Content-based Image Retrieval using Deep Convolution Neural Network

Mr. Pukale D. D.
Research Scholar
Bharti Vidyapeeth University
College of Engineering, Pune
pukaledd@gmail.com

Dr. S. G. Bhirud
Dept. of CE & IT
V.J.T.I., Mumbai
sgbhirud@vjti.org.in

Dr. V. D. Katkar
Information Technology Dept.
PCCOE, Pune
katkarvijayd@gmail.com

***Abstract*— Use of computationally expensive Neural Network for processing huge amount of data is increased in recent past. Deep Convolution Neural Network is widely used by researchers to analyze images for variety of applications. This paper presents a Content based Image Retrieval system using Deep Convolution Neural Network. Experiments are performed on Gray images, RGB color space, YCbCr color space images grouped into different clusters. Precision-recall crossover point is used as performance measurement criteria.**

***Keywords*— Deep Learning, Content based Image Retrieval, Convolution Neural Network**

I. INTRODUCTION

Use of digital image processing based applications has increased many fold with availability of low price disk storages and high speeds processors. Image databases containing millions of images are now cost effective to create and maintain [9].

Earlier image retrieval systems are text-based; in which images are manually annotated and then indexed according to the annotation. However, with the exponential increase in the volume of images database, the task of user-based annotation becomes very cumbersome [10]. Major Limitations of text annotation based image search are:

a. Image annotation is subjective
b. Image annotation may fails to convey the complete and appropriate details of image.

Content based Image Retrieval (CBIR) system extract and represents the visual features of images like color, shape, texture and uses these features to discover visually similar images. Color histogram is one of the important and widely used feature to describe the content of image. It represents the proportion of specific colors in an image [11]. Visual features are broadly categorized as Global and Local features. Global features are extracted from image by treating it as a single entity where as local features are extracted from a specific region of the image [12].

CNN is a Feed-Forward Neural Network that can extract topological properties from an image. It can recognize patterns with extreme variability as a result many researchers [1, 2, 4] have used Convolution Neural Network for image feature extraction & representation for CBIR.

Rest of the paper is organized as follows. Section 2 describes the related work carried out by other researchers using Neural Network for CBIR. Section 3 briefly explains the Convolution Neural Network and section 4 presents the experimental system architecture. Section 5 presents the experimental results and section 6 concludes the paper.

II. RELATED WORK

Ruigang Fu et. al. [1] have presented a CBIR system using Deep Convolution Neural Network and Linear Support Vector Machine. Deep features are extracted from images using Convolution Neural Network and similar images are identified using deep features and linear Support Vector Machine. Retrieved similar images are ranked based on distance between retrieved image and trained hyperplane.

Kun Hu et. al. [2] have proposed Multi-Level Pooling method to extract object-aware deep image features from different layers of Convolution Neural network. Features extracted from different layers are used to create a short representative feature vector for CBIR.

Domonkos Varga and Tamás Szirányi [3] have presented a supervised learning framework for CBIR which learns the probability-based semantic-level similarity and feature-level similarity simultaneously.

Xinran Liu et. al. [4] have described a novel method for CBIR using combination of Convolution Neural Network and Radon Barcodes. Initially Convolution neural Network is used for Global classification of images; then Radon Barcodes are used for retrieval of similar images.

Shun-Chin Chuang et. al. [5] have proposed a CBIR system using Multiple Instance Neural Network. It takes samples of related and unrelated images from user and uses color histogram of these images to train the Neural Network. Then this trained Neural Network is used to retrieve the similar images.

P. Muneesawang and L. Guan [6] have presented an adaptive CBIR system using Neural network. To improve the

978-1-5386-4008-1/17/$31.00 ©2017 IEEE

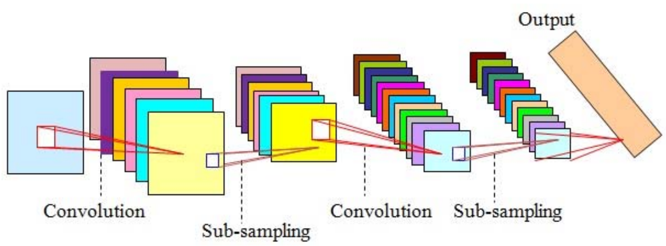
image retrieval performance; weights of trained Neural Network are adaptively changed using user relevance feedback.

