International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

COPY RIGHT

IJIEMR Transactions, online available on 12th Nov 2018. Link :http://www.ijiemr.org/downloads.php?vol=Volume-07&issue=ISSUE-12

Title: IMAGE RETRIEVAL BASED ON CNN ARCHITECTURES

Volume 07, Issue 12, Pages: 115–122.

Paper Authors

N VENKAT SUBBARAO, K.RAMANJANEYULU, N.SRAVANI

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

IMAGE RETRIEVAL BASED ON CNN ARCHITECTURES

N VENKAT SUBBARAO[1], K.RAMANJANEYULU[2], N.SRAVANI [3]
[1]M.Tech, E-mail:--vvnallamalli@gmail.Com
[2]Research scholar, JNTUK, Kakinada ,Assoc Professor, QISIT, Ongole, E-mail:--Ramu36nba@gmail.Com
[3]Assistant Professor, Department Of ECE, QISCET, ONGOLE, E-Mail:--sravanisrv@gmail.com

Abstract- With the advancement of multimedia technology, large number of images is used in different areas such as medical applications, military, computer vision applications, agriculture etc. For retrieval process feature representation is the main important task for an efficient system. In this paper we use 2 different models of CNN architectures i.e., VGG19, Inception and compared their performances. Using CNN architectures generated the feature vectors for a database images and query image. Compute the Euclidean distance between the query and database feature vectors & if the distance between these 2 vectors is less then it more similar to the query image. The CNN models are tested on the corel-1k data base. The system performance can be computed by using precision, recall.

Key Words: CBIR, CNN architecture, Precision, Recall.

I. INTRODUCTION

Recently, Pictures are extensively used because an image gives a lot of information compared to the text. Due to the fast increase of digital technologies like computers and internet services makes it possible to store and send large amounts of images. Instead of downloading, image retrieval has been needed in modern eras. Content Based Image Retrieval (CBIR) can be considered as the effective techniques to use visual data [1]. In this method mainly image retrieval done based on the image content itself like shape color, image structure etc., rather than the interpreted text of an image. Data mining is strongly challenged by traditional database technologies, but traditional text objects cannot meet the needs of the image database. The traditional way of notifying images using text does not automatically have an image description. In order to implement CBIR, the system must comprehend and deduce the contents of the image. The display index must be automatically generated, which affords a more graphical interface for downloaders. CBIR denotes that contents of an image are directly extracts itself that will be searched in the image database. The basic notion of CBIR is to evaluate image data from low-level image functions, including color, texture and shape, as well as create feature vector functions, such as your index. Extraction methods focus on such extraction and are basically implemented in line with multi-dimensional images.

Being images having a lot of information and no language barriers to enable global interactions. So, CBIR has extensive and significant applications in several fields such as, scientific, medical, educational, architectural design, Ministry of Justice etc. The feature is the key basis for CBIR, which is a visual image property. The function is universal

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

for the whole or basic image for a small pixel group. Depending on the method implemented in the CBIR, functions are categorized into low-level features and high-level features. Low-level features are used to remove the distinction between objects in the world and information in the description of the scene. High-level features are used to eradicate differences between info that can be retrieved from visible data & translations, where same data is delivered to users provided in a given condition. In general low-level features include objects that reflect color, shape or texture & spatial points in the image [2]. Due to its reliability, efficiency, ease of implementation and low color storage requirements, the most powerful features and most CBIR systems use color. Texture features are another feature that is used in CBIR, which aims to encapsulate details and patterns in the image area. Some researchers aim to reduce the gaps between the visual features and the diversity of human senses. To develop the high-level features for CBIR, the object-ontology is used to determine the advanced concept. Relative feedback was added to the into the retrieval loop for learning of user intents and semantic models were produced to foundation for high-level image retrieval.Basics related to CNN are discussed in part- II. CBIR using CNN architectures are deliberated in part- III. In section IV we presented the experimental results of a system. This paper concludes in part-V.

