International Journal of Interdisciplinary Innovative Research &Development (IJIIRD)
ISSN: 2456-236X
Vol. **06** Special Issue **01 | 2021**

A Review on Deep Learning for Content-Based Image Retrieval

Vrushali Anant Chopde[1], Prof. Y.P.Sushir[2]
*[1]*Student, Dept of ENTC, Padm.Dr.V.B.Kolte College of Engineering, Malkapur, SGBAU, India.*
*[2]*Asst. Professo, Dept of ENTC, Padm.Dr.V.B.Kolte College of Engineering, Malkapur, SGBAU, India.*

ABSTRACT

Learning effective feature representations and similarity measures are crucial to the retrieval performance of a content-based image retrieval (CBIR) system. Despite extensive research efforts for decades, it remains one of the most challenging open problems that considerably hinders the successes of real-world CBIR systems. The key challenge has been attributed to the well-known "semantic gap" issue that exists between low-level image pixels captured by machines and high-level semantic concepts perceived by human. Among various techniques, machine learning has been actively investigated as a possible direction to bridge the semantic gap in the long term. Inspired by recent successes of deep learning techniques for computer vision and other applications, in this paper, we attempt to address an open problem: if deep learning is a hope for bridging the semantic gap in CBIR and how much improvements in CBIR tasks can be achieved by exploring the state-of-the-art deep learning techniques for learning feature representations and similarity measures. Specifically, we investigate a framework of deep learning with application to CBIR tasks with an extensive set of empirical studies by examining a state-of-the-art deep learning method (Convolutional Neural Networks) for CBIR tasks under varied settings. From our empirical studies, we find some encouraging results and summarize some important insights for future research.

***Keywords**: Deep Learning; Content-Based Image Retrieval; Convolution Neural Networks; Feature Representation*

1. INTRODUCTION

The retrieval performance of a content-based image retrieval system crucially depends on the feature representation and similarity measurement, which have been extensively studied by multimedia researchers for decades. Although a variety of techniques have been proposed, it remains one of the most challenging problems in current content-based image retrieval (CBIR) research, which is mainly due to the well-known "semantic gap" issue that exists between low-level image pixels captured by machines and high-level semantic concepts perceived by human. From a high-level perspective, such challenge can be rooted to the fundamental challenge of Artificial Intelligence (AI), that is, how to build and train intelligent machines like human to tackle real-world tasks. Machine learning is one promising technique that attempts to address this grand challenge in the long term.

Recent years have witnessed some important advances of new techniques in machine learning. One important breakthrough technique is known as "deep learning", which includes a family of machine learning algorithms that attempt to model high-level abstractions in data by employing deep architectures composed of multiple non-linear transformations . Unlike conventional machine learning methods that are often using "shallow" architectures, deep learning mimics the human brain that is organized in a deep architecture and processes information through multiple stages of transformation and representation. By exploring deep architectures to learn features at multiple level of abstracts from data automatically, deep learning methods allow a system to learn complex functions that directly map raw sensory input data to the output, without relying on human-crafted features using domain knowledge. Many recent studies have reported encouraging results for applying deep learning techniques to a variety of applications, including speech recognition, object recognition, and natural language processing.

1.1Content Based Image Retrieval System

The Content of the image in terms of colour, texture and shape is the heart of content based image retrieval system. Figure 2 shows the simple structure of content based image retrieval system. Image is retrieved based on similarity matched between the query image and database images. Every image is comprised of colour, texture, shape as well as low level and high level features. The characteristics of an image retrieval that is based on low-level features can be obtained directly from the image whereas it hard to solve arithmetic problems in high level features. The image retrieval is made by three primal techniques that is image retrieval by colour, image retrieval by texture and image retrieval by shape. This paper provides a particular lane to use this feature to retrieve the desired image.