K. -H. Yap and K. Wu [7] have proposed an adaptive system using Fuzzy relevance feedback and Radial Bias Function Network. Users Fuzzy interpretation of image similarity is used to update the parameters of Radian Bias Function network in order to improve the CBIR performance.

Samy Sadek et. al. [8] have described the design of CBIR system using Cubic Splines Neural Network. With the help of Cubic Splines Neural Network, CBIR system is able to learn the non-linear relationship among the images and improves the performance retrieval performance.

III. CONVOLUTION NEURAL NETWORK

Figure 1 shows the general structure of Convolution Neural Network (CNN). It is a special type of Multilayer Feed-Forward Neural Network designed to recognize visual patterns from image pixels. It can recognize patterns with extreme variability



Fig, 1.General Structure of Convolution Neural network

Each convolution layer of a CNN is composed of multiple feature maps. Feature maps are in the form of a plane and all the neurons of feature map are constrained to share the same set of synaptic weights. Each neuron in CNN takes inputs from a receptive field in the previous layer, which enables it to extract local features.

Convolution layer is followed by a sub-sampling layer This layer performs local averaging and sub-sampling of pool if image pixels, which in-turn reduces the resolution of feature map. By reducing the spatial resolution of the feature map, certain degree of shift and distortion invariance is achieved.

IV. SYSTEM ARCHITECTURE

Figure 2 shows the process of training the Deep CNN model for CBIR. Images present in the image dataset are clustered into 'K' number groups using K-NN clustering algorithm and then given as a input to the Deep CNN model to train it. After training, features are extracted from trained CNN model to represent the images in training dataset.

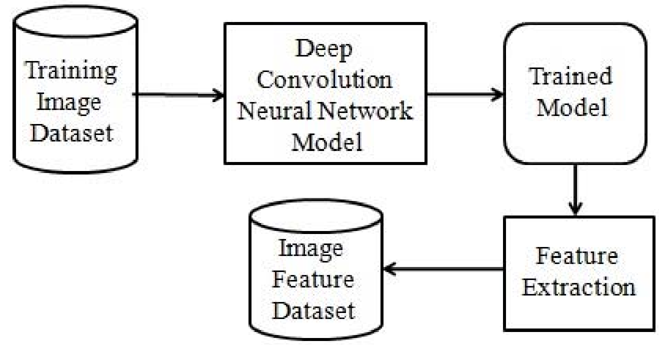


Fig. 2: Training Deep CNN Model for CBIR

Table 1 shows the internal structure of Deep CNN model used to train the CBIR system.

TABLE I. DETAILS OF CONVOLUTION NEURAL NETWORK

Layer Description	**Output Shape**
Input	(128, 128, 3)
Convolution (3*3*32)	(128, 128, 32)
Convolution (3*3*32)	(128, 128, 32)
Max Pool (2*2)	(64, 64, 32)
Convolution (3*3*32)	(64, 64, 32)
Convolution (3*3*32)	(64, 64, 32)
Max Pool (2*2)	(32, 32, 32)
AFFINE (512 Units)	(512, 1)
AFFINE (512 Units)	(512, 1)

Figure 3 shows the process of retrieving similar images using Deep CNN model. Features of query image are extracted using trained CNN model. These features are compared with features of available images using Euclidian distance measure to discover similar images.

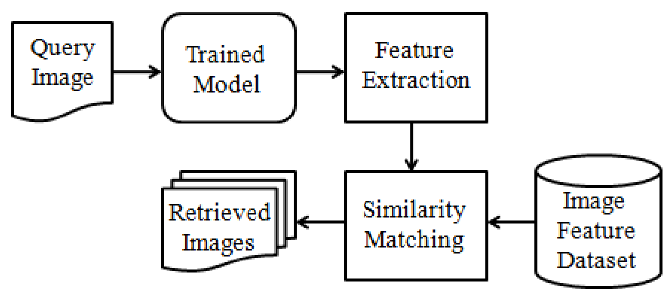


Fig. 3. Process to retrieve Similar Images

V. EXPERIMENTAL RESULTS

Wang 1000 image dataset [13, 14] is used to perform the experiments. It contains images belonging to various categories like Tribal persons, Sea shores, buses, manmade structures, flowers, etc... Experiments are performed using machine having 8 GB RAM and 2.30 GHz Quad-core processor. The images are converted into YCbCr, RGB color space and Gray. Experiments are conducted on three types of images by grouping them in 20 and 30 clusters. The following section describes the results obtained for all the six cases.

