building Materials, 2017. 157: p. 322-330.

[24]. He, K., et al. Deep residual learning for image recognition. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[25]. Shirkhorshidi, A.S., S. Aghabozorgi, and T.Y. Wah, A comparison study on similarity and dissimilarity measures in clustering continuous data. PloS one, 2015. 10(12).

[26]. Varish, N., A. K. Pal, R. Hassan, M. K. Hasan, A. Khan, N. Parveen, D. Banerjee, V. Pellakuri, A. U. Haqis and I. Memon ., Image retrieval scheme using quantized bins of color image components and adaptive tetrolet transform. IEEE Access, 2020. 8: p. 117639-117665.

[27]. Li, J. and J.Z. Wang, Real-time computerized annotation of pictures. IEEE transactions on pattern analysis and machine intelligence, 2008. 30(6): p. 985-1002.

[28]. Liu, G.-H., J.-Y. Yang, and Z. Li, Content-based image retrieval using computational visual attention model. pattern recognition, 2015. 48(8): p. 2554-2566.

**Cite this article as :**

Dr. J C Karur, Asma Hebbal, Dr. Jagadeesh Pujari, "Content-Based Image Retrieval Using Deep Learning", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 8 Issue 4, pp. 122-128, July-August 2022. Available at doi : https://doi.org/10.32628/CSEIT228418
Journal URL : https://ijsrcseit.com/CSEIT228418