SCALABLE IMAGE RETRIEVAL BY MULTILABEL DEEP HASHING

Sonal U Darade

***Abstract*: In computer vision Hashing have more scope and it is used to generate a binary hash value. So, we used Deep Hashing to learn a compact binary code for single and multiple layers of projections. There are many different methods which is used for binary code generation. We used supervised and multilabel supervised learning method for binary code value.**

***Index Terms*: Scalable image search, fast similarity search, hashing, deep learning, multi-label learning.**

I. INTRODUCTION

In recent years due to the rapid growth of large scale visual search in computer vision get more attracted. So, the objective of large scale visual search is to get more accurate results from large datasets. There are many different techniques for similarity search like nearest neighbor, KNN, etc.

The basic idea of hashing is to detect a hash value of each visual object into a binary feature vector. There are two existing method which is used for binary code: data-independent and data-dependent. Many methods are used to learn hashing functions such as LSH, BRE, KMH, ITQ, MLH, SPLH, etc.

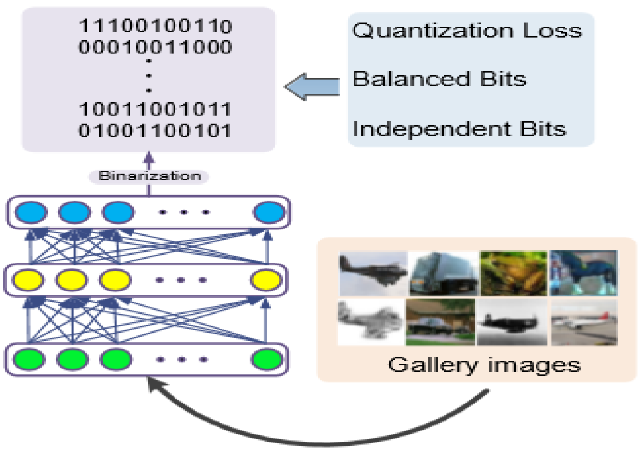

Fig.1: The basic idea of deep hashing

II. REVIEW OF LITERATURE

J. Masci et al. [1] in this paper, they proposed a new Approach which is based on a novel coupled Siamese neural network architecture. The neural network is used to map the hashing function on different modalities. Firstly, a single framework is combined by inter and intra modal similarity. Second, the compact binary code representation of data is done. Third, they solve optimization problem using the SSH methods. Fourth, the full optimization is done using the neural network Siamese under the Hashing framework.

Zhongbo Shi et al. [2] in this paper, they introduced that how the redundancy is removed from images and from inside the image. So, a generic image set compression scheme is used to remove the redundancy. In generic image set a no. of Scenes of different locations and view angles is stored. A correlations among all the images is predicated using the expected automatically generic image set. SIFT, RMSE, SSIM based prediction measurement is used for feature prediction. A different method of image set compression is used to reduce the redundancy like KLT, Centroid method, MMD, MMP, etc.

F. Shen et al. [3] in this paper, we studied that hashing gets more attention in computer application, Ml, IR, etc. The hashing techniques have no. of documents, images, videos and some sort of binary data. They propose a novel supervised Hashing framework for binary efficiency. To solve the Binary optimization problem they proposed a DCC Algorithm. In DCC Algorithm a hash code bit is generated bit by bit.

J. Wang et al. [4] in this paper, finding the similar image between two similar images is become an indispensable task. Category-level image similarity is known as having two similar image but they can only find the core similar image part. So, they proposed a Fine-grained similarity search. For finding the similarity of images first feature is extracted like Glitter's, RGB, HSV, Gray Scale, etc. A Deep Ranking Model is used to learn a fine-grained image similarity model.

F. Zhao et al. [5] in this paper, to learn hash function they proposed a framework on semantic ranking and deep learning. Hash function preserve multilevel similarity between multi-label images in the semantic space. So, to construct a hash function they used deep convolution neural network. CNN is used to learn directly from images. CNNs has achieved great success in various visual forms.

J. Wang et al. [6] in this paper, to generate a hash code recently used method is semi-supervised and supervised learning algorithm. So, they proposed a novel framework which shows the ranking-list. They focus on this 2 main methods: label-dependent hashing and ranking order statistics based hashing. In label-dependent hashing method the semi-supervised method is known as pointwise and supervised method is known as pairwise.

M. Ranzato et al. [7] in this paper, they use unsupervised learning method for object recognition. In object recognition the task is to learn an invariant features using unsupervised learning. Many unsupervised learning methods is based on encoder-decoder. To minimize the reconstruction error encoder-decoder is used. The invariant features can used two key concepts: invariant feature vector and transformation parameters.