Case 1: Color Model: YCbCr, Number of Clusters: 20

Figure 4 shows the experimental results obtained for case 1. Red curve indicates average precision and green curve indicates average recall obtained at different threshold values. It is observed that the crossover point is at 0.37.

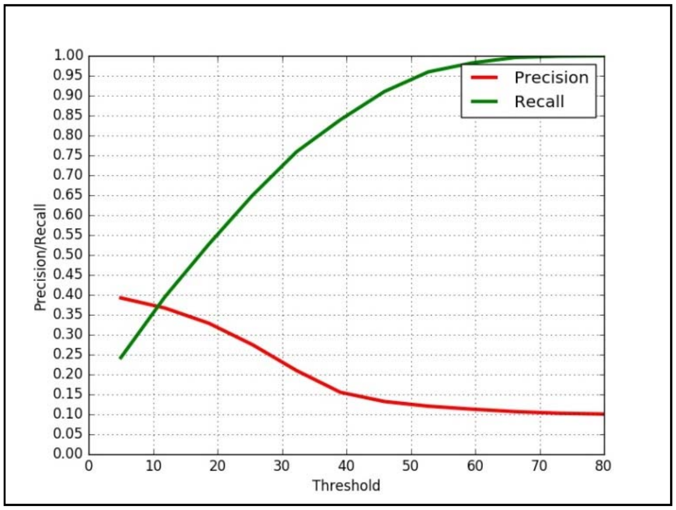


Fig. 4. Precision vs Recall Graph (Color Model: YCbCr, Clusters: 20)

Case 2: Color Model: YCbCr, Number of Clusters: 30

Figure 5 presents the experimental results for case 2. Red and Green curve indicates average precision and average recall obtained at different threshold values. It can be observed that the crossover point is obtained at 0.365.

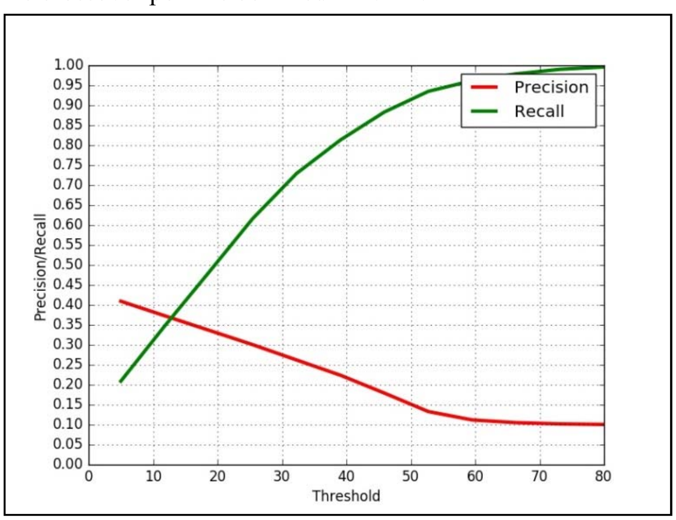


Fig. 5. Precision vs Recall Graph (Color Model: YCbCr, Clusters: 30)

Case 3: Color Model: RGB, Number of Clusters: 20

Experimental results obtained for case 3 are shown in Figure 6. Average precision obtained is shown using red curve and average recall obtained is shown using green curve. Crossover of precision and recall is observed at 0.37.

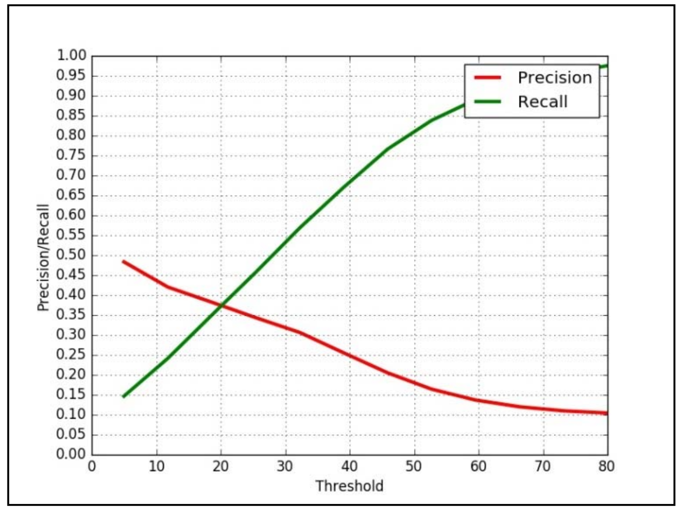


Fig. 6. Precision vs Recall Graph (Color Model: RGB, Clusters: 20)

Case 4: Color Model: RGB, Number of Clusters: 30

Experimental results for case 4 are presented in figure 7. Average precision and average recall obtained at different threshold values is shown using red curve, green curve respectively. Precision-recall crossover is observed at 0.35.