II. BASICS RELATED TO CNN

Basic architecture for a CNN consists of mainly convolutional layer, pooling layer and fully connected layers.

Convolutional layers:

It is the main building block for CNN architectures. In convolutional layer, convolution operation occurs in between each block of an image and the filter. To do this convolution operation size of the filter and size the block from an image should be same.

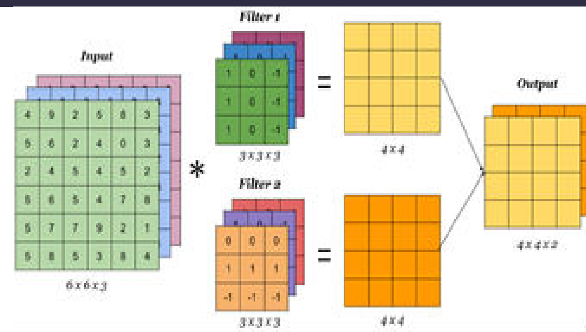
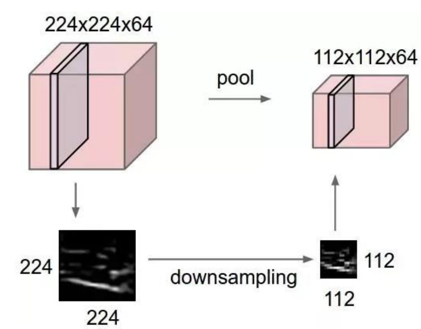


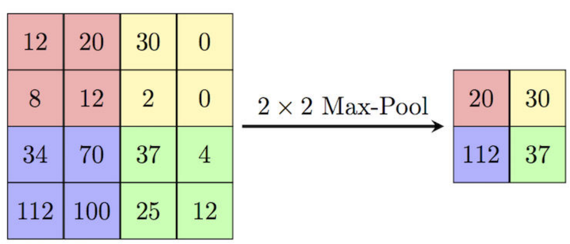
Fig 2.1: convolutional layer operation

Pooling layer:

The important layer in CNN is pooling. It is one of the methods to reduce the size of an image without loss of the most important features. The operation occurs in pooling layer is divide an image into small 2 × 2 blocks and take 1 each pixel from each block of an image. If the outcome pixel is the largest of all pixels in a block then it is known as max -pooling.



(A)



(B)

Fig 2.2: process for Max-pooling layer

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

III. CNN ARCHITECTURES

In this paper we use 2 different CNN architectures i.e., VGG19 model & INCEPTION models.VGG19 model having 19layers.It contains successive convolutional layers followed by polling layers.VGG19 architecture is as shown below.

In Google Net contains multiple of inception modules stacked one over other.

Because these Inception modules are placed on each other's statistics on the implications of the result of their change: as features of higher abstraction are captured by higher layers, their spatial concentration is expected to decrease. This shows that the ratio of the 3×3 and 5×5 convolutions should be increased when moving to a higher layer.

