A Novel Hybrid Approach to Visual Concept Detection Using Image Annotation

Sanjay M. Patil *sannit.99@gmail.com*
Research Scholar, Department of Computer Technology,
Yeshwantrao Chavan College of Engineering,
Nagpur, 441110, India

Kishor K. Bhoyar *kkbhoyar@yahoo.com*
Professor, Department of Information Technology,
Yeshwantrao Chavan College of Engineering,
Nagpur, 441110, India

Abstract

Millions of images are being uploaded on the internet without proper description (tags) about these images. Image retrieval based on image tagging approach is much faster than Content Based Image Retrieval (CBIR) approach but requires an entire image collection to be manually annotated with proper tags. This requires a lot of human efforts and time, and hence not feasible for huge image collections. An efficient method is necessary for automatically tagging such a vast collection of images. We propose a novel image tagging method, which automatically tags any image with its concept. Our unique approach to solve this problem involves manual tagging of small exemplar image set and low-level feature extraction of all the images, hence called a hybrid approach. This approach can be used to tag a large image dataset from manually tagged small image dataset. The experiments are performed on Wang's Corel Dataset. In the comparative study, it is found that, the proposed concept detection system based on this novel tagging approach has much less time complexity of classification step, and results in significant improvement in accuracy as compared to the other tagging approaches found in the literature. This approach may be used as faster alternative to the typical Content Based Image Retrieval (CBIR) approach for domain specific applications.

Keywords: Image Concept Detection, Low-level Features, Tags, Weighted Features.

1. INTRODUCTION

The objective of concept-based image retrieval is to index and retrieve images based on concepts. The visual content and its accompanying metadata of images are used to detect these concepts[1]. Humans can easily understand the content of images, but search engines have a limited ability to recognize the image or scene content. Image scene elements ("trees," "clouds"), objects ("elephant," "flower"), etc. are some of the examples of Concepts. To retrive a query image from non-annotated image collections, the information of concept can be used. The low-level features extracted from visual data are mapped to high-level features for indexing of images for concept detection. The high-level features are the one which are perceived by humans. A large number of both positive and negative examples are required for efficient concept detector training.

Robust training data can be built by manually labeling images, but this is a laborious task and time consuming. This leads to need of different training samples for each domain, resulting in high number of training samples to achieve high accuracy of the detection. Many approaches are proposed to overcome these difficulties by automatically generating concept training sets. Authors in [2] use active learning where repeated training cycles are involved in a dataset. However, instead of using approaches involving training set generation, Mandel et al. [3]

proposed the use of tags. Tags are words assigned to an image which may reveal the concept of an image. Usually, tags are single words. This work introduces an approach where color and texture features are extracted and applied for improving concept detection of the image. We select tagged prototype training images from a small data collection and apply our algorithm to determine the concept relevance of all images. We get a matching score which is converted into possibility matrix. Thresholding is used to automatically annotate the tag of prototype image to training images based on this possibility matrix. This information is used to detect the concept of the query image. The rest of the paper is organized as follows: Section 2 gives a brief overview of previous approaches to concept detection. Section 3 describes various image descriptors used in this paper as well as the summary of the dataset used. It also explains the classifier used. Section 4 presents experimental results. Section 5 gives result analysis. Section 6 concludes the paper.

