IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-10, Oct 2018 E-ISSN: 2347-2693

Content Based Image Retrieval using Learnt Features from Convolution Neural Networks

Vijayakumar Bhandi[1*], Sumithra Devi. K. A.[2]

[1]Dept. of Computer Science and Engineering, Jain University, Bangalore, India
[2]Dept. of Information Science and Engineering, Dayananda Sagar Academy of Technology & Management, Bangalore, India

Available online at: www.ijcseonline.org

Accepted: 20/Oct/2018, Published: 31/Oct/2018

***Abstract*—** Content based image retrieval (CBIR) relies on fetching relevant images from a dataset based on low level image features. Image features such as colour, texture and shape have been widely used in CBIR applications. These handcrafted features have been carefully designed and found to perform well in image retrieval task. The performance of these features greatly depends on the choice of the handcrafted features being used and domain knowledge. Hence there is a need to identify image features which are independent of domain knowledge and can be dynamically extracted from image data. Machine learning is a promising area here, since it focuses on learning representations from input data. Machine learning methods have been applied in various image processing tasks earlier. Convolution neural networks (CNN) models are able to create expressive features from image data and are successfully applied in image classification tasks. In this paper, we create a frame work to use CNNs to learn features from the image data and use these learned features for content based image retrieval. We test our proposed CBIR framework to retrieve images from a digital library database of art images. The results are compared against standard CBIR model which uses global colour histogram handcrafted feature. The results show that the learnt features extracted from a CNN model perform equally good as handcrafted features when applied to image retrieval task.

***Keywords* —** Content based image retrieval, Convolution neural networks, Machine learning, Global colour histogram

I. INTRODUCTION

The growth in digital technology has led to creation of huge multimedia databases containing various kind of data like images and videos. Content based image retrieval [CBIR] deals with retrieval of top matching images from a collection of images based on the content similarity. CBIR uses low level features of the image for image retrieval. There are various low level features used in CBIR, which have been carefully designed and used effectively in various image retrieval applications. These are usually referred as handcrafted features. Some of the important handcrafted features are colour, texture and shape features. colour is the most dominant and distinguished visual feature of a digital image. The colour feature is relatively robust to background complication and independent of image size and orientation. Hence colour is most widely used visual feature in CBIR.

Colour histogram is the basic colour content representation and it captures global Colour distribution of an image. Colour histogram describes the statistical colour distributions by quantizing the colour space. Colour histograms are easy to compute, but they result in large feature vectors which are difficult to index. Global colour histogram [GCH] is the most widely used handcrafted feature in content based image retrieval due to its good performance in colour analysis of digital images. In our work we have created a standard CBIR frame work which uses GCH as the handcrafted feature. We will use this method as a baseline method to compare the results of image retrieval of the proposed CBIR method.

The handcrafted image features perform good in image retrieval task but however the performance depends on the choice of the features to be used. The choice of handcrafted features requires good domain knowledge about the image retrieval task at hand. If one chooses inadequate features then the performance of the image retrieval will be greatly affected. Hence there is a need to create a content based retrieval system which can learn the features from the image data and perform image retrieval. Such a system should perform as good as or better than the CBIR using handcrafted features.

Machine learning is a branch of artificial intelligence and it is based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention. In today’s world, machine learning has been successfully utilized in many fields like financial services, healthcare, transportation, oil and gas and defense etc. Machine learning finds great use in various real world applications like classification, medical diagnosis, prediction, learning association and regression etc. Machine learning has been successfully applied in image processing tasks like classification, clustering, and object recognition etc.

Neural networks are an important class of models within machine learning domain. Neural networks are a specific set of algorithms that have revolutionized machine learning. Neural networks are inspired by biological neural networks. Neural Networks are themselves general function approximations, hence they can be applied to almost any machine learning problem about learning a complex mapping from the input to the output space.

A Convolutional Neural Network (CNN) is a stack of non-linear transformation functions that are learned from input data. CNNs derive their name from the “convolution” operator. The primary purpose of convolution in case of a CNN is to extract features from the input image. Convolution preserves the spatial relationship between pixels by learning image features using small squares of input data. CNNs are being used in various image processing applications like classification, clustering and object recognition etc. since CNNs are able to learn powerful features from complex image data.

