AN EMPIRICAL STUDY ON IMAGE RETRIEVAL SYSTEM

[1]Shreya Kukdeja, [2]Prof. A. G. Phakatkar
[1]Student, [2]Professor
[1]Department of Computer Engineering,
[1]Pune Institute of Computer Technology, Pune, India

Abstract : Due to exponential growth in the technology of digital cameras the availability of the cameras are easily in many devices like mobile phones, Tablets and even on smart TVs. This makes the device users to capture the every moment of their life through the camera and store them implicitly or explicitly in the cloud or any storage respiratory. This leads to the huge accumulation of the images in the storage house so this creates havoc in the retrieval process of these data on the basis of the user query. So, some traditional systems are existed which are retrieving these images based on the content, semantic, text and image that nowhere near the logical behavior of the system image that nowhere near the logical behavior of the system. To tackle this many models are proposed where the accuracy is always been a big question while retrieving. So, this research paper analyzes most of the well-known past works on image retrieving methodologies to understand the techniques in proper way.

IndexTerms **- CNN, Morphology, Simialrity Measure.**

I. INTRODUCTION

People say that, a picture is worth a thousand words, which holds truer than ever, owing to the digital age and the enormous influx of cameras and smart phones that are equipped with the latest sensor hardware. These factors contribute to an exponential growth in the number of photographs. An average Smartphone user generates an upwards of 1000 photographs every year. Most of these are uploaded onto the cloud and this calls for an efficient image retrieval system that can retrieve these images as and when needed.

Image retrieval (IR) has transformed into an eminent research field in computer vision where digital image collections are swiftly being created and made available to plenty of users through the internet. Formidable increment in the compilation of images from environmental agencies, medical institutes and art museums. In the commercial sector, companies have established that are composing large collections of photographic images of real-world scenes available to users who utilize them for illustrations in advertisements articles, books and other media meant for the public at large. Astonishingly, the indexing of these particular images is still being done by hand; a human indexer segregates and inputs a set of keywords for each image. Each keyword can be enhanced by terms derived from a thesaurus that appends synonyms and other terms that previous users have tried earlier in searches that led to relevant images. Keywords can also be extracted from captions, but these are barely dependable. The commercial image providers, for the major part, are not utilizing these methods. The central consideration is that most systems call for an example image and then retrieve similar images from their databases. Actual users do not possess example images; they start with an idea, not an image. Some systems allow users to sketch the images which they needed. These systems necessitate the users to consider their intentions in mind first and therefore can only be applicable in some distinct domains, like, purchase of painting and trademark matching. Therefore, the recognition of universal classes of concepts and objects is imperative to administer automated indexing of images. An image retrieval system is one of the significant research area. Which can be used for browsing, searching and retrieving images from a large database of digital images. Most traditional and common methods of image retrieval system utilize some method of adding metadata such as captioning, keywords, or descriptions to the images so that retrieval can be performed over the annotation words.

Text Based Image Retrieval - The text based image retrieval uses the technique of summing up the metadata, such as descriptions, captioning or keywords to the images. The retrieval is occupied with the annotation words and it is the time-consuming and introduces an annotation complex while also requiring huge labors to manually annotate those images. The semantic content isnt taken into consideration in TBIR. The visual and textual content descriptors are produced from the image query and text query. The descriptors are transformed into a vector format. Similarly, visual and textual descriptors are converted and calculated into vector depiction for the images stored in the database. The vector, developed by the user query is then coordinated with the vectors. The content and text based methods return two independent lists of images with distinct weights. These two lists must be amalgamated in a significant way to give the user a combined image list.

Content Based Image Retrieval - CBIR is the application of computer vision methods to the image retrieval dilemma, that is, the complication of searching for digitally produced images in huge databases. Content-based implies that the searching method evaluates the constituents of the image in place of the metadata such as descriptions, tags or keywords associated with the image. The word content in this scenario deals with textures, shapes, colors or any other data that can be extracted from the image itself. CBIR is useful because predominant web reliant image search engines depend completely on metadata and this produces a lot of undesirable results. Also utilizing humans manually assign keywords for images in a huge database can be expensive, in efficient and may not extract every keyword that explains the image. Therefore, a system that can segregate images based on their

composition would yield more definitive results and administer better indexing. The images are retrieved only by means of segregating them by their Texture, Color, and Shape in CBIR.