J. Xiao et al. [8] in this paper, they proposed a Scene UNderstanding (SUN) database for storing the images. Many different algorithms performed on SUN database. We get human scenes from SUN database. Accuracy of human scenes is measured by AMT.

A.Sablayrolles et al. [9] Hashing is used in documentation, Images, Video, etc. The representation also done by hashing. In supervised hashing the label train data set is known. Train classifier is used to classify the queries. They perform both task supervised and semi-supervised hashing.

K. Lin et al. [10] content based image retrieval is used to find the similar images from database. The large database is needed in visual search due to its growth. So, they used deep learning method for the large database.

E. Liong et al. [11] in this paper, they proposed a CBFD For face detection. Unsupervised learning method is used to learn a compact binary code of face representation. Mainly they focus on 3 topics: face representation, feature learning, binary code learning.

X. Wang et al. [12] we studied that, how to search semantic similarity search. They use supervised learning method for this purpose. Supervised learning method can perform 2 task: design of hash functions and how to preserve semantic similarity.

III. SYSTEM ARCHITECTURE

A.PROBLEM STATEMENT

To propose a deep hashing (DH) approach for large database of image and video using SDH and MSDH.

To learn a binary code we used Deep Hashing (DH). Existing method is used for only single layer of projection and proposed method is used for single and multiple layer of projection.

B.SYSTEM OVERVIEW

Fig. 2 shows the architecture of the Image Processing System.

Firstly we store the database by MIRFLICKR Dataset. Then we preprocess the Set of Images using the three techniques:

1. Gray Scale
2. RGB
3. HSV

Gray Scale - This is used to calculate the Intensity Of each pixel of images. The colorimetry is used to calculate the value of gray scale.

RGB (Red, Green, Blue) – The purpose of RGB color value is for sensing, representation and display of images. The RGB color model is additive in sense of light intensity.

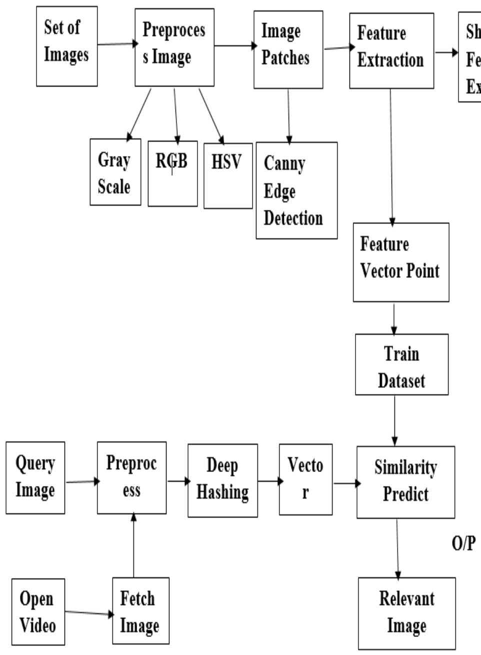

Fig.2: System Architecture

HSV (Hue, Saturation, Value) - This is used to reduce the certain amount of color from object and reduces the light intensity.

Image Patches – In this we use Canny Edge Detection Algorithm to detect the edge of image. Edge Detection is used for segmentation.

Feature Extraction – In this reduction of attributes is done. Features calculated at each pixel like color, location.

Feature Vector Point – In this we display the vector points of each pixel which is detected in feature extraction. Display Feature Vector and Display Descriptor this two vector points is displayed.

Train Dataset – In this we store the values of whole datasets which trained by Image Preprocessor like gray scale, RGB, HSV, Feature Patches, Feature Extraction. The MIRFLICKR Dataset is store for training purpose.

Deep Hashing – This is our Proposed approach by using Deep Hashing [DH] we learn how to compute compact binary hash code of image. Single and multi-layer projection is computed by Deep Hashing.

Open Video – In this we can run the AVI or any other File of Video. We also detect the feature of video image.

Fetch Image – In this we can fetch the video image after stopping the video and after that we preprocess that image.

Similarity Predict – In this we match the hash value of query image and stored image in database.

Relevant Image – After matching hash value which is having maximum accuracy that image is selected as an output.

IV. RESULTS AND IMPLEMENTATION

Below Figure shows you the Process of Image Processing using Deep Hashing.