In this paper, we explore the idea of extracting important features learnt by a CNN model and use these learnt features to content based image retrieval task. Also, we implement a frame work for using CNN features for content based image retrieval from a complex art images database.

This paper is organized as follows: Section II describes the related work. Section III explains the proposed methodology. Section IV details about the experimental results and Section V presents the conclusion of this research work.

II. Related Work

Retrieving images from a collection matching to the input query image has always been an interesting research problem. The traditional text-based image retrieval methods use manually annotated keywords for searching the relevant images. This is labour-intensive, time consuming and expensive method. Also the rich semantics of the image are difficult to describe and human perception affects it [1] [2]. Content based image retrieval [CBIR] methods rely on low level features for image matching. These low level features can be automatically extracted from the image. A typical CBIR system works on query by example method, where low level features of query image are matched against database image features to extract top N similar images [1] [2]. Colour, texture and shape features are most widely used and accepted low level visual features in CBIR [1] [2]. These features can be called as handcrafted features, as these are carefully designed and tested over the time for image retrieval task.

Colour feature is the most widely used handcrafted feature in CBIR. It is the most distinguishing and dominant low level visual feature in CBIR [2]. The global colour feature is simple to calculate and provides good discriminating power in image retrieval. The colour feature has invariance to the orientation and scaling of image hence it is most suitable for an effective image retrieval application [3]. The colour information of image can be extracted by different techniques but the mostly used and prominent technique is colour histogram [4]. Because of the robustness, effectiveness, implementation simplicity and low storage requirements advantages, colour has been the most effective feature and almost all CBIR systems employ colour feature. HSV or CIE Lab and LUV spaces are used to represent colour instead of the RGB space as they are much better with respect to human perception [6]. Many CBIR systems have been designed by using Global colour histogram feature and achieved good image retrieval results [5][7].

Since their introduction by LeCun et al., Convolutional Neural Networks (CNNs) have demonstrated excellent performance at tasks image classification and object detection [8]. CNN's are effectively used in various pattern and image recognition applications like gesture recognition, face recognition, and object classification [9] [10] [11]. Extracting effective features is the major important stage in a lot of object recognition and computer vision tasks [15]. Several researchers have focused on designing robust features for a variety of image classification tasks [12]. Nowadays, much attention is given to feature learning algorithms and CNNs [16]. Earlier research works show that it is possible to feed an image directly to a CNN network and utilize features for image classification [12] [13]. In our work, we are extending this idea to extract features from a CNN model and use them for image retrieval task.

Image retrieval from a collection of digital art images is quite challenging and interesting problem. From technical point of view, art image retrieval encounters many challenges, like high-dimension feature space, nonlinear distributions, insufficient training examples and the semantic gap between low-level features and high level contents. Earlier works have shown that colour aesthetics and semantics can be used for image retrieval from digital library of art images [14]. We have used a digital library art images dataset for our experiments in this work, since we can effectively make use of colour feature for image retrieval and proposed CBIR method can be validated effectively.

III. Methodology

The proposed CBIR methodology consists of an offline database indexing module and online query module as outlined in Fig. 1.

A. *Offline database indexing*

This is one time offline exercise to extract the required features [GCH handcrafted or CNN learnt features] from all images in the database. The features extracted are stored in a feature database.

B. *Online query module*

In this online module, the retrieval of matching images corresponding to input query image is performed. The query image features [GCH handcrafted or CNN learnt features] are computed and are compared against the features database. Based on the sorted distances, the matching top N images are returned as results.

