https://dx.doi.org/10.24001/ijaems.3.4.10 *ISSN: 2454-1311*

The MIR Flickr Retrieval Evaluation Proposal Based on User Tags and Textual Passwords

Pavan Gujjar Panduranga Rao[1], Dr. P. Srinivasa Rao[2] , Dr. G. Lavanya Devi[3]

[1]Research Scholar, Department of Computer Science and System Engineering, Andhra University College of Engineering(A),Vishkhapatanam, Andhra University, India
[2,3]Professor, Department of Computer Science and System Engineering, Andhra University College of Engineering(A), Andhra University, Vishkhapatanam, Andhra Pradesh, India

Abstract— *I*n most well known image retrieval test sets, the imagery typically cannot be freely distributed or is not representative of a large community of users. In this paper we present a collection for the MIR community comprising 69,000 images from the Flickr website which are redistributable for research purposes and represent a real community of users both in the image content and image tags. We have extracted the tags and EXIF image meta data, and also make all of these publicly available. In addition we discuss several challenges for benchmarking retrieval and classification methods and applications.

Keywords— Content-based image retrieval (CBIR), relevance feedback, image collections, benchmarking, Graphical Password Authentication.

Categories and Subject Descriptors

[Information Storage and Retrieval]: Digital Libraries – Collection Dissemination, Standards.

[Information Storage and Retrieval]: Information Search and Retrieval – Query Formulation.

General Terms

Experimentation, Human Factors, Measurement, Performance,
Standardization.

I. INTRODUCTION

Arguably, the most frequently used test set in content-based image retrieval is the Asia Corel Stock Photography collection. The total collection consists of more than 800 Photo CDs, each containing 100 broadly similar images of a certain category. In most cases, research groups have made their own selection from the available categories, usually amounting to a varying subset consisting of 6,000 to 20,000 images.

II. FLICKR TAGS

The experimentation is conducted on MIR Flickr dataset. The data set contain 69,000 unique tags and a cardinality of a tag for image.

The appealing feature in the Flickr dataset is that, it facilitates the users to retrieve their images of interest by associating the tags to the images. The tags can be further divided in sub categories. The useful tags are those that clearly portray the images, using a direct relation based on the visual content of the image. Each of the tags is processed by eliminating the capital letters, white spaces and other special characters. The usual number of tags per image is 8.94. In the set there exist 1386 tags which take place in at least 20 images, the most highlighted images in the dataset available with tags along with the frequency of occurrence is presented in the following table-1.

Table.1: Frequency of Content-Based Tags

Image Tag	Frequency
Sky	845
Water	641
Portrait	623
Night	621
Nature	596
Sunset	585
Clouds	558
Flower / Flowers	510/351
Beach	407
Landscape	385
Street	383
Dog	372
Architecture	354
Graffiti / Street art	335/184
Tree / Trees	331/245
People	330
City / Urban	308/247
Sea	301
Sun	290
Girl	262
Snow	256
Food	225
Bird	218
Sign	214
Car	212
Lake	199
Building	188
River	175
Baby	167
Animal	164

https://dx.doi.org/10.24001/ijaems.3.4.10 *ISSN: 2454-1311*

The Sample Dataset considered from Flicker, is presented, in figure-1

Fig.1: MIR Flickr Dataset

III. MIR FLICKR DATASET

The MIR Flickr dataset contains 25,000 images retrieved from the Flickr multimedia repository taking into account a high "interesting-ness" factor. This dataset supplies all original tags user annotations provided by the Flickr users and also annotations for all the images, which were obtained by majority voting. Annotations for 24 concepts are provided in this dataset. This is a significant dataset that encompasses various types of photographs with a varied set of concepts between them. There are about 69,000 unique tags in the dataset and a cardinality of 9 tags per image. The distribution between the various types is seen using a sample from the MIR Flickr dataset.

