IMAGE RETRIEVAL BASED ON IMAGE SIZE IN SUPERVISED LEARNING

MRS.S.DEEPIKA
ASSISTANT PROFESSOR, DEPARTMENT OF B.COM BUSINESS ANALYTICS
PSGR KRISHNAMMAL COLLEGE FOR WOMEN
deepika@psgrkcw.ac.in
N.SUGANYA
UG SCHOLAR, BCOM BUSINESS ANALYTICS
PSGR KRISHNAMMAL COLLEGE FOR WOMEN
suganyasugu2003@gmail.com

ABSTRACT:

This paper mainly aims at the image size retrieval system to display the images based on the image size. The image size may varies different size to be retrieved in the images. We have to find out and visualized the image size by sea born graph. Supervised learning of semantic searching is a method that allows users to draw generic and return similar drawn images. To analysis the top level of the images based on the image size. And to retrieve the images based on the image size. To analyse the image size based on the sea born graph and to find out the vector.

KEY WORDS: retrieval, image size, sea born graph, vector.

I. INTRODUCTION

Analysis is made with the python the dataset has been imported to jupyter notebook and thepython is a object oriented and a high level programming language. Supervised learning of semantic searching is a method that allows users to draw generic search queries and return similar drawn images, giving more user control over their search content. We have created various vectors with the attributes in the dataset that we created and extracted the results according to the image size. To retrieve the image based on image size. The analysis made through the sea born graph and to find the vector. The image size makes full use of the semantics in query and the top ranked images of the initial results. The Support Vector Machines (SVM), introduced by [1], is a supervised learning method for solving classification problems. SVM was applied for learning ranking functions in the context of information retrieval. It has also been employed specifically for CBIR the margin can be interpreted as a measure of separation between two classes and can be interpreted as a measure of the quality of the classification [3].

II. RELATED WORKS

Content-Based Image Retrieval (CBIR) emerged as a promising substitute to surpass the challenges met by text based image retrieval solutions. The online phase allows the user to start the retrieval task by providing his query image. Finally, typical CBIR system returns a set of images visually relevant to the user query. This assumption does not hold because of the semantic gap between the higher level meaning and the low-level visual features [10]. Researcher contributions to bridge the semantic gap can be categorised into different manner based on the adopted Technique can perceived as those focusing on scenery image retrieval [9].

Statistical modelling based clustering considers each cluster/category as a restrictively distributed pattern. Thus, the overall dataset is modelled as distribution mixture. The Expectation Maximization algorithm is usually used to estimate the parameters of the mixture components/distributions corresponding to the cluster properties [7]. More specifically, given the query image, clusters are learned in an unsupervised manner in order to enhance the retrieval accuracy. Objective function optimization is another traditional unsupervised learning technique. For instance, the popular K-means algorithm minimizes the sum of the intra cluster distances [6].

Bayesian classification in their proposed image retrieval approach. Their system aimed to capture high-level concepts of natural scenes using low-level features. Images were then automatically classified into outdoor or indoor images [11]. Three algorithms have been evaluated, representing very different learning strategies: CBIR-SVM, which is based on Support Vector Machines [4]. CBIR-GP, which is based on Genetic Programming, which is based on Association Rules. We have performed a systematic set of experiments using two image databases [2].

The Support Vector Machines (SVM), introduced by [1], is a supervised learning method for solving classification problems. SVM was applied for learning ranking functions in the context of information retrieval. It has also been employed specifically for CBIR the margin can be interpreted as a measure of separation between two classes and can be interpreted as a measure of the quality of the classification [3].

Genetic Programming (GP) is an inductive learning method introduced by Koza as an extension to Genetic Algorithms (GAs) [5]. It is a problem-solving system designed following the principles of inheritance and evolution, inspired by the idea of Natural Selection. The space of all possible solutions to the problem is investigated using a set of optimization techniques that imitate the theory of evolution .Whose non-leaf nodes are numerical operators and the leaf node set is composed of the similarity values obtained from different descriptors [8].