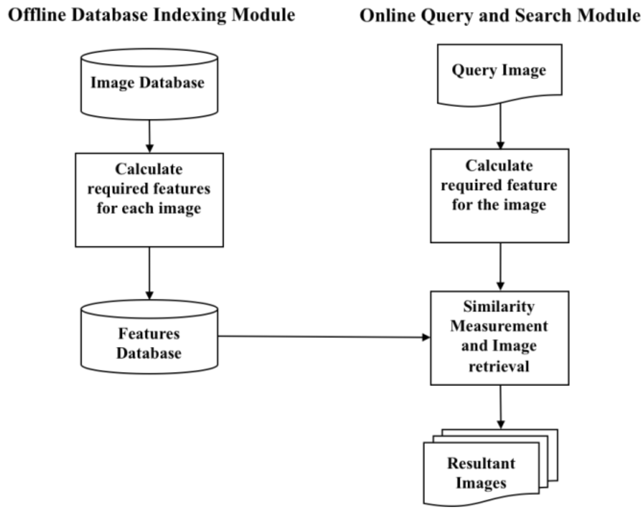

Fig. 1. Proposed CBIR method

In the proposed work we have created two CBIR frameworks, one a standard CBIR using GCH handcrafted feature and another one is the proposed CBIR with CNN learnt features.

C. *CBIR using Global colour Histogram handcrafted feature*

We created a basic CBIR framework which uses GCH handcrafted feature for image retrieval. We will compare results from our proposed CBIR framework with this baseline CBIR to gauge its effectiveness. The input image is converted to HSV colour space and GCH feature [H-8, S-12, V-3 bins] is calculated and normalized. The GCH feature gives us a vector of 288 values for each image.

D. *CBIR using CNN learnt features*

This is our proposed frame work, where we extract learnt features from Convolution neural network for image retrieval task. A CNN model usually has multiple layers that progressively compute features from input images as shown in figure 2

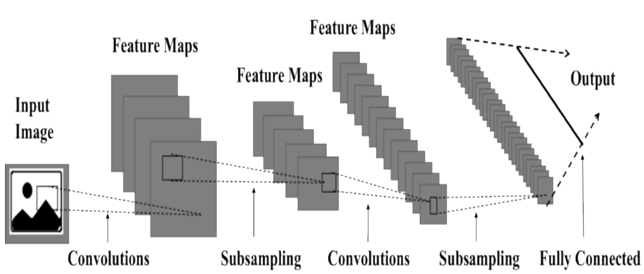

Fig. 2. Architecture of Convolution neural network model

In our proposed CNN architecture, we have used following configurations:

- 3 convolution layers with different weights [Conv layer 1 with 64 weights, layer 2 with 32 weights and layer 3 with 25 weights].
- Kernel size with 3X3 filter matrix
- Rectified Linear Activation [relu] as activation function for each layer
- Images are resized to 224X224 size while feeding to the CNN network
- RMSprop is used as an optimizer
- 3 Max pooling layers and 3 dense layers are used

E. *Performance Evaluation*

The performance of the image retrieval is measured in terms of precision as defined in equation (1).

$$Precision = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}} \quad (1)$$

F. *Implementation*

The proposed CBIR method is implemented using Python 3.6.5, Keras 2.2.2 module, Tensoflow 1.11 as backend on macOS High Sierra 10.13 (64 bit) operating system with 16GB RAM and Intel core i7 processor.

G. *Experimental Work*

Proposed CBIR method is subjected to experimental study on art images dataset derived from "The Russian Museum: the Virtual Branches" (website http://rusmuseumvrm.ru/). The dataset consists of 7000+ images which are grouped into 5 classes as described in Table 1 and sample images in Table 2.

Table 1: Art images dataset from Museum collection

Sr. No#	Art group/Class	Number of Images per class
1	Drawings and water colours	1100 Images
2	Paintings	2000 Images
3	Sculptures	1700 Images
4	Engravings (Graphic Art)	800 Images
5	Iconography (Old Russian Art)	2000 Images

Two images from each class are used as query images and these images are not included in the training set (dataset). Given a query image from a class, it is assumed that the user is searching for images from the same class, and therefore the images from the same class are considered relevant and the images fetched from all other classes are considered irrelevant. The precision rates are measured by varying the number of images retrieved.

Table 2: Sample images from the Art images dataset

Class - Drawings		
Class – Paintings		
Class – Sculptures		
Class – Engravings		
Class - Iconography		

IV. Results and Discussion

The image retrieval experiment is carried out in 2 steps as results are outlined in below sections. Same query images have been used in both steps.