The tag type taxonomy used in this chapter is based on the following categories:

1) EXIF: Metadata embedded in the image related to the device used to capture the image. An example of this is the information regarding the maker and model of the camera;
2) Geo-tag/EXIF-Location: Metadata available in certain devices that references a geographical location, usually where the photo was taken;
3) Low-level: Pertains to the visual content of an image;
4) Mid-level: The most common features containing any of the generic keyword;
5) High-level: These correspond to the annotations, keywords with ground-truth certainty;
6) Author : These are directly related to the owner / user subjective assessment of the image;
7) Collaborative: These are related to the interpretation of a group of people, hence collaborative;
8) Ambiguous: In this type of features there may exist multiple interpretations of the keyword based on the viewer and context.

   There can also exist an overlapping between other types of tags, as such. Ambiguous situations arise mostly among the datasets.

a) Semantic: A single keyword considered may have different meanings in different sense called as semantic tags.
b) Geographical: The tag, which corresponds to the location of a particular place.
c) Temporal: The tag, which correspond to multiple instances of the events.
d) Language: The tag considered may be having different meanings in different languages.
e) Generalization: The representation of a tag, which overcome the ambiguity is considered to be generalization. Some of the images from MIR Flickr dataset, together with their associated tag frequencies is presented in the following Figure-2

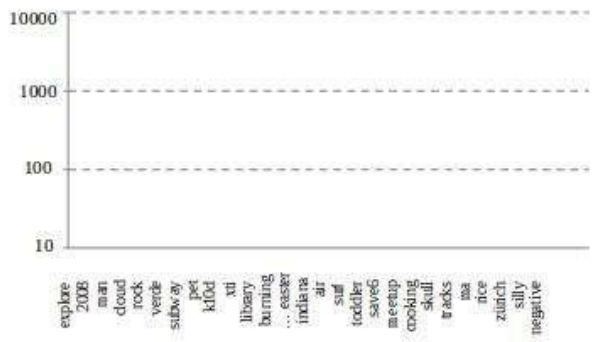


Fig.2: Tag Frequency Distribution in MIR Flickr Dataset.

IV. EXPERIMENTATION PERFORMANCE EVALUATION

In order to evaluate the performance of the proposed method precision and recall are considered. The formulas for computing are given below. The precision and recall values for the considered images is presented in table-3**.**

V. PRECISION

It is the ratio of the number of appropriate images retrieved to the total number of unrelated and related images retrieved. It is usually expressed as a percentage.

Precision = (A / (A + C))* 100;

A: Number of related images retrieved.

C: Number of irrelevant images retrieved.

A + C: Total number of irrelevant + relevant images retrieved.

VI. RECALL

It is the ratio of the number of relevant images retrieved to the total number of relevant images in the database. It is usually expressed as a percentage.

Recall = (A / (A + B)) * 100

https://dx.doi.org/10.24001/ijaems.3.4.10 ISSN: 2454-1311

A: Number of relevant images retrieved

B: Number of relevant images not retrieved

A + B: The total number of relevant images

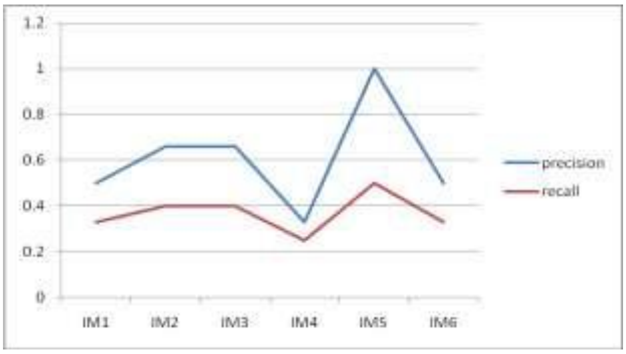
The results obtained are presented in table-2

Table.2: The Results Obtained

Query image	Model based approach	
	Precision	Recall
	26.8	18.5
	21.1	8.0
	23.3	4.2
	56.6	40.8
	44.0	4.4