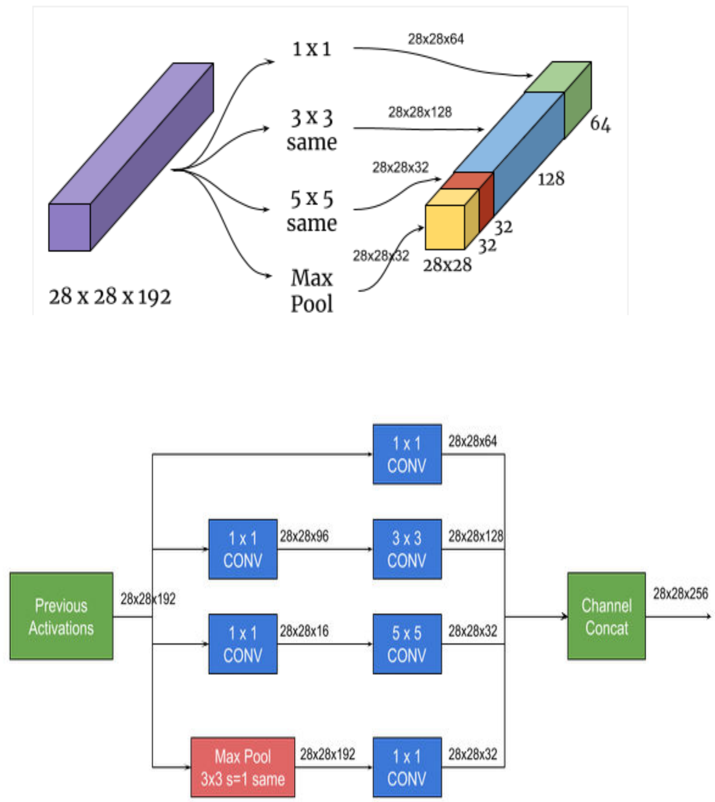


Figure 3.2 : Inception modules

After database is trained using CNN architecture and generate the feature vectors and construct the feature database from the images of a database. The extraction process has been initiated when a user asks a system using an example image or a drawing of the object. For a query image also feature vector computed using the same process that is used to create an entry database. Similarity measure is used to compute the Euclidean distance between the feature database images & the query image

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

EUCLIDEAN DISTANCE (ΔD) = $\sqrt{\sum_{i=1}^{n}(|Q_i - D_i|)^2}$

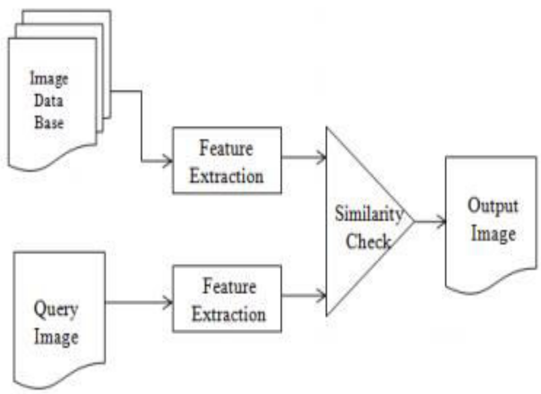


Fig 3.3: CBIR architecture

IV. EXPERIMENTAL RESULTS

We use Corel -1k database is used for testing the CNN architectures. Corel database consists of 1000 RGB color images having 10 various categories.

Figure 4.1: sample images in Corel image database

The performance of the system can be computed with the help of Recall, Precision, and F-score.

The following equations for calculating the Precision, Recall

Precision (P) = X/Y
Recall = X/Z
X=N.O of Relevant images Retrieved
Y= Total N.O of images retrieved.
Z=N.O of relevant images in the collection of Dataset.