2. LITERATURE REVIEW

In this section, a brief review of previous research work in the area of image concept detection is presented. We review the works that use tagged images to generate training sets and build concept detection model automatically. Authors in [4] provided an early discussion about generic concept detection approaches and state that concept detection is considered as a supervised pattern recognition problem mostly. The concept models are trained by a manually tagged set of training images. Several concepts may co-occur in an image when it is classified. Hence concept detection is regarded as a multi-class multi-label problem. The Fuzzy SVM classifier has been used in various works. In [5] a membership function is found based on the Euclidean distance of images of the training. This requires manually labeled images and relies on user feedback to balance the fuzzy membership of each image. The work in [6] implements a social assisted media tagging scheme which takes advantage of the large user generated images and its tags to train the classifiers. A similarity graph is built in between the labeled and unlabeled images of the collection in [7]. Further, features for training concept detectors use graph Laplacian eigenmaps. It offers multiple options for fusing different image features. [8] Presents a method for classification using edge-based features, namely edge direction histogram and edge intensity histogram which provide discriminative information useful for classification. Authors in [9] proposed a framework for individual concept inference and refinement by exploring the concept co-occurrence patterns in images with network community detection algorithms. A framework that extends Bag-of-Words with higher-order occurrences computed on mid-level features is proposed in [10]. Concept-based image retrieval based on automatic ground truth generation using tags is demonstrated in [11]. A set of words define a concept. The relevance of an image to a concept is determined by a score using its tags. The co-occurrences against a corpus of words are used to assign similarity values to pairs of words. The final score is computed by considering maximum or average values and used for automatic annotation of images. To measure the relevance of social tags with respect to the visual content, tag relevance fusion scheme is introduced in [12]. Recent developments in the area of object categorization and image region classification, from both theoretical and application perspectives, has been presented in [13]. In the implementation of a system for concept detection, the problems reported in the literature are:

1. Effective performance of concept detector requires variety of images and accuracy of tagging.
2. Tagging a large dataset is a laborious task and requires lot of time.
3. Complex classifiers are needed for classification of various concepts.

In this work we have used a small manually tagged image dataset to automatically tag a large dataset. By introducing manual tagging with feature matching, we eliminated the need of complex classifier. Our technique is based on simple classifier based on thresholding.

3. MATERIALS AND METHODS

Any concept detection system must extract suitable features and use some learning mechanism to detect the concept of the input image. This section outlines the dataset used in our experiment, the individual features and a novel approach used for classification.

3.1 Dataset Used

We have used Wang's Corel 1K Dataset which is subset of Corel stock photo database. This dataset has 10 classes of 100 images each and has size of either 384 × 256 or 256 × 384. Corel includes a wide range of images from natural scenes to artificial objects. The classes defined are beach, monument, bus, dinosaur, elephant, flower, horse, mountain, food and African. Figure 1 shows sample images of these classes.

3.2 Image Descriptors

Features are used to represent images for searching, indexing and browsing images in an image dataset. The image feature is defined as a function of one or more measurements, each of which specifies some quantifiable property of an image and is computed such that it quantifies some significant characteristics of the image. The direct way is the extraction of feature vectors using computer vision techniques to build visual concept detector model. Feature extraction is the process of generating features to be used in the selection and classification tasks. Proper use of feature space is required for efficient storage and computation.

CBIR techniques automatically extract low-level features (e.g. color, texture, shapes and layout of images) to measure the similarities among images by comparing the feature differences. Feature extraction from images and selection of appropriate features is the key to the success of any image mining task [1]. In our experiment we have used color and texture features as elaborated below.

3.2.1 Color Features

Color is a very important clue for searching images, as it is invariant to scaling, translation and rotation of an image. For example, a mountain scene can be characterized by blue sky on the top whereas a forest scene will contain substantial portions of green shades [8]. The color features used in this experiment are:

FIGURE 1: Sample Images from Each Class of Corel 1k Dataset.

3.2.1.1 Color Histogram

A histogram is a graphical representation of the tonal distribution in a digital image. The number of bits used to represent each pixel of an image decides the total number of grey levels in histogram. Color histogram [9] is the simplest and most common way of expressing the statistical distribution of colors and the essential tone of an image. The color histogram is invariant to translation and rotation of the imaging axis.

Human visual system characterizes a color image by its brightness and chromaticity. Brightness is a subjective measure of luminous intensity. Hue and Saturation define the chromaticity. Hue is a color element and represents a dominant color. Saturation is an expression of the degree to which white light dilutes a pure color. The HSV model is motivated by the human visual system as it better describes a color image than the RGB model [16].