Sketch Based Image Retrieval - Sketch based image retrieval uses the input as sketches and based on the sketches the relevant images are retrieved and also stated that the SBIR is still a young research area, there are many applications capable of using this retrieval paradigm, such as web searching and pattern detection. It utilizes a feature that drawing a sketch query that utilizes a very simple touch screen based technology is being explored. An innovative concept for SBIR which relies on detecting simple shapes which are called key shapes, works as a local approach, but rather than identifying key points, it identifies key shapes through which local descriptors are gauged. It relies on key shapes that permit the user to represent the design of the objects in an image which could be utilized to boost the effectively in the retrieval undertaking.

Semantic Based Image Retrieval - Concentrating on semantic analysis of images, it strengthens intelligence-assisted multimedia evaluation and uniting the gap between semantics and low level visual features. The framework executes at semantic level utilizing probable semantic labels, ceremoniously characterized as fuzzy sets, to take decisions on approaching image regions instead of visual features used conventionally. In order to emphasize its autonomy of a specific image distributed approach some researchers have altered two well acknowledged region growing algorithms, i.e., watershed and recursive shortest spanning tree, and correlated them to their conventional counterparts.

In this paper, section 2 is dedicated for literature review of past work and Finally Section 3 concludes this paper.

II. LITERATURE REVIEW

Xie [1] Proposes to consolidate retrieval algorithms and image classification into (Online Nearest-neighbor Estimation) ONE. The researchers demonstrate that, with the assistance of high-quality regional features, both classification and retrieval tasks could be accomplished with a simple NBNN search. The authors take benefit of PQ and PCA estimation along with GPU parallelization to mitigate heavy computational costs. Despite the simplicity, our algorithm achieves advanced retrieval performance and image classification.

A.Razavian,J.Sullivan [2] Presents an efficient pipeline based on ConvNets image representations, for visual instance retrieval. This approach benefits from multi-scale scheme to derive local features that take geometric invariance into precise account. For the representation, the authors preferred to use last Convolutional layer while adapting max-pooling, which made the layer accessible in terms of its dimensional but also helped to boost the performance. Throughout the experiments with five standard image retrieval datasets, the researchers demonstrated that generic ConvNet image depictions surpassed other conventional methods in all the tested cases if they are extracted appropriately. The pipeline does not depend on the bias of the dataset, and the only remedy the authors make to specially designed training data is in the PCA whitening.

Y. Kalantidis, C. Mellina [3] Outlines a hypothesized framework with cross-dimensional weighting which incorporates recent related works for aggregated deep Convolutional features. The authors propose elementary, weighting schemes which are non-parametric for channel wise and spatial weighting and present insights for their performance by visualizing and studying the distributional attributes of the layer output responses. Using this approach, the researchers report results that surpass the conventional in popular image search benchmarks.

E. Mohedano [4] Explores a cumulative strategy established on Bag of Words to encode features from Convolutional neural networks into a sparse representation for instance search. The authors achieved aggressive performance in comparison to other CNN-based representations in Oxford and Paris benchmarks, while being more extensible in terms of index size, cost of indexing, and search time. They also correlated our BoW encoding scheme with some pooling of CNN features in the far more assertive TRE CV id instance search task, and established that our method consistently and significantly performs better.

Y. Li, X. Kong Proposes a method for image retrieval called Multi-layer Order less Fusion (MOF). MOF accumulates hierarchical activations from different layers of the Convolutional Neural Network (CNN), with the activations from each layer incorporated into the (BoW) Bag-of-Words architecture separately. With the benefit that all activations from different layers can be extracted together by-passing image patches across the deep network only once, the presented MOF concept is more efficient than earlier multi-scale and multi-category feature fusion methods. The experimental results on two benchmark datasets suggest that the CNN activations from different layers are reciprocal with each other. The fusion of them under the basic BoW framework can already achieve competitive performances against other conventional algorithms [5].

T. Yu [6] Presents the Fuzzy Objects Matching (FOM) scheme, in which object proposals are exploited to detect the probable regions of the query object in reference images. To achieve high search effectiveness, from one recommendation image, the researchers factorized the feature matrix of all the object proposals into the output of sparse codes and a set of fuzzy objects in order to accomplish efficient and effective instance search. The authors conduct comprehensive experiments using object proposal representation generated from four different aggregation methods on public datasets. The outstanding precision and the efficient implementation verify the outstanding performance of the method.