$$precision = \frac{true\ positives}{true\ positives + false\ positives}$$

$$recall = \frac{true\ positives}{true\ positives\ + false\ negatives}$$

Fig 4.2: Retrieval results of query 'Roses' for VGG19 model

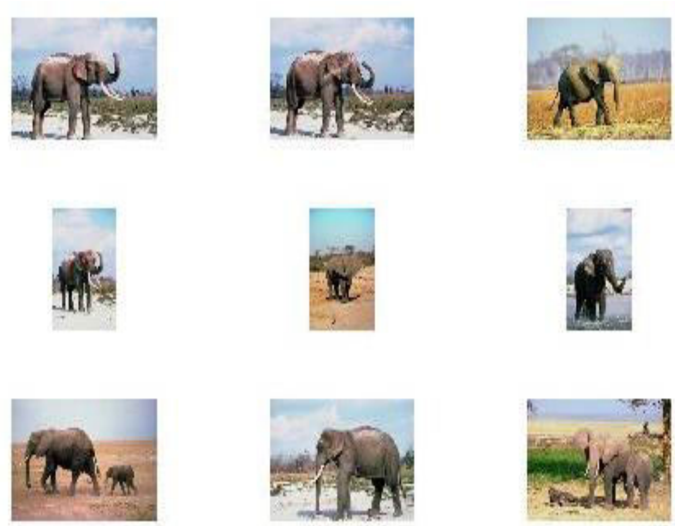
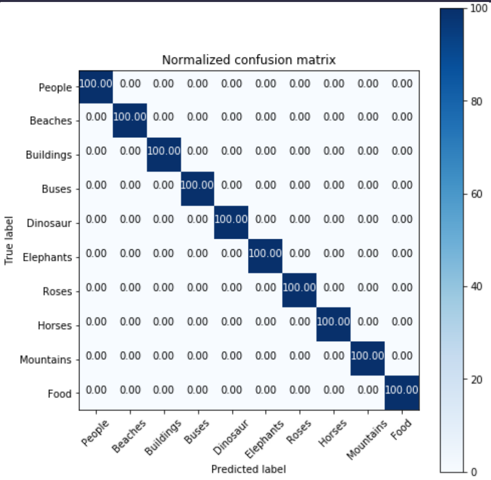
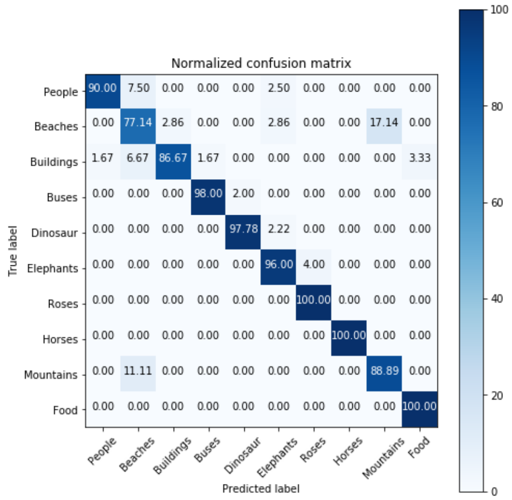


Fig 4.3: Retrival results of query 'Roses' for Inception model

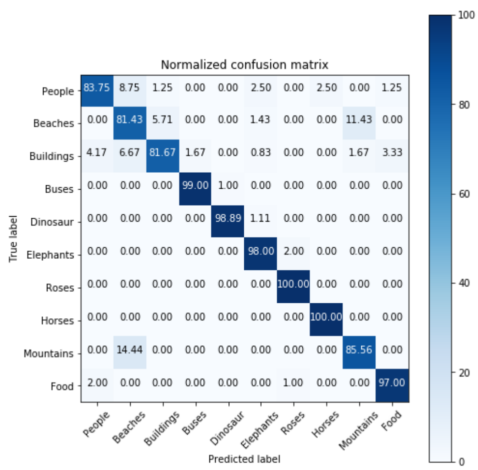
International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org



(a)

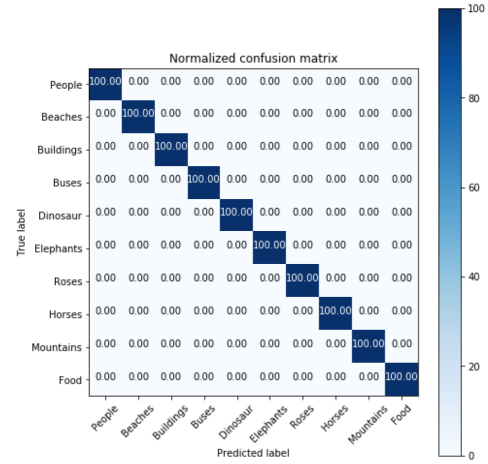


(b)