3.2.1.2 Color Moments

Color moments are measures that characterize color distribution in an image. Equation 1, 2 and 3 shows the mean, variance and standard deviation moments for each channel of a color space where i is the index of each channel, I_{ij} is the value of the j^{th} pixel in channel i and N is a total number of image pixels.

$$E_i = \frac{1}{N}\sum_{i=0}^{N} I_{ij} \qquad \text{K (1)}$$

$$\sigma_i = \sqrt{\left(\frac{1}{N}\left(\sum_{j=1}^{N}(I_{ij} - E_i)^2\right)\right.} \qquad \text{K (2)}$$

$$s_i = \sqrt[3]{\left(\frac{1}{N}\sum_{j=1}^{N}(I_{ij} - E_i)^3\right)} \qquad \text{K (3)}$$

Color moments are invariant to scaling and rotation. They can be used as features in order to compare how similar two images are based on color. Most of the color information is contained in the low-order moments. We have used two lower order color moments, Mean and Standard Deviation for each channel in RGB space.

3.2.2 Texture Features

Texture is an important visual feature used in domain-specific applications. It can give us information about the content of an image efficiently. It is a repeated pattern of information or arrangement of the structure with regular intervals. It quantifies the properties such as smoothness, coarseness and regularity in an image. The texture feature used in our system is:

3.2.2.1 Wavelet Transform

The wavelet transform is one of the current popular feature extraction methods used in texture classification. The wavelet transform is able to de-correlate the data and provides orientation sensitive information which is vital in texture analysis. It uses wavelet decomposition to significantly reduce the computational complexity and enhance the classification rate. Table1 summarizes the feature set used in our experiments.

3.3 Classifier Design

In this section we are presenting a novel hybrid approach to concept detection. The proposed approach is implemented using two stages. First stage is automatic annotation of images and second stage is concept detection. Figure 2 shows the steps for automatic annotation. We have

Feature	Feature Description	Dimension
Color Moments (CM)	Low order moments (mean and standard deviation)	6
HSV Color Histogram (HSV)	Each of h, s and v channel is quantized to 8x2x2 bins respectively	32
Wavelet Transform (WT)	Mean square energy and standard deviation	40

TABLE 1: Feature Set used In Our Approach.

selected two image datasets, a small manually tagged image dataset (MIDS) which consists of tagged prototype images of each individual concept and a large untagged image dataset (UIDS). The low-level features, color moment, HSV histogram and wavelet moment of both datasets are extracted and stored. The features of these two datasets are matched by comparing each image in MIDS with each image in UIDS. The matching score is used for automatic image tagging. If the matching score is above a predetermined tagging threshold T_g, the concept tag of prototype image is assigned to image in UIDS automatically. These recorded matching scores and concept tags assigned to images of UIDS are stored in a concept possibility matrix file. Here we define the term 'Concept Possibility' as, *the degree of certainty with which a particular image belongs to a particular concept*. The sum of the matching scores for a particular image over different tags is greater than one; hence we consider the matching score as the possibility.

In the next stage, query image, concept possibility matrix file with matching scores and tags of concept are given as input to algorithm 2. Figure 3 shows the steps used for concept detection and assigning tags to the unknown query image. The features of the query images are extracted. These features are normalized in the range 0 to 1 and stored as query image feature vector. Query image feature vector is compared with image feature vectors in tagged image dataset (TIDS).All the images in TIDS with matching score above the concept matching threshold T_c, are stored as subset of images(M). Now we find concept of query image from the tags and their recorded matching scores of TIDS images. For each concept, we extract tags from concept possibility matrix file. For each tag of a concept, possibilities of all images of individual concept in the subset images are averaged. Thus, we obtain vector giving possibility of query image belonging to every concept and tags of that corresponding concept are assigned to query image. The concept with maximum possibility is the concept of query image.

Experimental results show that the model created increases concept detector effectiveness. In the proposed workflow the user gives a query image and its concept is detected automatically.

Algorithm-1. Automatic Image Annotation

Input: Untagged image dataset *UIDS* with *n* images per concept and manually tagged dataset *MIDS* of *m* original concepts.
Output: Concept possibility matrix files with matching score and tags of concept.
1. Select one prototype image per concept from *MIDS* and manually tag it, extract and store their features.
2. Extract features of all untagged images in dataset *UIDS*.
3. Compare features of *UIDS* images with prototype images using normalized Euclidian distance. Find matching score *S.*
4. If the matching score $S[0,1]>T_g$(Tagging Threshold), then assign all the tags of concept prototype image to the *UIDS* image with the matching score as S. Store the concept and its matching score in a concept possibility matrix file.