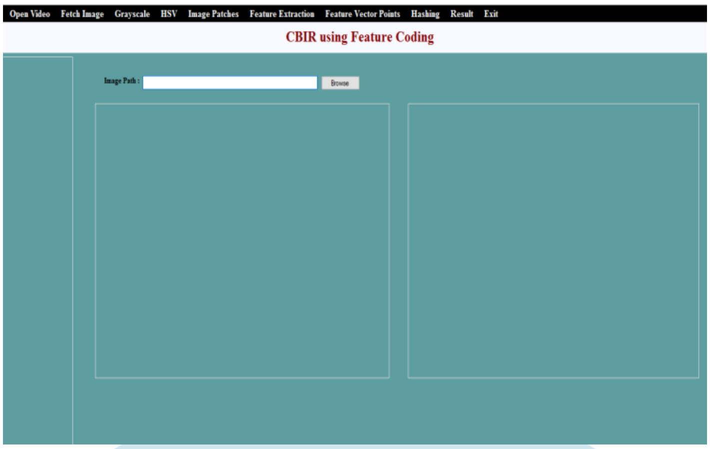

Fig.3: The Mainframe

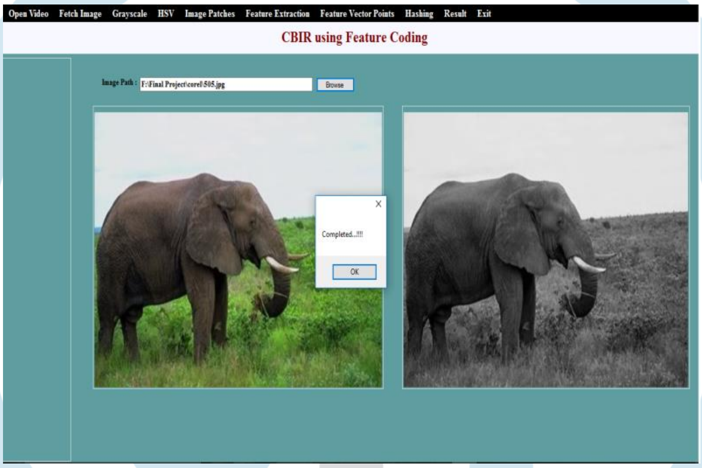

Fig.4: Gray Scale

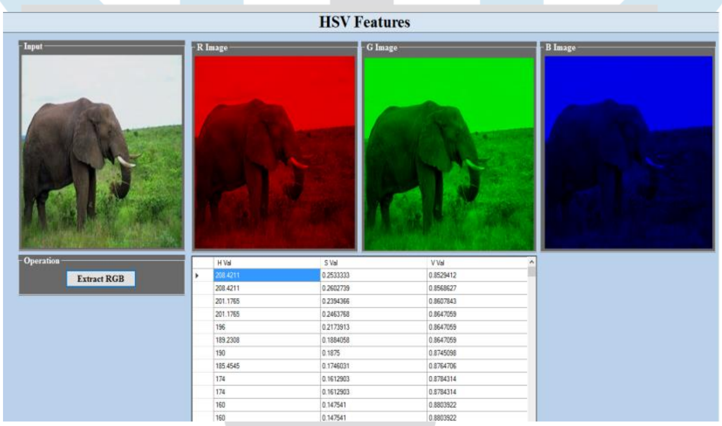

Fig.5: RGB and HSV

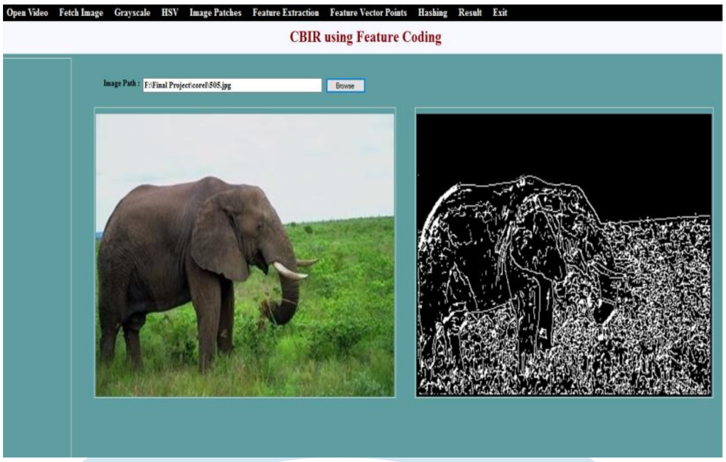

Fig.6: Image Patches

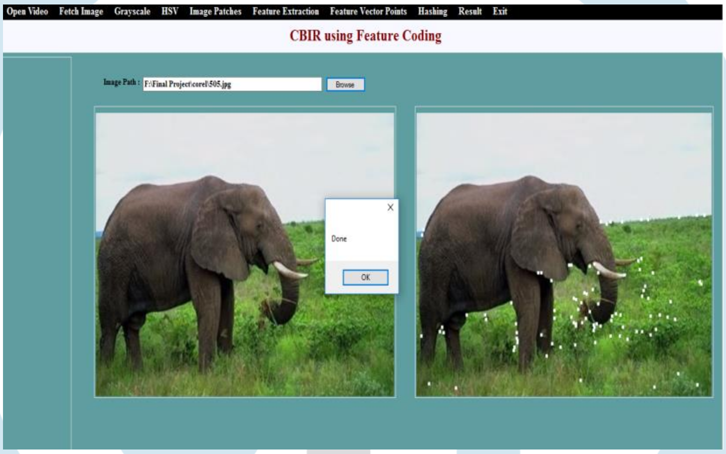

Fig.7: Feature Extraction

Fig.8: Display Feature Vector Point

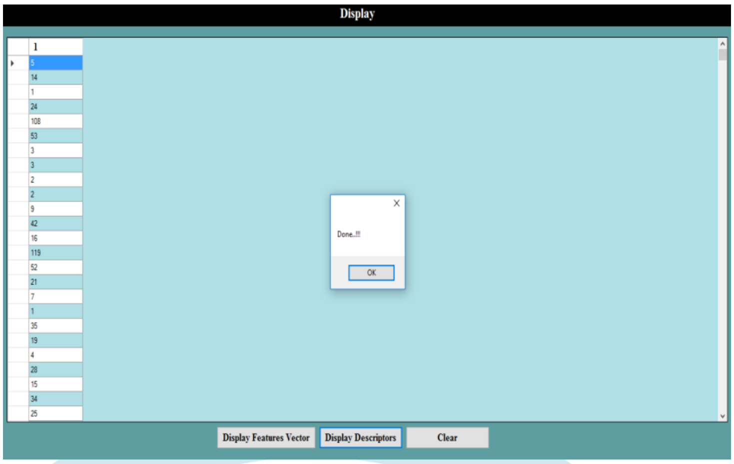

Fig.9: Display Descriptors

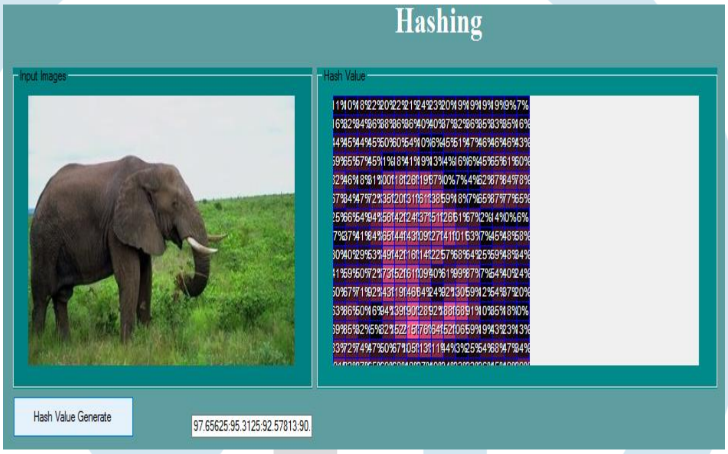

Fig.10: Hashing

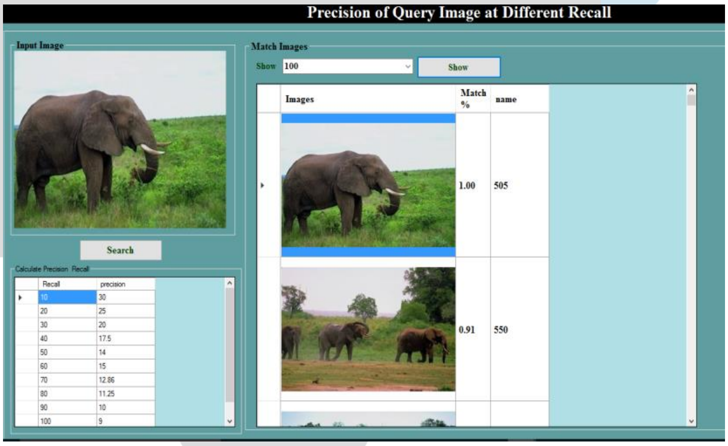

Fig.11: Precision and Recall

Contribution Part

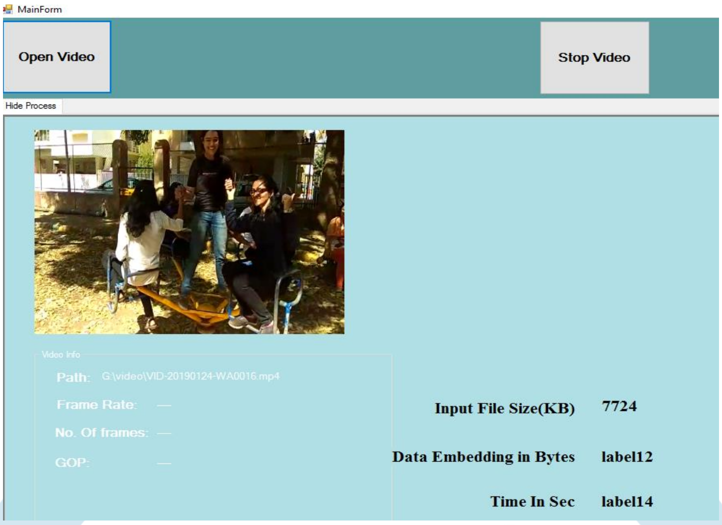

Fig.12: Open & Stop Video

Fig.13: Fetch Image from Video

V. CONCLUSION

In this paper, we propose a deep hashing for wide range of datasets. In normal hashing only single layer of projection is detected and by using deep hashing neural network we detect the multiple layer of projection. Then Deep Hashing [DH] is extended by SDH and MSDH.

ACKNOWLEDGMENT

I have a tremendous pleasure in presenting the project "Scalable Image Retrieval by Multilabel Deep Hashing" under the guidance of Prof. C. R. Barde and PG coordinator Prof. A. S. Vaidya. I am really obligated and appreciative to Head of the Department