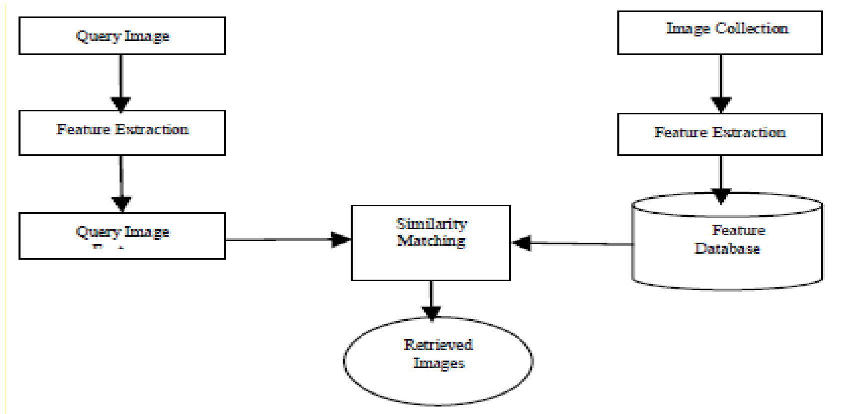

Fig.1 Representation of CBIR system

1.2 Image Retrieval by Color

Color feature is a sensitive and understandable feature of the image, and normally histogram techniques are used to demonstrate it. The advantage of Colour histogram technique is that, it has high speed, it does not require huge memory and not susceptible with the change in images size and other parameters. On the basis of Hue Saturation Value of color space, the color feature vector of query image and database image is calculated. To increase the accuracy of image retrieval, histogram technique is treated to be superior to Hue Saturation Value. R.VijayaArjunan and Dr.V.Vijaya Kumar have proposed the technique to obtain the color histogram.

Following steps are involved in retrieval of image by using color features

1. Feature Extraction
2. Histogram calculation
3. Similarity Matrix calculation
4. Dissimilarity calculation
5. Arrangement of Images in ascending order.

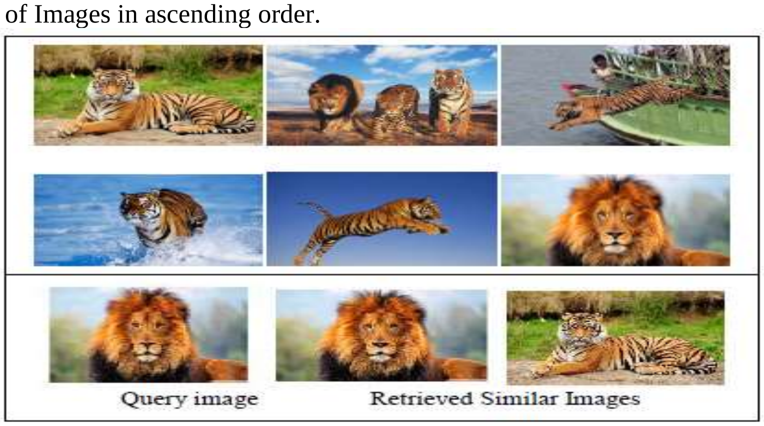

Fig.2 Image retrieving using colour histogram.

1.3 Image Retrieval by Texture

Texture is used to illustrate a basic constituent ofimages which are arranged uniformly.The region of an imagecan be determined by texture segmentation. After finding theregion of image, their bounding boxes may be used to retrievethe formation like an R-tree. The problem occurring withdimension and cross correlation also affects the texture and itcan be solved by comparable technique used in colourretrieval. Periodicity and scale are the merits of texture.Texture has qualities like periodicity and scale; it can beexpressed in terms of Contrast, direction and thickness. Thenatural possessions of surface is texture and it demonstratevisual pattern, it include significant information associatedto the structural arrangement of the surface like buildings, sea,plants, textile. Accessing the related image by using texturetechnique, two most important issues involved namelyarithmetic analysis and structural analysis. When the texturecomponent can be clearly identified, structural analysis isused whereas for micro texture arithmetic analysis is used. Arithmetic process distinguishes dissimilarity of power in atexture window. For example process contains contrast (highcontrast tiger skin vs. low contrast monkey skin), coarseness(dense stones vs. coarse gravel), and directionality(directedcloth vs. undirected grass). Texture element in the imageobtained from structural technique, identifies their shapes andcalculates their placement policies. This set of laws are usedto illustrate how texture element are to be found in image anddetermined the number of instantaneous neighbor, number ofunit space element and whether they are layed out equally ornot.