FIGURES AND TABLES

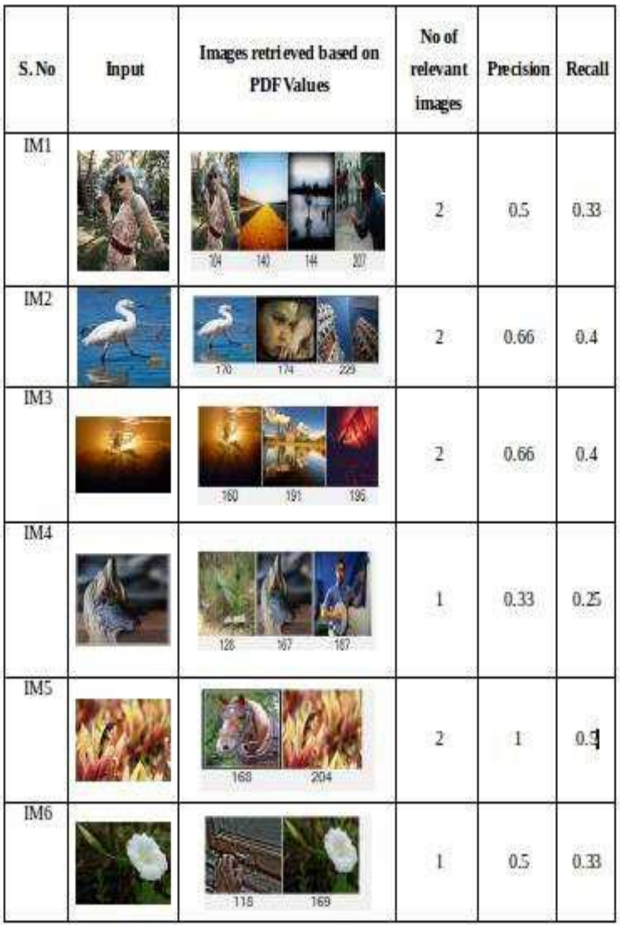
From the below Graph-1, it can be clearly seen that the precision rate increases which clearly helps in authenticating the individual.



Graph -1: Graph Representing Precision and Recall

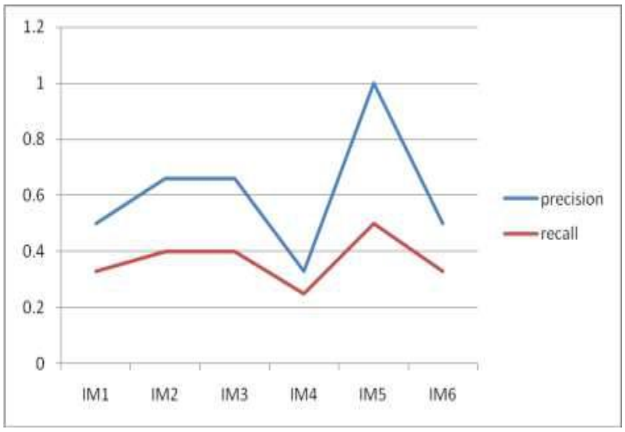
The authentication based on the developed model is also tested against the Flickr dataset based on the PDF values is presented in table-3.

Table.3: Images Retrieved based on PDF Values



S.No	Input	Images retrieved based on PDF Values	No of relevant images	Precision	Recall
IM1		104 140 144 207	2	0.5	0.33
IM2		170 174 229	2	0.66	0.4
IM3		160 191 196	2	0.66	0.4
IM4		128 167 187	1	0.33	0.25
IM5		168 204	2	1	0.5
IM6		118 169	1	0.5	0.33

The relevancy of the retrievals based on the PDF is evaluated using the metrics; precision and recall and the results obtained are presented in the Graph-2



Graph-2: Graph Representing Precision and Recall

https://dx.doi.org/10.24001/ijaems.3.4.10 ISSN: 2454-1311

VII. EXPERIMENTATION ON MIR FLICKR DATASET

Database Considered:

Fig.3: Dataset Considered from MIR Flickr Dataset

In order to present the methodology, experimentation we have considered two datasets namely Flickr database consisting of 25,000 images with tags like nature, cigarette, flowers and watches etc. Each image is associated with a tag description; among these images 450 have unique tags. The experimentation is performed by considering 2000 images, query image is considered with the size 100 x 100