A. Image retrieval using Global colour Histogram handcrafted feature [Menthod 1]

We created a CBIR framework for image retrieval by using GCH feature on the entire image. We use this as a baseline model to measure performance of proposed method over it. GCH feature [H-8, S-12, V-3 bins] is calculated in HSV colour space for each image in the dataset and stored in features database. Similarly, GCH for query image is calculated and compared against features database. Cosine distances method is used for features similarity measurement and matching images are retrieved. We repeat the image retrieval experiment for each query image by varying the number of images retrieved. In this baseline CBIR, we get 65.32% average precision across all classes. The class wise average precision is tabulated in table 3.

B. Image retrieval using Convolution neural networks learnt features [Method 2]

In this step, our proposed CBIR method using learnt features from CNN method is subjected to experimentation. A convolution neural network model is created with the configurations as highlighted in the methodology section of this paper. The CNN model is trained on the training set from the art images dataset. For our experimentation we extract the features from the first dense layer of the model. These are the outputs from the first dense layer. From this layer we get 512 features per image. During offline database indexing module, these features are stored in the features database. During the online query experiment, we compare 512 features obtained from each query image with the features database. We use cosine distances for calculating the image similarity and retrieve top N results. We repeat the image retrieval experiment for each query image by varying the number of images retrieved. Image retrieval results from our proposed approach are tabulated in table 3.

Table 3: Average precision rates for image retrieval.

Image Class	[Method 1] Baseline CBIR using GCH handcrafted feature	[Method 2] Proposed CBIR method using CNN learnt features
Drawings	36.07%	57.57%
Engravings	25.94%	44.40%
Iconography	97.37%	79.50%
Paintings	99.84%	97.74%
Sculptures	67.40%	51.80%
Overall Precision	65.32%	66.12%

Using proposed CBIR method we get 66.12% average precision. The proposed method performs better or equally good in retrieving images from 3 classes, Drawings, Engravings and Paintings. However, the precision of proposed method is lower than baseline method for two classes, Iconography and sculptures. For lower fetch sizes (5,10,15) we see that the proposed CBIR method performs very well. Looking at the overall results, the new proposed CBIR method using CNN learnt features performs better as compared to baseline CBIR using handcrafted features.

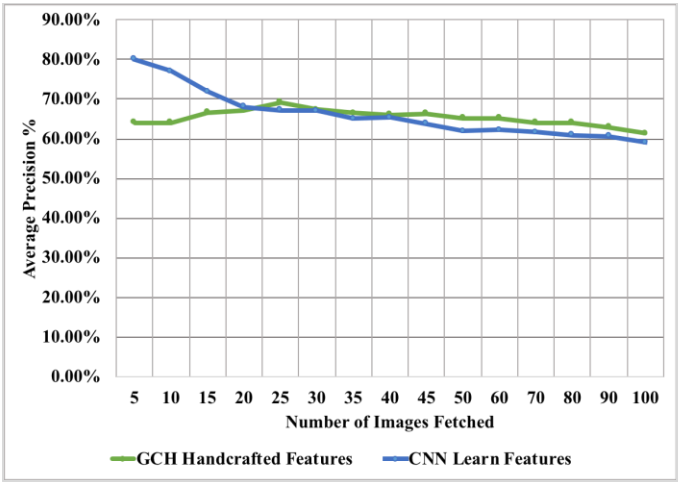

Fig. 3. Average precision vs number of images retrieved for baseline CBIR and proposed CBIR

V. Conclusion and Future Scope

In this work, we demonstrated that it is possible to extract learnt features from Convolution neural network machine learning algorithms and use these features for image retrieval task. We created a content based image retrieval framework which uses CNN learnt features. The proposed CBIR framework was tested on a multiclass digital database of art images containing more than 7000 images. We achieved average retrieval precision of 66.12% for the proposed CBIR framework, which is slightly better than the average retrieval precision of 65.32% from CBIR framework using global colour histogram feature. From experimental results we conclude that learnt features from CNN network can be successfully used for image retrieval task and perform as good as handcrafted features.