1.4 Image Retrieval by Shape

Next important method for retrieving image is byusing its shape feature. Shape representations can be usuallydivided into two categories, Region based and Boundarybased. Simply the external edge of the shape is used inBoundary based representation. In this method, the outerdescription of the region, like the pixels near the object edgeis measured. However another technique is completelydifferent from previous method. Region based method usesthe complete shape area of image by relating the consideredarea via its inner quality; i.e., the pixels enclosed in that area.

1.5 Deep Learning

Deep learning has gained massive popularity in scientific computing, and its algorithms are widely used by industries that solve complex problems. All deep learning algorithms use different types of neural networks to perform specific tasks. This article examines essential artificial neural networks and how deep learning algorithms work to mimic the human brain. It covers the following topics

1.5.1 What is Deep Learning?

Deep learning uses artificial neural networks to perform sophisticated computations on large amounts of data. It is a type of machine learning that works based on the structure and function of the human brain.
Deep learning algorithms train machines by learning from examples. Industries such as health care, eCommerce, entertainment, and advertising commonly use deep learning.

1.5.2 Defining Neural Networks

A neural network is structured like the human brain and consists of artificial neurons, also known as nodes. These nodes are stacked next to each other in three layers:

- The input layer
- The hidden layer(s)
- The output layer

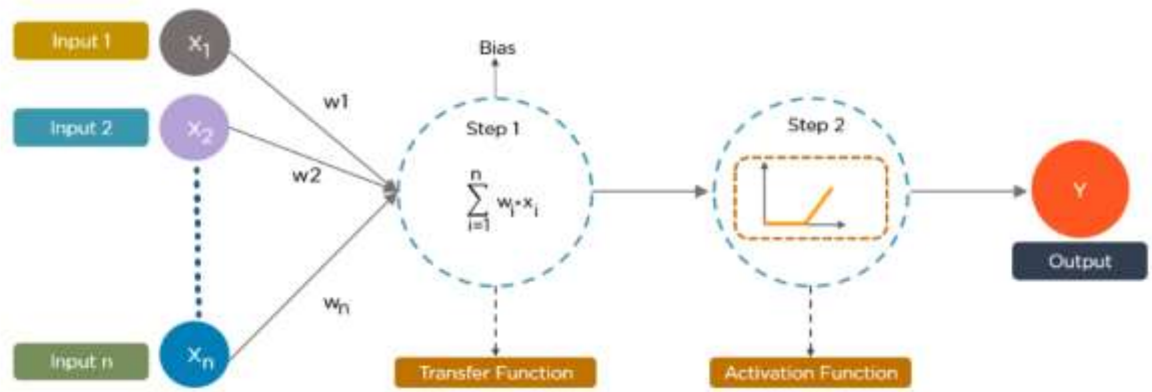

Fig.3 Image retrieving using color histogram.

Data provides each node with information in the form of inputs. The node multiplies the inputs with random weights, calculates them, and adds a bias. Finally, nonlinear functions, also known as activation functions, are applied to determine which neuron to fire. It refers to a class of machine learning techniques, where many layers of information processing stages in hierarchical architectures are exploited for pattern classification and for feature or representation learning. It lies in the intersections of several research areas, including neural networks, graphical modeling, optimization, pattern recognition, and signal processing, etc. Deep learning has a long history, and its basic concept is originated from artificial neural network research. The feed-forward neural networks with many hidden layers are indeed a good example of the models with a deep architecture. Back-propagation, popularized in 1980's, has been a well-known algorithm for learning the weights of these networks.