III. METHODOLOGY

SEABORN GRAPH:

Sea born is a library for making statistical graphics in Python. It builds on top of matplotlib and integrates closely with pandas data structures. Sea born helps you explore and understand your data. Its plotting functions operate on data frames and arrays containing whole datasets and internally perform the necessary semantic mapping and statistical aggregation to produce informative plots. Its dataset-oriented, declarative API lets you focus on what the different elements of your plot.

VECTOR:

To create a new vector in jupyter consists of attribute and retrieve the data according to the image size. A vector is a single dimensional array. Vector are ordinary numbers and attributes.

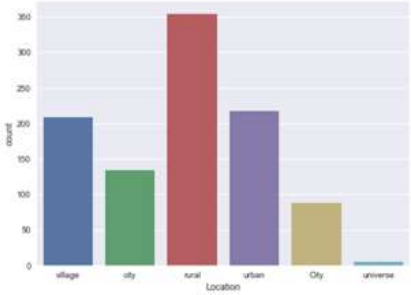
IV. RESULT

FIGURE: 3.1

```
In [14]: import seaborn as sns
         print(sns.countplot(data['Location']))

         plt.show()

Axes(0.125,0.125;0.775x0.755)
```



The graph represents the location attribute is taken here for visualization location is like where we search like Rural, urban, village, city. So here more people from rural location searched more than other city, village location people. It shows more than 0 to 350 in a rural location.

FIGURE:3.2

```
In [47]: str4 = data[['Image feature','Image Quality','Pixel Ratio', 'Imgae Type','Location']]

In [48]: str4
```

Out[48]:

	Image feature	Image Quality	Pixel Ratio	Imgae Type	Location
0	biriyani	HD	04:08	bmp	village
1	biriyani	VHD	16:32	jpeg	city
2	biriyani	UHD	04:08	jpg	rural
3	biriyani	4K	16:32	ico	urban
4	biriyani	Normal	04:08	mpeg	village
5	biriyani	HD	16:32	bmp	city
6	biriyani	VHD	04:08	jpeg	rural
7	pen	UHD	16:32	jpg	urban
8	pen	4K	04:08	ico	rural
9	pen	Normal	16:32	mpeg	rural
10	pen	HD	04:08	bmp	village
11	pen	VHD	16:32	jpeg	city
12	pen	UHD	04:08	jpg	rural
13	ball	4K	16:32	ico	urban
14	ball	Normal	04:08	mpeg	village
15	ball	HD	16:32	bmp	city
16	ball	VHD	04:08	jpeg	rural
17	ball	UHD	16:32	jpg	urban

In this analysis, We need to use the image retrieval dataset and it consists of 'Image feature', 'Image quality', 'Pixel ratio', 'Image type', 'location'. We retrieved the data based on image size which is in the dataset by comparing with other features. First, We created vector consist of some attributes like category, sub category, image type, image feature, pixel ratio, image quality , background noise percentage all the attributes.

FIGURE: 3.3

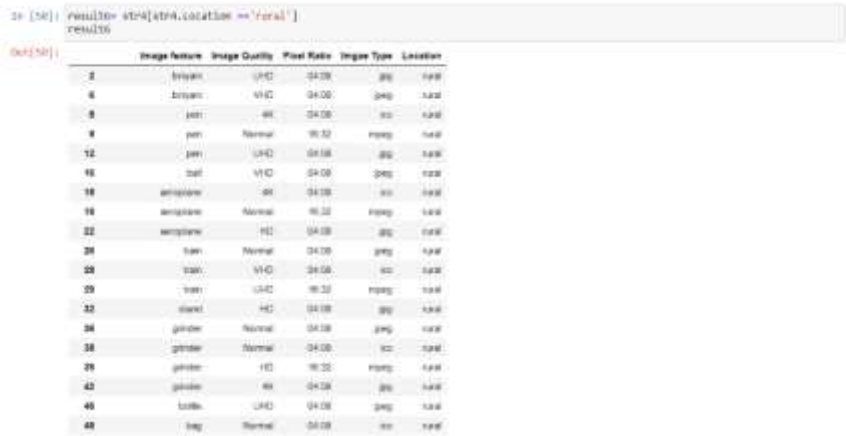
```
In [49]: result5= str4[str4.Location =='village']
         result5
```