Algorithm-2. Visual Concept Detection and Tagging

Input: Query image, concept possibility matrix file with matching score and tags of concepts.
Output: Concept of Query image with tag.
1. Extract features of the query image.
2. Compare these features with features of tagged image dataset and find subset of images (M) containing matched images above score T_c (Concept Matching Threshold). Find concept of query image from the tags and their recorded matching scores from images in concept possibility matrix file.
3. For each concept, extract tags from concept possibility matrix file. For each tag of that concept, find the average of possibilities over M images. Find the average of these tag possibilities. The concept with maximum possibility is the concept of query image.

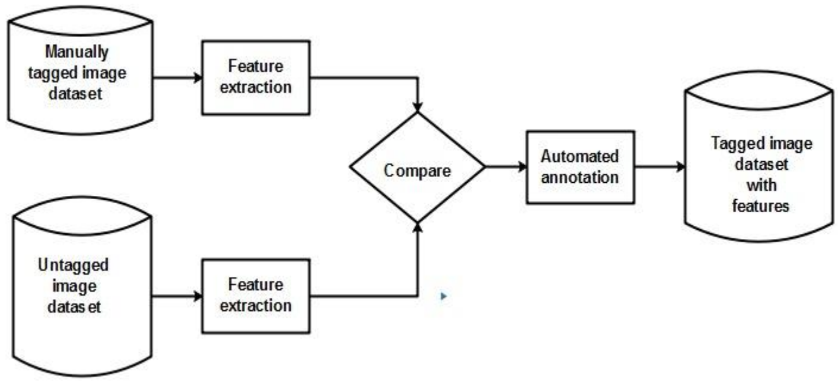

FIGURE 2: Automatic Tagging of UIDS using MIDS.

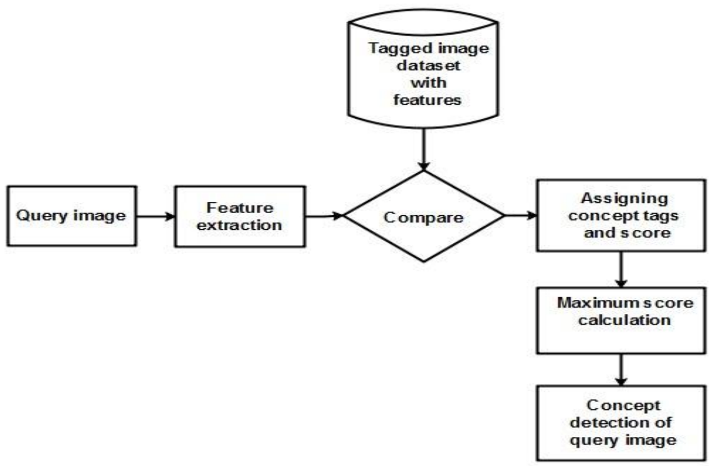

FIGURE 3: Concept Detection of Unknown Images using TIDS.

4. EXPERIMENTAL SETUP

The goal of our experiments is to detect and annotate the concept of an unknown image from a large image collection. The dataset consists of 1,000 images. Out of these 600 images are used for training and 400 images are used for testing. The experiment was repeated three times; for each run we have selected different combinations for training and testing images, in the ratio 600:400.Two classifiers K-means Nearest Neighbor (KNN) and Artificial Neural Network (ANN) along with our proposed approach are tested on Corel dataset.

Concept detection is implemented in two stages. In the first stage, untagged images are automatically assigned tags and possibities from tagged prototypes images using thresholding. Here 0.85 is used as a tagging threshold T_g for tagging untagged images based on feature matching score of these images with tagged images. Ideally this threshold must be one (for identical images). But two images from the same class may have some feature dissimilarities. Hence 15 percent tolerance is kept intuitively to take care of these dissimilarities. The second stage detects concept of query image from tagged training image dataset. For every query image passed to the model, a query image feature is extracted and compared with tagged image dataset. For finding the subset of images we have used concept matching threshold T_c varying from 0.65 to 0.95 in steps of 0.05. The decision for optimal threshold can be made based on the F-Score [17], which measures the test's accuracy. Based on the analysis of F-Scores obtained for various values of thresholds, it is observed that threshold value of 0.85 gives the highest F-Score. Hence it is selected as optimal threshold that gives the best possible precision and recall combination.