Recently, it has become a hot research topic in both computer vision and machine learning, where deep learning techniques achieve stat-of-the art performance for various tasks. The deep convolution neural networks (CNNs) came out first in the image classification task of ILSVRC-2012.Deep learning is one of the classifications of soft computing phenomenon in which extraction of data from millions of segregated images can be retrieved using this phenomenon. The retrieval performance of a content-based image retrieval system crucially depends on the feature representation and similarity measurement, which have been extensively studied by multimedia researchers for decades Although a variety of techniques have been proposed, it remains one of the most challenging problems in current content-based image retrieval (CBIR) research, which is mainly due to the well-known" semantic gap" issue that exists between low-level image pixels captured by machines and high-level semantic concept perceived by humans. From a high-level perspective, such challenge can be rooted to the fundamental challenge of artificial intelligence (AI) that is,how to build and train intelligent machines like human to tackle real-world tasks.

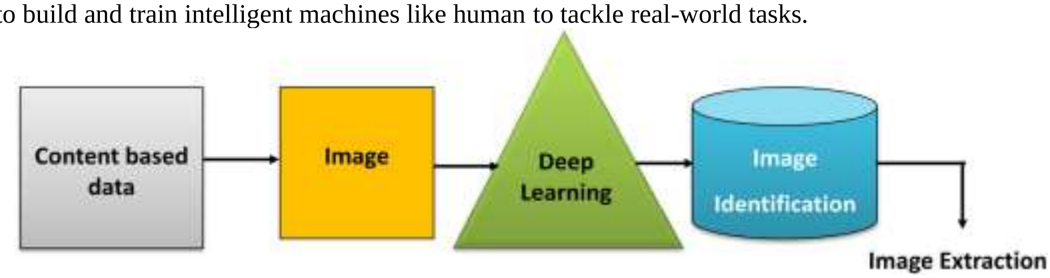

Fig.7 Content based architecture

Machine learning is one promising technique that attempts to address this challenge in the long term. Recent years have witnessed some important advanced new techniques in machine learning. Deep learning is the part of machine learning, which includes a family of machine learning algorithms that attempt to model high-level abstractions in data by employing deep architectures composed of multiple on linear transformations. Unlike traditional machine learning techniques that are often using "shallow" architectures, deep learning mimics the human brain that is organized in a deep architecture and processes information through multiple stages of transformation and representation. By exploring deep architecture features at multiple levels of abstracts from data automatically, deep learning methods allow a system to learn complex functions that directly map raw sensory input datum to the output, without relying on human-crafted features using domain knowledge. Many recent studies have reported encouraging results for applying deep learning techniques to a variety of applications, including speech recognition, object recognition, and natural language processing, among others.

2. ANALYSIS OF PROBLEM

Inspired by the successes of deep learning, in this paper, we attempt to explore deep learning techniques with application to CBIR tasks. Despite much research attention of applying deep learning for image classification and recognition in computer vision, there is still limited amount of attention focusing on the CBIR applications.
We try investigate deep learning methods for learning feature representations from images and their similarity measures towards CBIR tasks. In particular, we aim to address the following open research questions:

(i) Deep learning methods effective for learning good feature representations from images to tackle CBIR tasks
(ii) Improvements can be achieved by deep learning techniques when compared with traditional features crafted by experts in multimedia and computer vision
(iii) Apply and adapt an existing deep learning model trained in one domain to a new CBIR task in another domain effectively

In order to answer the above questions, we investigate a framework of deep learning for content-based image retrieval (CBIR) by applying a state-of-the-art deep learning method, that is, convolution neural networks (CNNs) for learning feature representations from image data, and conduct an extensive set of empirical studies for a variety of CBIR tasks. From the empirical studies, we obtain some encouraging results and reveal several important insights for addressing the open questions. As a summary, we make the following major contributions in this work:

- We introduce a deep learning framework for CBIR by training large-scale deep convolution neural networks for learning effective feature representations of images;
- We conduct an extensive set of empirical studies for comprehensive evaluations of deep convolution neural networks with application to learn feature representations for a variety of CBIR tasks under varied settings.