Out[49]:

	Image feature	Image Quality	Pixel Ratio	Imgae Type	Location
0	biriyani	HD	04:08	bmp	village
4	biriyani	Normal	04:08	mpeg	village
10	pen	HD	04:08	bmp	village
14	ball	Normal	04:08	mpeg	village
20	aeroplane	4K	04:08	bmp	village
24	aeroplane	UHD	04:08	mpeg	village
30	train	4K	04:08	bmp	village
34	stand	UHD	04:08	mpeg	village
40	grinder	VHD	04:08	bmp	village
44	bottle	HD	04:08	mpeg	village
50	bag	VHD	04:08	bmp	village
54	bag	4K	04:08	mpeg	village
60	shawarma	Normal	04:08	bmp	village
64	shawarma	4K	04:08	mpeg	village
70	chips	Normal	04:08	bmp	village
74	chips	VHD	04:08	mpeg	village
80	samsung	UHD	04:08	bmp	village
84	micro oven	VHD	04:08	mpeg	village
89	printer	HD	04:08	bmp	village
93	stick	Normal	04:08	mpeg	village

We have create the vector called result1 that for comparing the attributes Location 'village' to display the attributes has been retrieved. It shows the image quality, pixel ratio and image type to analysis the image size. There are many image features to find out in the location.

FIGURE: 3.4



We have created the vector called result2 for comparing the attributes location 'rural' to display the attributes has been retrieved. It shows the image quality, pixel ratio and image type to analysis the image size. The images are mostly searched in a rural location.

V. CONCLUSION

In this paper shows the image retrieval is used to find out the image size. So we find out the result by visualized that more people has searched through the rural location by comparing village location. The village location is very low to retrieval the images.

FURTHER WORK:

It is suggested that the method of solution can be extended to different location and it represent the number of clusters that have occurred in attributes based on image quality and image type should be improve, so it is suggested in logistic regression.

REFRENCES

[1] Boser .B, Guyon .I, and Vapnik .V A training algorithm for optimal margin classifiers. In COLT, pages 144–152. Springer, 1992.
[2] Fan.W, Gordon.M.D, and Pathak.P. Genetic programming-based discovery of ranking functions for effective web search. J. Manage. Inf. Syst., 21(4):37–56, 2005.
[3] Han.J, McKennaS.J, and Wang.R. Learning query-dependent distance metrics for interactive image retrieval. In M. Fritz, B. Schiele, and J. H. Piater, editors, ICVS, volume 5815 of Lecture Notes in Computer Science, pages 374–383. Springer, 2009.
[4]Joachims.T Training linear SVMs in linear time. In SIGKDD, pages 217–226, 2006.
[5]Koza. J Genetic Programming: On the programming of computers by natural selection. MIT Press, 1992.
[6]Mac Queen, J. B.: "Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability", Berkeley, University of California Press, (1), 281-297, 1967.
[7] MCLACHLAN, G. AND PEEL, D. 2000. Finite Mixture Models. Wiley-interscience.
[8] R. da S. Torres, A. X. Falcao, M. A. Goncalves, J. P. Papa, B. Zhang, W. Fan, and E. A. Fox. A genetic programming framework for content-based image retrieval. Pattern Recogniti] on, 42(2):283–292, 2009.
[9]Shi.R, Feng.H, T.-S. Chua, C.-H. Lee, An adaptive image content representation and segmentation approach to automatic image annotation.
[10] Smeulders A.W, Worring. M, Santini, S., Gupta, A. and Jain.R, Content-based image retrieval at the end of the early years. IEEE Trans. Pattern Anal. Mach. Intell. 22 (12) (2000), 1349-1380.
[11] Vailaya .A, M.A.T. Figueiredo, Jain.A.K, Zhang.H.J, Image classification for content-based indexing, IEEE Trans. Image Process. 10 (1) (2001) 1171