

Visual Information Retrieval using Deep learning with Visual Attention Model

T.Venkata Raman

Research Scholar, JNTU Hyderabad, Telangana, India

K.Venugopal Rao

Professor & HOD CSE Department, Narayanamma Engineering College, Hyderabad

Abstract. In to-days world there is a huge growth of visual data. These data is increasing beyond the capacity of human brain. All the existing concepts like CBIR, Image Retrieval models are not meeting the current requirements. To overcome these difficulties in the field of visual information retrieval, we are proposing the solution by constructing “visual attention model” by taking the saliency features of the object. To implement this concept and to have the high accuracy and to improve the retrieval efficiency, we are implementing the Deep learning features in our work. Deep learning makes our work simple in moving further to achieve the result set. This paper discusses about the construction of visual attention model, Saliency features, Deep learning and the proposed architecture which will be implemented in achieving high accuracy when compared to existing models.

Keywords: Visual attention model, saliency model, CBIR, similarity measures, Image retrieval.

I. INTRODUCTION

Day by day there is huge growth in the visual world but at the same moment the processing capacity of the human brain is not reaching to visualize the world. To analyze an object from the group of objects or to compare one object with the other object normal visual features are not enough to identify or to differentiate between the objects. To overcome such problems, saliency features of an object is considered. With the help of saliency feature we construct a “Visual attention model” which takes the support of deep learning to extract the saliency features and improves the retrieval accuracy when compared to existing ones.

Saliency model construction: This model is constructed based on the Triesman’s feature integration theory[1] and also based on the Julesz’ Texton theory [3], by only considering traditional features of CBIR like Color, intensity and orientation are considered as the primary visual features which are commonly used in many saliency models, apart from these some other pop-up features are also considered like color volume, with edge information together which are used to detect saliency regions.

Feature set of an object or an image which have been designed manually may be over specified, ambiguous, incomplete and may take long time to process or to validate them. Deep learning is the mechanism which provides very flexible, easily understandable and learnable frameworks for representing the visual world. Deep learning is also used in un supervised data like raw text, audio, images, videos etc and also in supervised content. This is the main reason of having Deep Learning as part of our implementation.

II. RELATED WORK

In my previous paper, I had a survey on Visual attention models and also on Saliency features. My previous paper explains the disadvantages of CBIR and also covered the importance of Visual attention models. With my previous work I am very clear that CBIR is not a correct approach in reaching retrieval efficiency because it only focuses on image features.

CBIR system is one of such experts systems that highly rely on appropriate extraction of features and similarity measures used for retrieval [5]. This approach only used for traditional feature extraction. The area has gained wide range of attention from researchers to investigate various adopted methodologies, their drawbacks, research scope, etc [6-12].

To implement visual attention model several factors to be considered like similarity measures, Saliency model. Even I have presented some analysis of different concepts which will be used in improving the image retrieval efficiency. After analyzing it was understood that there exists some gap to concentrate in increasing the effectiveness of image retrievals. In accomplishing the gap my paper presented a kind of scope which improvises the

performance issues in image retrievals[2]. Now, in this paper I am considering another concept which is known as Deep Learning which makes some issues in simpler way in achieving the retrieval efficiency.

Now a days “ visual attention model” is gaining the popularity in the field of visual information retrieval. Majority of these models are built with the help of integration of the features of the object[1] and some other models are also built with guided search. Out of different saliency models observed. Itti’s Saliency model became popular in achieving success on image comparisons. Saliency model will be constructed with the physical and distinctive features of the object. These features are relative to one another. Visual attention models will also be constructed with the help of visual features of an image like color, shape, intensity etc. Tsotsos et al. proposed a model by using these features and tuning these features accordingly.

III. PROPOSED ARCHITECTURE AND SOLUTION

In the proposed Architecture, four major components play key roles. They are Saliency model, higher order statistics, Deep learning and decision rules library. The architecture diagram of the proposed system is illustrated in Figure 1. Being motivated by the saliency model proposed in [4], I have intended to use this model to represent the database images as well as query image. The resultant image will be subjected to determine the higher order statistical relationship persists in the image intensity and hence the feature vectors will be constructed. Based on the extracted features, decision rules library will be constructed in which the rules to determine the visually similar images for the given query image will be stored. Such rules will be extracted using the context of decision tree. The class decision block of the image retrieval phase will identify the class label of the query based on the decision rules and hence the images of similar class will be retrieved. I have even used deep learning to know the retrieval efficiency between using saliency model and deep learning.