3. PROPOSED METHOD

In our method, a multi-feature image retrieval method is introduced by combining the features of color histogram, edge, edge directions, edge histogram and texture features, etc. In this model, the content based image will be extracted from a collection of intended image groups. After performing some pre-processing steps like selection removal, its above features are extracted and are stored as small signature files. Similar images should have similar signatures. These signatures are compared with the content based signature. During the similarity measure, the distances between the different features are measured. Appropriate weights are applied to normalize the distance coefficients. These normalized coefficients are sorted and indexed based on the distance values and their optimized state of functioning.

4. CONCLUSION

Inspired by recent successes of deep learning techniques, we attempt to address the long-standing fundamental feature representation problem in Content-based Image Retrieval (CBIR).We aim to evaluate if deep learning is a hope for bridging the semantic gap in CBIR for the long term, and how much empirical improvements in CBIR tasks can be achieved by exploring the state-of-the-art deep learning techniques for learning feature representations and similarity measures. In particular, we investigate a framework of deep learning with application to CBIR tasks with an extensive set of empirical studies by examining a state-of-the-art deep learning method (convolution neural networks) for CBIR tasks under varied settings.

REFERENCES

[1] Lew, M.S., Sebe, N., Djeraba, C., Jain, R.: Content-based multimediainformation retrieval: State of the art and challenges. ACMTrans. Multimed. Comput. Commun. Appl. **2**(1), 1–19 (2006)
[2] Bringer, J., Chabanne, H., Patey, A.: Privacy-preserving biometricidentification using secure multiparty computation: an overviewand recent trends. IEEE Signal Process. Mag. **30**(2), 42–52 (2013)
[3] Aghasaryan, A., Bouzid, M., Kostadinov, D., Kothari, M., Nandi,A.: On the use of LSH for privacy preserving personalization.In: Proceedings of the 12th IEEE International Conference Trust,Security, Privacy in Computing and Communications (TrustCom).pp. 362–371 (2013)
[4] Fanti, G., Finiasz, M., Ramchandran, K.: One-way private mediasearch on public databases: the role of signal processing. IEEESignal Process. Mag. **30**(2), 53–61 (2013)
[5] Acar, G., et al.: FPDetective: dusting the web for fingerprinters. In:Proceedings of the 2013 ACM SIGSAC Conference on Computerand Communications Security (CCS), pp. 1129–1140 (2013)
[6] Balsa, E., Troncoso, C., Diaz, C.: OB-PWS: Obfuscation-basedprivate web search. In: Proceedings of the IEEE Symposium onSecurity and Privacy, pp. 491–505 (2012)
[7] Meunkaewjinda, A., Kumsawat, P., Attakitmongcol, K., Srikaew, A., 2008. Grape leaf disease detection from color imagery using hybrid intelligent system. In: Proceedingsof ECTI-CON 2008. IEEE, pp. 513e516.
[8] Gulhane, Viraj A., Gurjar, Ajay, 2011. Detection of diseases on cotton leaves and its possible diagnosis. Int. J. Image Process. 5 (5), 590e598.
[9] Kebapci, Hanife, Yanikoglu, Berrin, Unal, Gozde, 2011. Plant image retrieval using color, shape and texture features. Comput. J. 54 (9), 1475e1490.
[10] Li, Jinghui, Gao, Lingwang, Shen, Zuorui, 2010. Extraction and analysis of digital images feature of three kinds of wheat diseases. In: International Congress onImage and Signal Processing. IEEE, pp. 2543e2548.