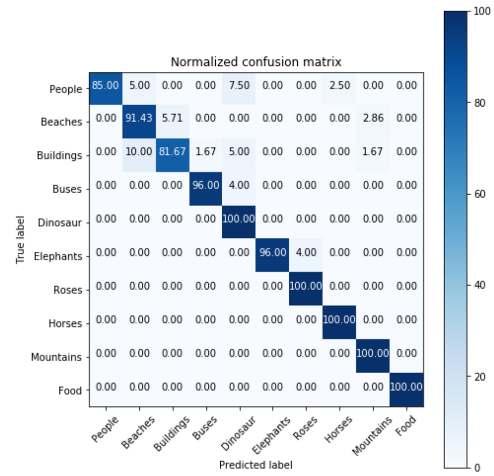


(C)

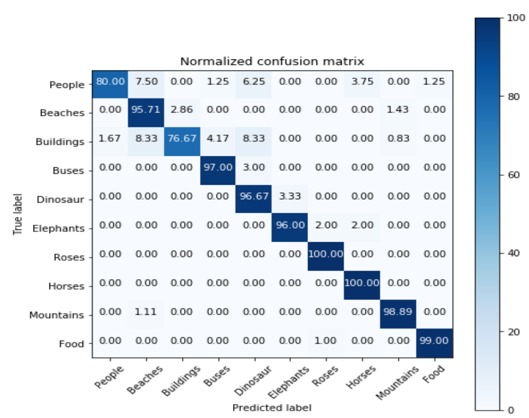
Fig: 4.4 (a),(b), (c) Confusion Matrix --- Top -1,Top- 5,Top - 10 images for VGG19 model



(a)



(b)



(c)

Fig: 4.5 (a),(b) ,(c) Confusion Matrix --- Top -1,Top -5,Top- 10 images for Inception model

Table: precision (%) for Corel Dataset (VGG19 architecture)

categories	Top- '1'	Top-'5'	Top - '10'
People	100	98.21	93.13
Beach	100	75.31	73.16
Buildings	100	96.8	92.14
Buses	100	98.32	98.34
Dinosaurs	100	97.99	98.99
Elephants	100	92.68	94.34
Flowers	100	96.15	97.08
Horses	100	100	97.56
Mountains	100	83.83	86.72
Food	100	96.77	95.49

Table: Recall (%) for Corel Dataset (VGG19 architecture)

categories	Top -'1'	Top- '5'	Top- '10'
People	100	90	83.75
Beach	100	77.14	81.43
Buildings	100	70	81.67
Buses	100	97.14	99
Dinosaurs	100	97.14	98.89
Elephants	100	80	98
Flowers	100	100	100
Horses	100	100	100
Mountains	100	80	85.56
Food	100	75.71	97

Table: precision (%) for Corel Dataset (Inception architecture)

categories	Top '1'	Top- '5'	Top- '10'
People	100	100	97.95
Beach	100	85.9	84.96
Buildings	100	93.46	96.4
Buses	100	98.29	94.7
Dinosaurs	100	85.83	84.61
Elephants	100	100	96.64
Flowers	100	96.15	97.08
Horses	100	97.56	94.56
Mountains	100	95.66	97.76
Food	100	100	98.75

Table: Recall (%) for Corel Dataset (Inception architecture)

categories	Top- '1'	Top -'5'	Top- '10'
People	100	85	80
Beach	100	91.43	95.75
Buildings	100	81.67	76.67
Buses	100	96	97
Dinosaurs	100	100	96.67
Elephants	100	96	96
Flowers	100	100	100
Horses	100	100	100
Mountains	100	100	98.89
Food	100	100	99

VI. CONCLUSION

In this paper we implement CBIR systems using VGG19 & inception CNN models and its comparison. When compared to inception, VGG19 needs more computational time as well as memory due to their huge width of the convolutional layers and also inception yields good performance in terms of precision and recall on the Corel-1k.

REFERENCES

[1] Arnold. W. M. Smeulders, M. Worring, S. Satini, A. Gupta, R. Jain. Content – Based Image Retrieval at the end of the Early Years,

IEEE Transactions on Pattern analysis and Machine Intelligence, Vol. 22, No. 12, pp 1349-1380 , 2000.

[2] Gupta. A. Visual Information Retrieval Technology: A Virage erspective, Virage Image Engine. API Specification, 1997.