5. ANALYSIS OF RESULTS

Our experiments are performed on Wang's dataset, which contains natural images of different categories. We have experimented with various thresholds for annotation and classification. The optimum threshold has been used in our experiment. We have used color and texture features. The performance of the implemented system is evaluated by comparing the results of our proposed classifier with KNN and ANN classifiers with different sets of feature weights.

5.1 Performance Benefits of Selected Weighted Features

Our model has achieved highest concept detection accuracy with the weighted combination of Color moments, HSV histogram (color features) and Wavelet Transform (texture feature). The weighted fusion of features is advantageous as it correlates the multiple features at the initial stage of annotation and helps in achieving better detection rate [18]. Table 2 shows performance of features weights of our model.

As seen from the bar graph (Figure 4) the combination of weights 1, 3, 1 (1*CM, 3*HSV Hist., 1*WT) has the highest accuracy point. These bar graphs were calculated with 1000 images and then mean accuracy was plotted. The effect of different weights was computed in order to explore different global features can contribute differently to the classification scheme. After classification the tags were annotated to the images. This annotation can further be used to classify new images added to the dataset. The weighing factor was chosen to be 1 and 3 because effect of specific global features can only be seen once it is enhanced by the factor of 2 and more. It was computationally difficult for us to calculate all the weighing factors. Hence we limited our studies to only 1 and 3.

Feature Weight (CM,HSV,WT)	Run Number			% Accuracy
	1	2	3	
(1,1,1)	89.72	89.54	90.01	89.75
(1,1,3)	89.31	89.37	89.12	89.26
(1,3,1)	**92.29**	**92.93**	**93.22**	**92.81**
(1,3,3)	89.77	89.13	90.11	89.67
(3,1,1)	90.32	89.89	90.28	90.16
(3,1,3)	89.26	88.02	89.49	88.92
(3,3,1)	92.01	91.21	91.21	91.47
(3,3,3)	90.19	90.34	90.19	90.24

TABLE 2: Performance of Feature Weights.

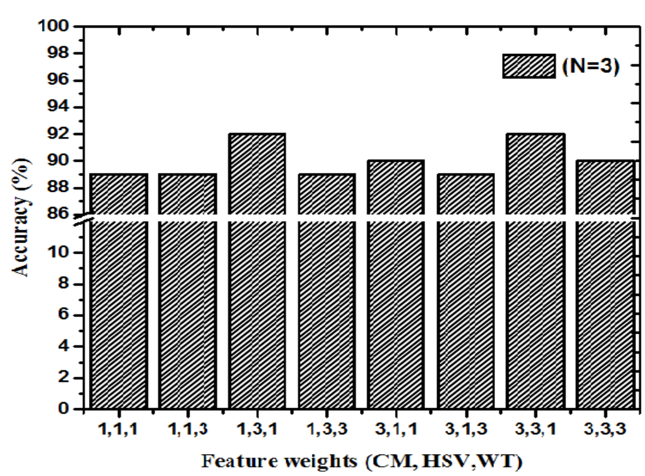

FIGURE 4: Feature Weights versus Accuracy.

5.2 Classifier Evaluation

Table 3 shows the confusion matrix which checks for any confusion between two classes, i.e., it gives information regarding mistagging of an image of one class to another class, if any. It also gives information about the true and false predictions made by concept detection model. The performance of our proposed system is evaluated by computing accuracy, true positive rate (TP Rate) and false positive rate (FP Rate) defined in equations 4, 5 and 6 respectively. Figure 5 shows bar chart of actual versus predicted class (test images) for our proposed approach. It is observed that few images are misclassified. This may be due to similar background, thus making them semantically similar and causing confusion.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \qquad \text{K (4)}$$

$$TPRate = \frac{TP}{TP + FN} \qquad \text{K (5)}$$

$$FPRate = \frac{FP}{FP + TN} \qquad \text{K (6)}$$

where *TP* is the set of true positive images that are related to the corresponding concept class and are classified correctly, *FP* is the set of true negative images that are irrelevant to the corresponding concept class and are classified incorrectly, *FN* is the set of false positive images that are related to the corresponding concept class but are misclassified, *TN* is the set of true negative images that are irrelevant to the corresponding concept class and are classified correctly.