Dr. D. V. Patil for their significant direction and consolation. I might likewise want to thank the Gokhale Education Society's R. H. Sapat College of Engineering, Management Studies Research, and Nashik-5 for giving the required offices, Web get to and vital books. At last I must express my sincere heartfelt gratitude to all the Teaching Non-teaching Staff members of Computer Department of GESRHSCOE who helped me for their important time, support, remarks, and thoughts.

REFERENCES

[1]. J. Masci, M. M. Bronstein, A. Bronstein, and J. Schmidhuber. Multimodal similarity-preserving hashing. *TPAMI*, 36(4):824–830, 2014.

[2]. Z. Shi, X. Sun, and F. Wu, “Feature-based image set compression,” in Multimedia and Expo (ICME), 2013 IEEE International Conference on, 2013, pp. 1–6.H. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1985, ch. 4.

[3]. F. Shen, C. Shen, W. Liu, and H. Tao Shen. Supervised discrete hashing. In CVPR, pages 37–45, 2015.E. H. Miller, “A note on reflector arrays (Periodical style—Accepted for publication),” *IEEE Trans. Antennas Propagat.*, to be published.

[4]. J. Wang, Y. Song, T. Leung, C. Rosenberg, J. Wang, J. Philbin, B. Chen, and Y. Wu. Learning fine-grained image similarity with deep ranking. In CVPR, pages 1386–1393, 2014.