Y. Uchida Proposes the application of the Fisher vector representation to binary features to improve the accuracy of binary feature based image retrieval. The researchers derived the closed-form estimation of the Fisher vectors of binary features that can be modeled according to the Bernoulli mixture. Furthermore, the authors also recommended a fast approximation method that hastens the computation of the proposed Fisher vectors with comparable performance by magnitude of the first order. The efficiency of the Fisher vectors of binary features was confirmed [7].

A. Alzubi Introduces compact bilinear architectures for diverse CBIR tasks based on CNN using both parallel feature extractors without precedent proficiency about the semantic meta-data of image contents. Image features are precisely siphoned from the activations of Convolutional layers then using the root bilinear compact pooling they are largely reduced to very low dimensional representations. The medium architecture and very deep architecture are calibrated for these CBIR tasks: object images, general contents, and landmarks [8].

L. Zheng, [9] Shares several findings with the community on the effective usage of CNN features during transfer. First, evidences accumulate that using larger images other than 224 224 yields superior accuracy. Second, the application of average pooling on the activation maps of intermediate CNN features consistently improves recognition performance over raw features. Specifically, the authors find that the pooled Conv5 feature produces superior or competitive performance to fully connected features. Finally, the combination of features across multiple CNN layers further promotes recognition accuracy, and the system is capable of pushing the modern application forward to a large margin.

H. Azizpour [10] Explores ConvNet representations trained on Image Net are becoming the standard image representation. The authors presented a systematic study, lacking until now, of how to effectively transfer such representations to new tasks. The most important elements of the study are: identify and define several factors whose settings affect transfer ability. The systematic experiments have allowed the researchers to achieve the following. First, by optimizing the identified factors it improved the conventional performance on a very diverse set of standard computer vision databases. Second, it is observed with present empirical evidence that the effectiveness of a factor is highly correlated with the distance of the target task from the source task of the trained ConvNet. Finally, the authors empirically verify that the categorical grouping and ordering of tasks involving visual recognition is meaningful as the optimal setting of the factors remain constant within each group and vary in a consistent manner across the ordering.

III. CONCLUSION

It is very important to retrieve the images for the given query image based on the morphology of the objects of the stored images. This intact unleashes more and more challenges as the size of the images are varies with high quality. So, this research paper analyzes most of the past work and come to a conclusion that there is lot to be achieved in image retrieving techniques through which time and space complexity of the system can be improved in proper manner.

REFERENCES

[1] L. Xie, R. Hong, B. Zhang, and Q. Tian, Image Classification And Retrieval Are One, in ACM International Conference on Multimedia Retrieval, 2015.

[2] A.Razavian,J.Sullivan, and S. Carlsson, Visual Instance Retrieval With Deep Convolutional Networks, ITE Transactions on Media Technology and Applications, vol. 4, no. 3, pp. 251258, 2016.

[3] Y. Kalantidis, C. Mellina, and S. Osindero, Cross-Dimensional Weighting For Aggregated Deep Convolutional Features, in European Conference on Computer Vision, 2016.

[4] E. Mohedano, K. McGuinness, N. E. OConnor, A. Salvador, F. Marques, and X. Giroi Nieto, Bags Of Local Convolutional Features For Scalable Instance Search, in ACM International Conference on Multimedia Retrieval, 2016.

[5] Y. Li, X. Kong, L. Zheng, and Q. Tian, Exploiting Hierarchical Activation Of Neural Network For Image Retrieval, in ACM International Conference on Multimedia, 2016.

[6] T. Yu, Y. Wu, S. D. Bhattacharjee, and J. Yuan, Efficient Object Instance Search Using Fuzzy Objects Matching, AAAI Conference on Artificial Intelligence, 2017.

[7] Y. Uchida, S. Sakazawa, and S. Satoh, Image Retrieval With Fisher Vectors Of Binary Features, ITE Transactions on Media Technology and Applications, vol. 4, no. 4, pp. 326336, 2016.

[8] A. Alzubi, A. Amira, and N. Ramzan, Content-Based Image Retrieval With Compact Deep Convolutional Features, Neuro computing, vol. 249,pp. 95105, 2017.

[9] L. Zheng, Y. Zhao, S. Wang, J. Wang, and Q. Tian, Good Practice In CNN Feature Transfer, 2016.

[10] H. Azizpour, A. Sharif Razavian, J. Sullivan, A. Maki, and S. Carlsson, From Generic To Specific Deep Representations For Visual Recognition, IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2015.