Output Class

27	3	0	2	2	0	0	1	3	2
4	22	0	2	5	2	1	4	0	0
3	6	20	0	4	0	1	2	0	4
0	0	0	40	0	0	0	0	0	0
3	1	0	9	26	0	0	0	0	1
1	0	0	0	0	33	3	0	2	1
1	0	0	2	2	1	34	0	0	0
2	1	3	7	4	0	0	21	0	2
0	5	0	3	2	1	1	0	25	3
4	2	1	2	5	1	2	0	0	25

Target Class

TABLE 3: Confusion matrix for concept matching threshold $T_c = 0.85$.

The ROC curve (Figure 6) is plotted for different values of thresholds (T_c) as discussed in section 4. From the ROC curve the images belonging to class elephant is nearest to the main diagonal of the ROC curve, indicating the lowest performance. The dinosaurs and African are very close to the top left corner of the graph, whereas bus, flower and horse are closest to this ideal region of the curve. The reason for beach and elephant not classified accurately is large variation in the ratio of foreground to background pixels. For the mountain, monument, beach and elephant classes the percentage background sometimes exceeds the size of the object leading to misclassification and hence reducing the overall accuracy of the algorithm. The class beach (yellow for sand and blue for water) and elephant (green for jungle and black for elephant) has bicolor distribution this affects overall classification and leads to more number of false positives and false negatives. However, in the case of dinosaur this bicolor distribution did not affect any result because the background was uniform white.

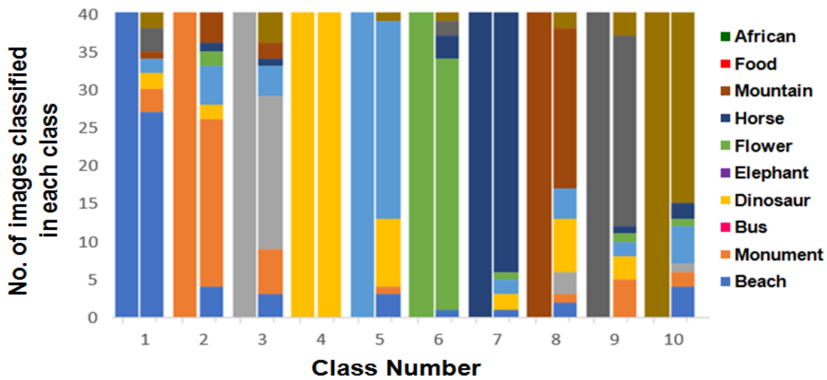

FIGURE 5: Bar chart showing number of correctly and incorrectly classified images for each class.

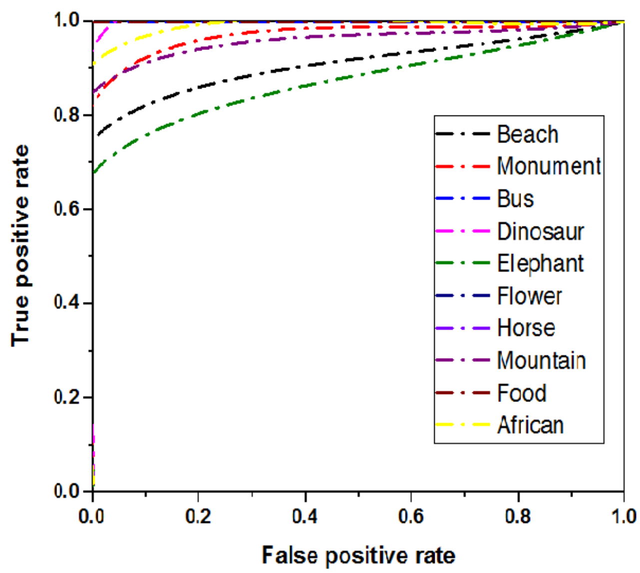

FIGURE 6: ROC Curve for The Proposed Classifier.