Figure-4 : Database Considered from Flicker

VIII. TAG BASED AUTHENTICATION

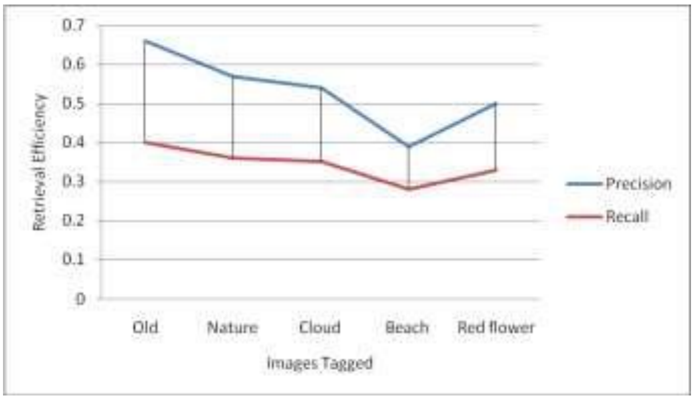
Generally, the process of feature extraction is carried out in the background, hence for efficient and effective retrievals, these features are to be associated with semantic interpretations. The semantic interpretations help to extract the data using the semantic attribute and also minimize the semantic gap. These semantic features are straight-forwardly understood by the users in contrast to the low level image attributes which include dissimilarity, proportion, homogeneity and uniformity.

The retrieval of images is carried out in two methods. In the 1st model the dataset with meta-tags are considered as input, for the experimentation of this model we have considered Flickr dataset is considered for model-1, and both Flickr and MIR Flickr are considered for model-2. For each tag, the synonyms are extracted and a code book of tags is generated.

The retrieval is based on the relevancy between the synonyms in the code book and the corresponding image meta-tags in table-4.

Table.4: Images Retrieved based on Code Book

Tags	Code Book	Images Retrieved	No of Images Retrieved	Relevant Images	Precision	Recall
Old	aged, elderly, older, mature, getting on, not getting any, younger		3	2	0.66	0.4
Nature	natural, world, natural, history, scenery, life, environment		21	12	0.57	0.36
Cloud	blur, obscure, make unclear, darken, shade, dim, confuse		11	6	0.54	0.35
Beach	seashore, seaside, coast, shore, coastline, shoreline, sand		28	11	0.39	0.28
Red Flower	crimson, crimson, scarlet, ruby, burgundy, cherry, blossom		4	2	0.5	0.33

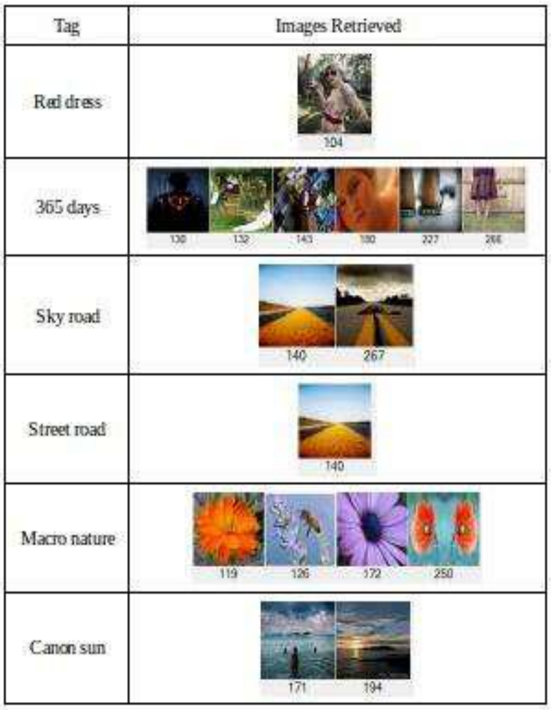


Graph-3: Graph Showing Precision and Recall

https://dx.doi.org/10.24001/ijaems.3.4.10 ISSN: 2454-1311

In the 2nd model, to reduce the search space, the dimensionality of the data is reduced by identifying the most relevant images, based on color, texture is considered . The most relevant images are considered by using the concept of phrases, where each image is tagged and these tags are stored in the meta-data. The methodology is presented using two datasets, namely Flickr and MIR Flickr The retrieval is based on this tag as input and the results obtained are present in graph-3.