Future Scope: In the current work we have created a basic convolution neural network. It is possible to fine tune the CNN by adding more layers, optimal dropouts, and other optimizers to achieve better results. Also, we can look at utilizing some of the pretrained neural networks like VGG16, RestNet, and GoogleNet etc.

References

[1] Yong Rui and Thomas S. Huang, “Image Retrieval: Current Techniques, Promising Directions and Open Issues”, Journal of Visual Communication and Image Representation, 1999.

[2] Wei-Ying Ma and HongJiang Zhang, “Content-Based Image Indexing and Retrieval”, Handbook of Multimedia Computing, CRC Press, 1999.

[3] Lei, Z., Fuzong, L. and Bo, Z., “A CBIR method based on colour-spatial feature”, Proceedings of IEEE Region 10 Conference: TENCON 99, Multimedia Technology for Asia-Pacific Information Infrastructure, Vol 1: 166-169, 1999.

[4] Park, J., An, Y., Kang, G., Rasheed, W., Park, S. and Kwon, G., “Defining a New Feature Set for Content-Based Image Analysis Using Histogram Refinement “, International Journal of Imaging Systems and Technology - Multimedia Information Retrieval, Volume 18 Issue 2-3, 2008.

[5] N. Sharma, P. Rawa, and J. Singh, “Efficient CBIR using colour histogram processing,” Signal & Image Processing, vol. 2, no. 1, p. 94, 2011.

[6] H. Muller, N. Michoux, D. Bandon, A. Geissbuhler, "A review of content-based image retrieval systems in medical applications—clinical benefits and future directions", International Journal of Medical Informatics, 73, pp. 1-23, 2004.

[7] C.H. Lin, R.T. Chen and Y.K. Chan, “A smart content-based image retrieval system based on colour and texture feature”, Image and Vision Computing vol.27, pp.658–665, 2009.

[8] LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D.: Backpropagation applied to handwritten zip code recognition. Neural Comput. 1(4), 541–551, 1989.

[9] Bobić, V., Tadić, P., Kvaščev, G, “Hand gesture recognition using neural network based techniques.” in Neural Networks and Applications (NEUREL), 13th Symposium on (pp. 1-4). IEEE, 2016.

[10] Lawrence, S., Giles, C. L., Tsoi, A. C., & Back, A. D. “Face recognition: A convolutional neural-network approach.”IEEE transactions on neural networks, 8(1):98-113,1997.

[11] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Rabinovich, A, “Going deeper with convolutions.” in Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9). 2015.

[12] S. Minaee, A. Abdolrashidi, and Y. Wang, "Iris recognition using scattering transform and textural features," in Signal Processing and Signal Processing Education Workshop (SP/SPE), 2015 IEEE, pp. 37-42, 2015.

[13] S. Minaee, A. Abdolrashidi, and Y. Wang, "Face Recognition Using Scattering Convolutional Network," arXiv preprint arXiv:1608.00059, 2016.

[14] Ivanova, Krassimira, “Content-Based Image Retrieval in Digital Libraries of Art Images Utilizing colour Semantics”, Institute of Mathematics and Informatics – BAS, Sofia, Bulgaria, 515-518. 10.1007/978-3-642-24469-8_62, 2011.

[15] D. Sherlin, D. Murugan, "A Case Study on Brain Tumor Segmentation Using Content based Imaging", International Journal of Scientific Research in Network Security and Communication, Vol.6, Issue.3, pp.1-5, 2018

[16] N. Rezazadeh, "Initialization of weights in deep belief neural network based on standard deviation of feature values in training data vectors", International Journal of Scientific Research in Computer Science and Engineering, Vol.5, Issue.4, pp.1-8, 2017

Authors Profile

Vijayakumar Bhandi is a research scholor in Department of Computer Science and Engineering, Jain Univerisity, Bangalore, India. His research work focuses on Image processing and machine learning field.

Dr. Sumithra Devi. K. A. is currently serving as Professor & HOD, Information Science and Engineering Dept, Dayananda Sagar Academy of Technology & Management, Bangalore, India. She has several of years experience in teaching and research. She has successfully guided many research scholars in the field of computer science & engineering and information science.