5.3 Comparison of The Proposed System with Other Systems

In this section, we compare result of our proposed system with some earlier work. The results of proposed system are compared with K-Nearest Neighbor (KNN) and Artificial Neural Network (ANN) classifiers. KNN is supervised machine learning classifier and is implemented for K = 5. ANN is implemented using backpropagation learning algorithm with feed forward neural network having fifteen neurons in hidden layer. A confusion matrix is used to validate the accuracy of both the classifiers.

As seen from Table 4 the proposed algorithm has better accuracy. For the same color and texture features there was 28 percent improvement in accuracy. The performance of KNN and ANN were almost equivalent, but ANN has more variance and slightly more accuracy. The last column in this table shows the average classification time over entire image dataset, which clearly shows the superiority of proposed approach over KNN and ANN. The performance comparison of various concept detection systems found in literature is given in Table 5. It is found that, the proposed concept detection system based on a novel tagging approach has much less time complexity of classification step, and results in significant improvement in accuracy as compared to the other tagging approaches. The proposed approach is specifically useful when a small

Algorithm	Run Number			% Accuracy	Classification time per image (Sec)
	1	2	3		
KNN	74.7	75.2	75.8	75.23	0.126
ANN	74.36	77.10	75.40	75.62	0.083
Proposed approach	92.29	92.93	93.22	92.81	0.064

TABLE 4: Comparative result of proposed approach with KNN and ANN classifiers.

Authors/Year	Tagging	Approach/ Classifier	Features	No. of Classes	No. of images in Dataset	Accuracy %
As listed	Yes (Tagging a small exemplar dataset)	Proposed (Thresholding and Euclidean distance)	CM, HSV and WT	10	1000	92.81
Gupta et.al [19] 2012	No	ANN	Color Moment and DB wavelets	3	900	82.66
N. Ali et al[20] 2016	No	SVM	Bag of Features	10	1000	86.27
Z. Mehmood et.al[21]2018	No	SVM	HOG and SURF	10	1000	80.61
N. Zhou et al[22]2011	Yes	HPM framework	Bag of words (ASPH and CSPH)	50	5000	65.68
R. Hong et al[23]2014	Yes	Multiple-Instance Learning	CC,CM,HSV, PWT and EH	20	2000	69.9
J. Kim et al[24]2012	No	KNN	Bag of words	4	3500	78.03

TABLE 5: Performance Comparison with Contemporary Works.

tagged image dataset is available, that can be used for fast concept detection and retrieval from huge image dataset.

6. CONCLUSION AND FUTURE WORK

In this paper we have presented a concept detection system based on the novel tagging approach combined with global, low level image feature extraction. The approach allows us to automatically annotate a large image dataset from the small dataset of tagged images. Even though we have not used segmentation and object recognition steps, we have achieved a very good accuracy. It is shown that (Table 5), the proposed approach outperforms all types of existing approaches including tag based, low level feature based as well as hybrid approach (combination of tags and low level features), in terms of accuracy. The classification time is also less as we are using simple Euclidean distance-based image feature matching.

The proposed approach is very useful for domain specific applications where a small tagged image dataset is available, that facilitates quick concept detection as well as retrieval from potentially very huge image dataset of the same domain.

In our further research we plan to incorporate local features within our model. We propose to detect objects of an image and extract their features, combine these local features with global features. Use of these combined features with our algorithm is expected to significantly reduce the concept detection error and improve the performance of our model.

7. REFERENCES

[1] P.G. Foschi, D. Kolippakkam, H. Liu and A. Mandvikar, “Feature Extraction for Image Mining”, *Proc. in Int. workshop on Multimedia Information System*, 2002, pp. 103-109.

[2] M. Wang and X. Hua, “Active Learning in Multimedia Annotation and Retrieval: A Survey”, ACM Transactions on Intelligent Systems and Tech., 2011.

[3] M.I. Mandel, R. Pascanu, D. Eck, Y.Bengio, L. M. Aiello, R. Schifanella, and F. Menczer, “Contextual Tag Inference”, ACM Trans. on Multimedia Computing, Communications and Applications, 2011.

[4] M. Naphade and J. R. Smith, “On the Detection of Semantic Concepts at TRECVID”, Proc. of the 12th Annual ACM Int. Conf. on Multimedia, New York, USA, 2004, pp.660-667.