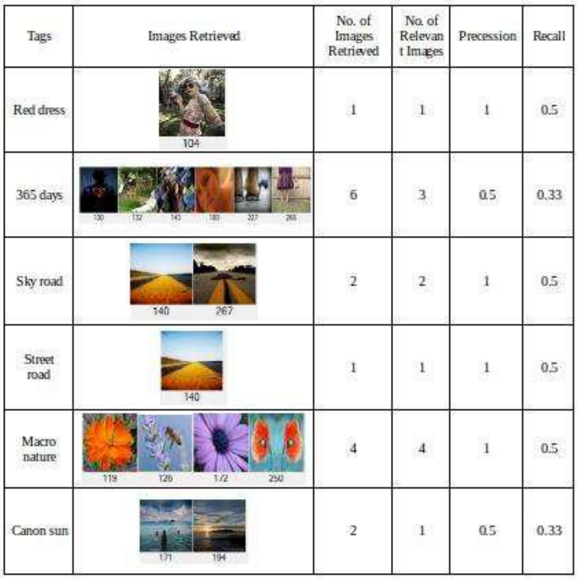
Table.5 : Tags and Image Retrieved



Tag	Images Retrieved
Red dress	104
365 days	130 132 143 180 227 266
Sky road	140 267
Street road	140
Macro nature	119 126 172 250
Canon sun	171 194

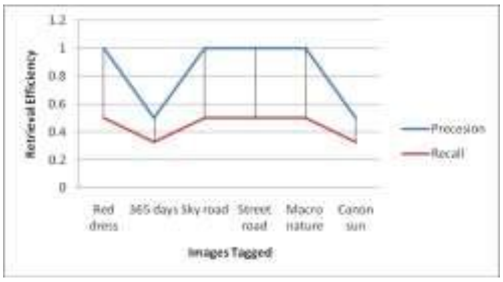
Performance evaluation based on image retrievals in table-6 obtained as below,

Table.6: Performance Evaluation



Tags	Images Retrieved	No. of Images Retrieved	No. of Relevant Images	Precession	Recall
Red dress	104	1	1	1	0.5
365 days	130 132 143 180 227 266	6	3	0.5	0.33
Sky road	140 267	2	2	1	0.5
Street road	140	1	1	1	0.5
Macro nature	119 126 172 250	4	4	1	0.5
Canon sun	171 194	2	1	0.5	0.33

The images retrieved from the flicker dataset based on the tags are presented in table-5.



Graph-4: Based on PDF values Graph of Precision and Recall

IX. CONCLUSION

In this research article two methodologies are presented for authenticating a user, namely tag based, synonym based retrieval, on datasets namely Flickr and MIR Flickr and the results obtained are presented in tables and graphs. From the above tables it could be noted that the authentication based on synonym based tagging helps towards better retrieval accuracy when compares to the retrievals based on tagging method. The performance evaluation is calculated using precision and recall and the results obtained are tabulated. From the above table it could be noted that the present model authenticates the users relevantly. Applications of CBIR in Architecture and engineering Design, Crime Prevention, Geographical information, Photograph archives, Textile Industry and Intellectual Property.

ACKNOWLEDGEMENTS

Pavan Gujjar Panduranga Rao is currently a Research Scholar from Andhra University and working as professor in the Department of Computer Engineering, Image Engineering Laboratory at St. John College of Engineering and Management, Palghar affiliated to Mumbai University, Mumbai. He received his B.E. and M.E degrees, both in Computer Science and Engineering, from Mysore University and Bangalore University, India. His research interests can be summarized as developing effective and efficient MIR Flickr data analysis techniques for Graphical Password Authentication and the related applications.

REFERENCES

[1] Dirik, A E, Memon, N and Birget, J C, "Modelling user choice in the Passpoints graphical password scheme", in Proceedings of the 3rd symposium on Usable privacy and security (SOUPS), ACM Press, 2007, Pg. 20-28. *(references)*

[2] Djamila, R. and Perrig, A. (2000) “Deja Vu: A user study using images for authentication”, Proceedings of the 9th USENIX Security Symposium. Denver, Colorado, USA August 14-17, Pg. 45-58

[3] GVS Rajkumar, K.Srinivasa Rao, And P.Srinivasa Rao,”Image Segmentation and Retrievals based on Finite Doubly Truncated Bivariate Gaussian Mixture Model and K-Means”, International Journal of Computer Applications (IJCA) ,2011, Vol. 25, No. 4, Pg. 5-13.