[5]. F. Zhao, Y. Huang, L. Wang, and T. Tan. Deep semantic ranking based hashing for multi-label image retrieval. In CVPR, pages 1556–1564, 2015.

[6]. J. Wang, W. Liu, A. X. Sun, and Y.-G. Jiang. Learning hash codes with listwise supervision. In CVPR, pages 3032–3039, 2013.

[7]. M. Ranzato, F. J. Huang, Y.-L. Boureau, and Y. LeCun. Unsupervised learning of invariant feature hierarchies with applications to object recognition. In CVPR, pages 1–8, 2007.

[8]. J. Xiao, J. Hays, K. Ehinger, A. Oliva, A. Torralba, et al. Sun database: Large-scale scene recognition from abbey to zoo. In CVPR, pages 3485– 3492, 2010.

[9]. A. Sablayrolles, M. Douze, H. J´egou, and N. Usunier. How should we evaluate supervised hashing? arXiv preprint arXiv:1609.06753, 2016.

[10]. K. Lin, H.-F. Yang, J.-H. Hsiao, and C.-S. Chen. Deep learning of binary hash codes for fast image retrieval. In CVPRW, pages 27–35, 2015.

[11]. J. Lu, V. E. Liong, X. Zhou, and J. Zhou. Learning compact binary face descriptor for face recognition. PAMI, 37(10):2041–2056, 2015.

[12]. X. Wang, T. Zhang, G.-J. Qi, J. Tang, and J. Wang. Supervised quantization for similarity search. In CVPR, pages 2018–2026, 2016.