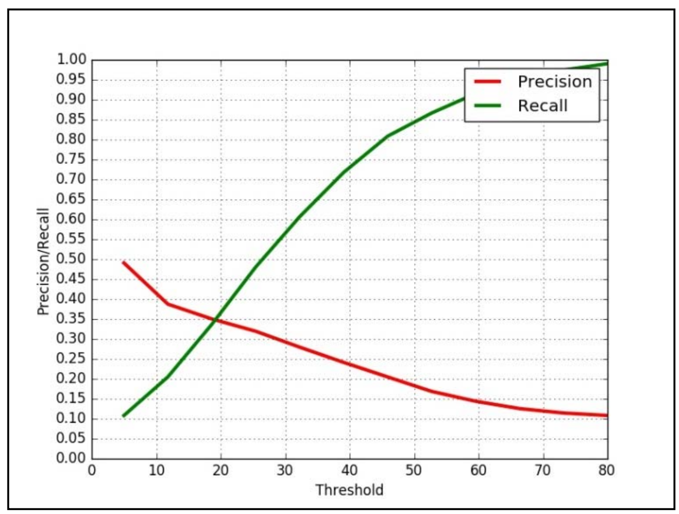


Fig. 7. Precision vs Recall Graph (Color Model: RGB, Clusters: 30)

Case 5: Color Model: Gray, Number of Clusters: 20

Figure 8 shows the experimental results obtained for case 5. Red curve indicates average precision and green curve indicates average recall obtained at different threshold values. It is observed that the crossover point is at0.34.

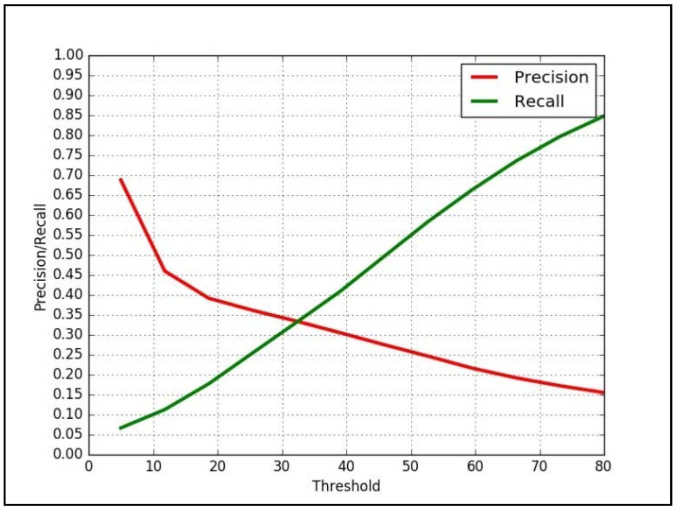


Fig. 8. Precision vs Recall Graph (Color Model: Gray, Clusters: 20)

Case 6: Color Model: Gray, Number of Clusters: 30

Experimental results obtained for case 6 are shown in Figure 9. Red and Green curve indicates average precision and average recall obtained at different threshold values. It can be observed that the crossover point is obtained at 0.37.

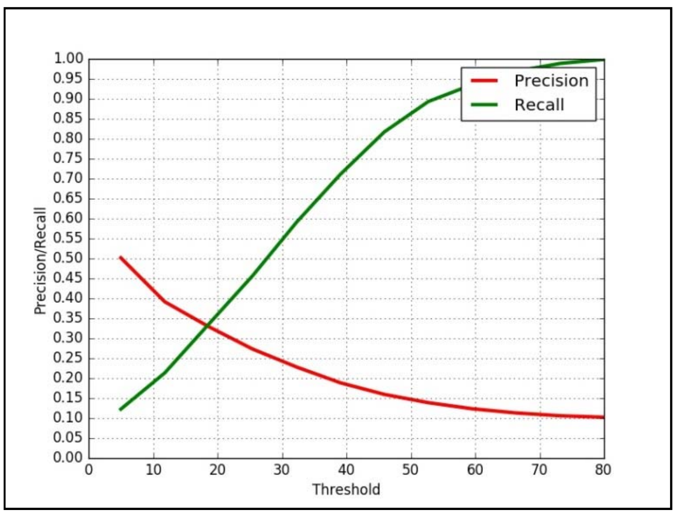


Fig. 9. Precision vs Recall Graph (Color Model: Gray, Clusters: 30)