[4] H C Sateesh Kumar, K B Raja, Venugopal K R and L M Patnaik, Automatic Image Segmentation using Wavelets", IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.2, February 2009.

[5] Pavan Gujjar Panduranga Rao, P.Srinivasa Rao, G.Lavanya Devi. “Hybrid Model Based On User Tags And Textual Passwords And Pearsonian Type III Mixture,” International Journal of Advanced Engineering Research and Science (IJAERS)

[6] P.CLOUGH,H.MUELLER, AND M. SANDERSON. 2004. The CLEF cross-language image retrieval track (imageCLEF) 2004. In Fifth Workshop of the Cross-Language Evaluation Forum (CLEF 2004), LNCS, New York, NY, USA, ACM Press.

[7] P. SINHA AND R. JAIN (2008). Classification and annotation of digital photos using optical context data. ACM International Conference on Content-Based Image and Video Retrieval (CIVR '08), pages 309-318, Niagara Falls, Canada.

[8] M. S. LEW, N. SEBE, C. DJERABA, AND R. JAIN. 2006. Content-based multimedia information retrieval: State of the art and challenges. ACM Trans. Multimedia Computer. Communication. Applications., 2(1) :1–19.

[9] C. FLUHR, P.-A. MOELLIC, AND P. HEDE. 2006. Usage oriented multimedia information retrieval technological evaluation. In MIR ’06: Proceedings of the 8th ACM international workshop on Multimedia information retrieval, pages 301–306, New York, NY,USA. ACM Press.

[10] M. J. Huiskes and M. S. Lew. The MIR Flickr Retrieval Evaluation. In MIR ’08: Proceedings of the 2008 ACM International Conference on Multimedia Information Retrieval , New York, NY, USA, 2008. ACM.

[11] F. Jurie and B. Triggs. Creating efficient codebooks for visual recognition. In ICCV ’05: Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV’05) Volume 1 , pages 604–610,Washington, DC, USA, 2005. IEEE Computer Society.

[12] T. K. Landauer and M. L. Littman. Fully automatic cross-language document retrieval using latent semantic indexing. In Proceedings of the Sixth Annual Conference of the UW Centre for the New Oxford English Dictionary and Text Research , pages 31–38, UW Centre for the New OED and Text Research, Waterloo, Ontario, Canada, October 1990.

[13] S. Lazebnik, C. Schmid, and J. Ponce. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. Computer Vision and Pattern Recognition, IEEE Computer Society Conference on , 2:2169–2178, 2006.

[14] Torralba, A., Fergus, R., and Freeman, W. T. 2008. 80 Million Tiny Images: A Large Data Set for Nonparametric Object and Scene Recognition. IEEE Trans. Pattern Anal. Mach. Intell.30, 11 (Nov. 2008), 1958-1970. DOI= http://dx.doi.org/10.1109/TPAMI.2008.128

[15] Torralba, A. Fergus, R. Weiss, Y. 2008. Small codes and large image databases for recognition. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2008, 1-8, Anchorage, AK.

[16] Uijlings, J. R., Smeulders, A. W., and Scha, R. J. 2009. Real-time bag of words, approximately. In Proceedings of the ACM international Conference on Image and Video Retrieval (Santorini, Fira, Greece, July 08 - 10, 2009). CIVR'09.ACM, New York, NY, 1-8. DOI= http://doi.acm.org/10.1145/1646396.1646405

[17] Ulges, A., Schulze, C., Keysers, D., and Breuel, T. 2008. Identifying relevant frames in weakly labeled videos for training concept detectors. In Proceedings of the 2008 international Conference on Content-Based Image and Video Retrieval (Niagara Falls, Canada, July 07 - 09, 2008). CIVR '08. ACM, New York, NY, 9-16. DOI= http://doi.acm.org/10.1145/1386352.1386358

[18] Verbeek, J., Guillaumin, M., Mensink, T. and Schmid, C. 2010. Image Annotation with TagProp on the MIRFLICKR set. In Proceedings of the 2nd ACM international Conference on Multimedia information Retrieval (MIR '10).