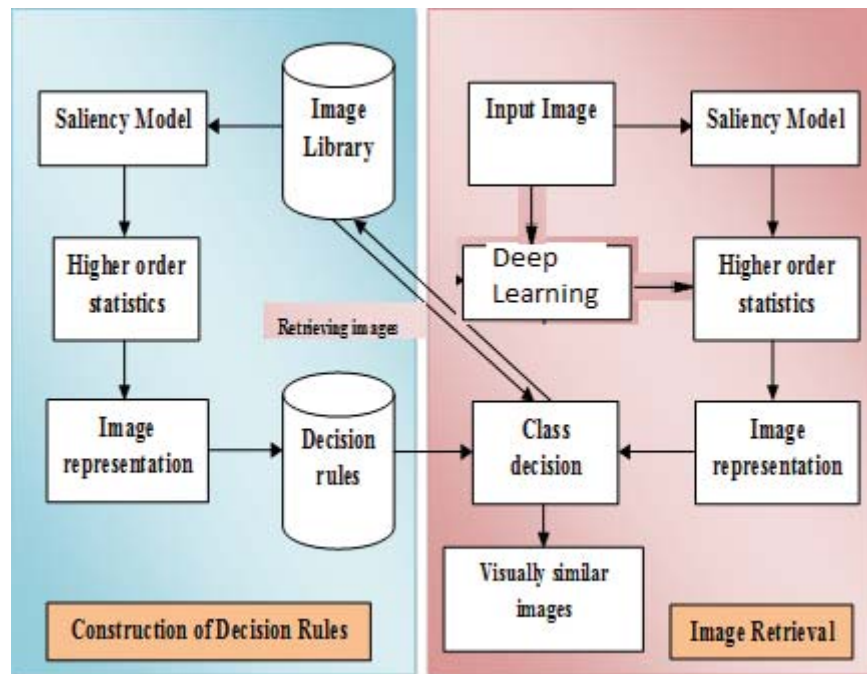


Figure 1: Proposed Architecture for Visual attention model based on Saliency model and Deep learning

Implementation Steps:

The implementation steps are described in Figure 1. Yet, the stepwise processes are described below.

Step 1: Construct saliency model for every image in the image database. The saliency model is constructed using saliency structure histogram as described in [4].

Step 2: For every saliency model, which is constructed for every database image, higher order statistics such as energy entropy, higher order deviation, etc, will be determined.

Step 3: Even by using Deep learning also feature set construction and based on that images are identified and given to the metrics.

Step 4: The statistical metrics will be combined to represent the image in low dimensional format.

Step 5: Hence defined images are constructed as rules for which the class label is given as the decision of the rule. For example, if F1 is high, F2 is low, and then I refer to tiger, where F1, F2 and I refer to the feature 1, feature 2 and the image, respectively.

Step 6: The rules construct the decision tree classifier where in it decides the image classification.

IV.CONCLUSION

The entire methodology will be developed in MATLAB and by using R-Programming. The performance of the implemented model will be investigated by making Comparisons will be made against existing models like distance based similarity measures to ensure the performance of the proposed methodology over the conventional methodologies. The performance of the methodology will be quantified using renowned performance metrics such as efficiency, precision, recall and F-scores.

REFERENCES

- [1] A. Treisman., A feature in integration theory of attention, *Cogn. Psychol.* 12 (1) (1980) 97–136..
- [2] T.venkata Ramana, K.V.G. Rao, G.Shyama Chandra Prasad, “Literature Survey to Improve Image Retrieval Efficiency by Visual Attention Model”, Springer conference on *Information Systems Design and Intelligent Applications*, Volume 434 of the series *Advances in Intelligent Systems and Computing*, pp 121-127.
- [3] B. Julesz, Texton gradients: the texton theory revisited, *Biol. Cybern.* 54 (1986), 245–251.
- [4] Guang-Hai Liu, Jing-Yu Yang, ZuoYong Li, “Content-based image retrieval using computational visual attention model”, *Pattern Recognition*, In Press, Corrected Proof, 2015
- [5] Ja-Hwung Su ; Wei-Jyun Huang ; Yu, P.S. ; Tseng, V.S., “Efficient Relevance Feedback for Content-Based Image Retrieval by Mining User Navigation Patterns” *IEEE Transactions on Knowledge and Data Engineering*, Vol. 23, No. 3, pp. 360-372, 2011
- [6] G. Qiu, “Color Image Indexing Using BTC,” *IEEE Trans. Image Processing*, Vol. 12, No. 1, Jan. 2003.
- [7] C. H. Lin, R. T. Chen, and Y. K. Chan, “A smart content-based image retrieval system based on color and texture feature,” *Image and Vision Computing*, vol. 27, no. 6, pp. 658–665, May 2009.
- [8] N. Jhanwar, S. Chaudhuri, G. Seetharaman and B. Zavidovique, “Content based image retrieval using motif co-occurrence matrix,” *Image and Vision Computing*, vol. 22, pp. 1211–1220, Dec. 2004.
- [9] P. W. Huang and S. K. Dai, “Image retrieval by texture similarity,” *Pattern Recognition*, vol. 36, no. 3, pp. 665–679, Mar. 2003.
- [10] C. C. Lai, and Y. C. Chen, “A user-oriented image retrieval system based on interactive genetic algorithm,” *IEEE Trans. Inst. Meas.*, vol. 60, no. 10, October 2011
- [11] A. Borji, L. Itti., “State-of-the-art in visual attention modeling”, *IEEE Trans. Pattern Anal. Mach. Intell.* Vol. 35, No. 1, pp. 185–207, 2013.
- [12] Stella vetova, Ivan ivanov., “ Image Features Extraction Using The Dual-Tree Complex Wavelet transform”,- *Advances in Applied and Pure Mathematics*, 978-960-474-380-3.