TABLE II. SUMMARY OF EXPERIMENTAL RESULTS

Sr. No.	Color Model	Number of clusters	Crossover point
1	RGB	20	0.37
2	RGB	30	0.35
3	YCbCr	20	0.37
4	YCbCr	30	0.365
5	Gray	20	0.34
6	Gray	30	0.37

Table 2 summarizes the experimental results obtained for Gray Images, RGB color space, YCbCr color space with different image clusters.

The result shows that, more number of clusters are required for Gray images to obtain the same crossover point as compared to RGB and YCbCr color spaces.

VI. CONCLUSION

This paper presents the CBIR system using Deep CNN. Wang 1000 image dataset is used to perform the experiments. For RGB, YCbCr color space and Gray images precision-recall crossover point is obtained at 0.37 for different number of clusters. Further improvement in performance can be verified by increase in number of Convolution layers.

REFERENCES

[1] Ruigang Fu, Biao Li, Yinghui Gao, Ping Wang, "Content-based image retrieval based on CNN and SVM", 2nd IEEE International Conference on Computer and Communications (ICCC), 2016, pp. 638-642

[2] Kun Hu, Yuan Dong, Hongliang Bai, "Multi-level convolutional channel features for content-based image retrieval ", Visual Communications and Image Processing (VCIP), 2016, pp. 1-4

[3] Domonkos Varga and Tamás Szirányi, "Fast content-based image retrieval using convolutional neural network and hash function ", IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2016 , pp. 26-36

[4] Xinran Liu, H.R. Tizhoosh, J. Kofman, "Generating binary tags for fast medical image retrieval based on convolutional nets and Radon Transform", International Joint Conference on Neural Networks (IJCNN), 2016

[5] Shun-Chin Chuang, Yeong-Yuh Xum Hsin Chia Fu and Hsiang-Cheh Huang, "A Multiple-Instance Neural Networks based Image Content Retrieval System", First International Conference on Innovative Computing, Information and Control - Volume I (ICICIC'06), pp. 412-415

[6] P. Muneesawang and L. Guan, " A neural network approach for learning image similarity in adaptive CBIR", IEEE Fourth Workshop on Multimedia Signal Processing 2001, pp. 257-262

[7] K.-H. Yap and K. Wu, "Fuzzy relevance feedback in content-based image retrieval systems using radial basis function network", IEEE International Conference on Multimedia and Expo, 20015

[8] Samy Sadek, Ayoub Al-Hamadi, Bernd Michaelis, and Usama Sayed, "Image Retrieval Using Cubic Splines Neural Networks", International Journal of Video & Image Processing and Network Security, Vol: 9 No: 10, pp. 17-22

[9] Syed Hamad Shirazi, Noor ul Amin Khan, Arif Iqbal Umar, Muhammad Imran Razzak, Saeeda Naz, Bandar AlHaqbani, "Content-Based Image Retrieval Using Texture Color Shape and Region", International Journal of Advanced Computer Science and Applications, Vol. 7, No. 1, pp. 418-426

[10] Miss. Aboli W. Hole1, Prof. Prabhakar L. Ramteke, "International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, Issue 10", pp. 45-49

[11] Ms. Pragati Ashok Deole,, Prof. Rushi Longadge, "Content Based Image Retrieval using Color Feature Extraction with KNN Classification", , International Journal of Computer Science and Mobile Computing, Vol.3 Issue.5, pp. 1274-1280

[12] Rehan Ashraf Khalid Bashir, Aun Irtaza and Muhammad Tariq Mahmood, "Content Based Image Retrieval Using Embedded Neural

Networks with Bandletized Regions", Entropy 2015, Vol. 17, pp. 3552-3580

[13] Jia Li, James Z. Wang, ``Automatic linguistic indexing of pictures by a statistical modeling approach," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 9, pp. 1075-1088, 2003. (download)

[14] James Z. Wang, Jia Li, Gio Wiederhold, ``SIMPLIcity: Semantics-sensitive Integrated Matching for Picture LIbraries," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol 23, no.9, pp. 947-963, 2001