[19] Yang, J. and Hauptmann, A. G. 2008. (Un)Reliability of video concept detection. In Proceedings of the 2008 international Conference on Content-Based Image and Video Retrieval (Niagara Falls, Canada, July 07 - 09, 2008). CIVR '08. ACM, New York, NY, 85-94. DOI= http://doi.acm.org/10.1145/1386352.1386367

[20] Zhang, J., Marsza ł ek, M., Lazebnik, S. and Schmid, C. 2007. Local Features and Kernels for Classification of Texture and Objecs. Categories: A Comprehensive Study . IJCV , 73(2): 213–238.

[21] S. Cohen and L. Guibas, "The earth mover's distance: Lower bounds and invariance under translation," Stanford University, DTIC Report, 1997.

[22] Hare, J.S and Lewis, P.H. 2010. Automatically Annotating the MIR Flickr Dataset. In Proceedings of the 2nd ACM international Conference on Multimedia information Retrieval (MIR '10).

[23] Hörster, E. and Lienhart, R. 2008. Deep networks for image retrieval on large-scale databases. In Proceeding of the 16th ACM international Conference on Multimedia(Vancouver, British Columbia, Canada, October 26 - 31, 2008). MM '08. ACM, New York, NY, 643-646. DOI= http://doi.acm.org/10.1145/1459359.1459449

[24] Huiskes, M.J. 2006. Image Searching and Browsing by active aspect-based relevance learning. In Proceedings of CIVR06, LNCS4071, 211-220. Springer

[25] A. Tung, R. Zhang, N. Koudas, and B. C. Ooi, "Similarity search: A matching based approach," in VLDB, 2006. UCI. (2013) Corel image features data set. [Online].Available: http://archive.ics.uci.edu/ml/datasets/Corel+Image+Features

[26] Srinivas Yerramalle and K Srinivasa Rao (2007), Unsupervised image classification using finite truncated Gaussian mixture model, Journal of Ultra Science for Physical Sciences, Vol.19, No.1, Pg.107-114.

[27] T. Yamazaki (1998), "Introduction of EM algorithm into color image segmentation," in Proceedings of ICIRS'98, Pg. 368–371.

[28] Yampolskiy, R. V. (2007) 'User Authentication via Behavior Based Passwords', Systems, Applications and Technology Conference, 2007. LISAT 2007. IEEE Long Island. 4-4 May 2007. Pg. 1-8.

[29] Yan, J., and et al. Password memorability and sec.: Empirical results. IEEE Sec. and Priv. 2, 5 (Sept. 2004), Pg. 25–31.

[30] Yan-yan Wang, Jia-yong Liu, Feng-xia Xiao, Jing Dan, "A More Efficient & Secure Dynamic Id-Based Remote User Authentication Scheme", Computer Communications, Vol 32, Issue 4, March 2009, Pg 583-585.

[31] Yardi, S., Krolikowski, P., Marshall, T., and Bruckman, A. (2008). "An HCI Approach to Computing in the Real World." Journal of Educational Resources in Computing. 8, 3 (Oct. 2008), Pg. 1-20.

[32] Prasad Reddy P.VG.D, Srinivas Rao K, Srinivas Yarramalle, "Unsupervised Image Segmentation using Finite Generalized Gaussian Distribution with EM and K-Means Algorithm", International Journal of Computer Science and Network Security(IJCSNS), 7(4) (2007).

[33] Yee, K.-P., "Aligning Security and Usability". IEEE Security and Privacy, 2004, Vol 2, Issue 5, Pg. 48-55.

[34] Zhao, Z., Dong, Z. & Wang, Y. (2006) 'Security Analysis of a Password-based Authentication Protocol Proposed to IEEE 1363. Theoretical Computer Science, 352 (1–3). Pg. 280-287.