[11] SathyaBama, B., MohanVali, S., Raju, S., Abhai Kumar, V., 2011. Content based leaf image retrieval (CBLIR) using shape, color and texture features. Indian J. Comput.Sci. Eng. 2 (2), 202e211.
[12] Kulkarni, Anand H., Patil, Ashwin, 2012. Applying image processing technique to detect plant diseases. Int. J. Mod. Eng. Res. 2 (5), 3661e3664.
[13] Kailey, Kamaljot Singh, Sahdra, Gurjinder Singh, 2012. Content Based Image Retrieval (CBIR) for identifying image based plant disease. Comput. Technol. Appl. 3 (3), 1099e1104.Conference on Machine Learning and Cybernetics (ICMLC), pp. 719e724.
[14] Guo-Dong Guo, Anil K.Jain, Wei-Ying Ma and HongJiang, ”Learning similarity measure for natural image retrieval with relevance feedback”, IEEE transactions on neural networks,Vol-13, No.4, 2002
[15] Erkin, Z.: Protection and retrieval of encrypted multimedia content:when cryptography meets signal processing. EURASIP J. Inf.Secur. **2007**, 20 (2007)
[16] Lagendijk, R.L., Erkin, Z., Barni, M.: Encrypted signal processingfor privacy protection: conveying the utility of homomorphicencryption and multiparty computation. IEEESignal Process. Mag.**30**(1), 82–105 (2013)
[17] Shashank, J., Kowshik, P., Srinathan, K., Jawahar, C.V.: Privatecontent based image retrieval. In: Proceedings of the IEEE Conferenceon Computer Vision and Pattern Recognition (CVPR), pp.1–8 (2008)
[18] Sabbu, P.R.,Ganugula, U.,Kannan, S., Bezawada,B.: An obliviousimage retrieval protocol. In: Proceedings of the IEEE InternationalWorkshops ofAdvanced InformationNetworking and Applications(WAINA), pp. 349–354 (2011)
[19] Erkin, Z., Franz,M., Guajardo, J., Katzenbeisser, S., Lagendijk, I.,Toft, T.: Privacy-preserving face recognition. In: Proceedings of the9th International Symposium on Privacy Enhancing Technologies(PETS), pp. 235–253 (2009)
[20] Sadeghi, A.-R, Schneider, T., Wehrenberg, I.: Efficient privacypreservingface recognition. In: Proceedings of the 12th InternationalConference on Information Security and Cryptology(ICISC), pp. 229–244 (2009)
[21] Osadchy, M., Pinkas, B., Jarrous, A., Moskovich, B.: SCiFI—asystem for secure face identification. In: Proceedings of the IEEESymposium on Security and Privacy (SP), pp. 239–254 (2010)
[22] Barni, M., Failla, P., Lazzeretti, R., Sadeghi, A., Schneider, T.:Privacy-preserving ECG classification with branching programsand neural networks. IEEE Trans. Inf. Forensics Secur. **6**(2), 452–468 (2011)

[23] Hsu, C.-Y., Lu, C.-S., Pei, S.-C.: Image feature extraction inencrypted domain with privacy-preserving SIFT. IEEE Trans.Image Process. **21**(11), 4593–4607 (2012)
[24] Annamalai, R., Srikanth, J.: Integrity and privacy sustenance ofshared large scale images in the cloud by ring signature. Int. J.Comput. Appl. **114**(12), 13–18 (2015)
[25] Agrawal, R., Srikant, R.: Privacy-preserving data mining. In: Proceedingsof the ACM SIGMOD International Conference on theManagement of Data, pp. 439–450 (2000)
[26] Duchi, J.C., Jordan, M.I., Wainwright, M.J.: Advances in NeuralInformation Processing Systems 25. Privacy aware learning, pp.1430–1438. Curran Associates, Red Hook (2012)
[27] Rane, S., Boufounos, P.T.: Privacy-preserving nearest neighbormethods: comparing signals without revealing them. IEEE SignalProcess. Mag. **30**(2), 18–28 (2013)
[28] Lu, W., Swaminathan, A., Varna, A.L., Wu, M.: Enabling searchover encryptedmultimedia databases. In: Proceedings of the SPIE,Media Forensics and Security, vol. 7254, pp. 725418-1–725418-11,(2009)