[3] Ping-Cheng Hsieh,Pi-Cheng Tung,A Novel Hybrid Approach Based On Subpattern Technique and Whitened PCA for Face Recognition, Pattern Recognition 42 (2009) 978-984.

[4] Vytautas Perlibakas Distance measures for PCA based face recognition, Pattern Recognition letters 25 (2004) 711-724

[5] S.I. Choi, C. Kim, C.H. Choi, Shadow compensation in 2D images for face recognition, Pattern Recognition 40 (2007) 2118–2125.

[6] A. Pentland ,M. Turk, , Eigenfaces for recognition, J. Cognitive Neurosci. 3 (1) (1991) 71–86.

[7] Yalefacedatabase

http://cvc.yale.edu/projects/yalefaces/yalefaces.html.

[8] J. Yang, D. Zhang, , Is ICA significantly better than PCA for face recognition? in: Proceedings of IEEE International Conference on Computer Vision, vol. 1, 2005, pp. 198–203.

[9] T.J. Sejnowski, A.J. Bell, The independent components of natural scenes are edge filters, Vision Res. 37 (23) (1997) 3327–3338.

[10] H.M. Lades,M.S. Bartlett, , T.J. Sejnowski, Independent component representation for face recognition, in: Proceedings of SPIE Symposium on Electronic Imaging: Science and Technology, 1998, pp. 528–539.

[11] S. Chen, Y. Zhu, Subpattern-based principle component analysis, Pattern Recognition 37 (2004) 1081–1083.

[12] M.Safayani, M.T.Manzuri Shalmani, and M.Khademi. Extended two- dimensional pca for efficient face representation and recognition,ICCP pages 295–298, Aug 2008.

[13] R. Gottumukkal and V.K. Asari. An improved face recognition technique based on modular pca approach. *Pattern Recogn.Lett.*, 25(4):429–436, 2004.

[14] S.Chen and Y.Zhu. Subpattern-based principle component analysis *Pattern Recognition*, 37:1081–1083, 2005.

[15] K.Tan and S.Chen. Adaptively weighted sub-pattern pca for face recognition. *Neurocomputing*, 64:505–511, 2005.

[16] R. Gottumukkal and V.K. Asari. An improved face recognition technique based on modular pca approach. *Pattern Recogn.Lett.*, 25(4):429–436, 2004.

[17] Yang J and Zhang D. Two-dimensional pca:a new approach to appearance-based face representation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(1):131–137, 2004.

[18] S. H. Ou, Q. L. Wang, and Z. Y. Zhu. *The Application and technology of Digital Image Processing(In Chinese)*. Tsinghua Press,Bejjing, China, 2004.

[19] Vytautas Perlibakas “distance measure for PCA-based face recognition. Pattern Recognition Letters 25 (2004) 711–724

[20] Gupta Navarrete, P., Ruiz-del-Solar, J., 2001. Eigenspace-based recognition of faces: Comparisons and a new approach. In: International Conference on Image Analysis and Processing ICIAP2001. pp. 42–47.

[21] Graham, D.B., Allinson, G.N.M., 1998. Characterizing virtual Eigen signatures for

general purpose face recognition. Face recognition: From theory to applications. NATO ASI Series F, Computer and Systems Sciences 163, 446–456.

[22] Swets, D.L., Pathak, Y., Weng, J.J., 1998. An image database system with support for traditional alpha numeric queries and content-based queries by example. Multimedia Tools Appl,(7),181-212.

[23] Yilmaz, A., Gokmen, M., 2001. Eigen hill vs. eigen face and eigen edge. Pattern Recognition 34, 181–184.

[24] D.Q. Zhang, S.C. Chen, J. Liu, Representing image matrices: Eigen images vs. eigenvectors, in: Proceedings of the Second International Symposium on Neural Networks (ISNN'05), Lecture Notes in Computer Science, vol. 3497, Chongqing, China, 2005, pp. 659–664.