[5] R. Min and H.D. Cheng, “Effective Image Retrieval using Dominant Color Descriptor and Fuzzy Support Vector Machine”, Pattern Recognition, Vol.42, No.1, Jan.2009, pp. 147-157.

[6] Y. Yang, Y. Gao, H. Zhang, J. Shao and T. S. Chua, “Image Tagging with Social Assistance”, Proceedings of International Conference on Multimedia Retrieval, ICMR’14,Glasgow,United Kingdom, Apr. 2014.

[7] S. Papadopoulos, C. Sagonas, L. KompatsiarisI and A. Vakali, “Semi-supervised Concept Detection by Learning the Structure of Similarity Graphs”, Advances in Multimedia Modeling, Lecture Notes in Computer Science, Vol. 7732, Springer, Berlin, 2013.

[8] C. Krishna Mohan and B. Yegnanarayana, “Classification of sport videos using Edge Based Features and Autoassociative Neural Network Models”, SIViP, Vol.4, No. 1, Mar. 2010, pp. 61–73.

[9] L. Feng and B. Bhanu, “Semantic Concept Co-Occurrence Patterns for Image Annotation and Retrieval”, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.38, No. 4. Apr. 2016.

[10] P. Koniusz, F. Yan, P.Gosselin and K. Mikolajczyk, “Higher-order Occurrence Pooling for Bags-of-Words: Visual Concept Detection”, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 39, No. 2, 2016.

[11] T. Tsirelis and A. Delopoulos, “Automatic Ground-truth Image Generation from user Tags”, 12th International Workshop on Image Analysis for Multimedia Interactive Services, 2011.

[12] X. Li, “Tag Relevance Fusion for Social Image Retrieval”, Multimedia Systems, Vol.23, No. 1, Feb. 2017, pp 29–40.

[13] A. Elmoataz, F. Nouboud, O. Lézoray and D. Mammass, “Editorial of the Special Issue on Advances in Low-Level Image Representations for Processing and Analysis”, SIViP, March 2016, Vol.10, Issue 3, pp. 421–422.

[14] P. Shrivastava, K.K. Bhoyar and A.S. Zadgaonkar, “Image Classification Using Fusion of Holistic Visual Descriptions”, Int. Jour. Image, Graphics and Signal Processing, 2016, pp.47-57.

[15] M.J. Swain and D.H. Ballard, “Color Indexing”, Int. Jour. of Computer Vision, Vol. 7, Issue 1, 16- 21, Nov.1991, pp. 11-32.

[16] G.V. Sangamnerkar and K.K. Bhoyar, “A Neural Network Color Classifier in HSV Color Space”, Int. Conf. on Industrial Automation and Comp., 12-13th Apr. 2014.

[17] Z.C. Lipton, C. Elkan and B. Naryanaswamy , "Optimal Thresholding of Classifiers to Maximize F1 Measure", Calders T., Esposito F., Hüllermeier E., Meo R. (eds) Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2014. Lecture Notes in Computer Science, Vol. 8725. Springer, Berlin, Heidelberg.

[18] E.Guldogan and M.Gabbouj, "Feature Selection for Content Based Image Retrieval", SIViP, Vol. 2, No. 3, Sep. 2008, pp. 241–250.

[19] D. Gupta, A.K. Singh., D. Kumari and Raina, "Hybrid Feature Based Natural Scene Classification using Neural Network", Int. Jour. of Computer Apps., 2012, pp. 48-52.

[20] N. Ali, K.B. Bajwa, R. Sablatnig, Z. Mehmood, "Image Retrieval by Addition of Spatial information based on Histograms of Triangular Region", Computer and Electrical Engineering., Vol. 54, Aug. 2016, pp. 539-550.

[21] Z. Mehmood, F.Abbas, T. Mahmood, M.Arshad Javid, A. Rehman, "Content-Based Image Retrieval Based on Visual Words Fusion versus Features Fusion of Local and Global Features", Arabian Jour. for Science and Engineering, Jan. 2018.

[22] N. Zhou, W.K. Cheun g, G. Qiu, and X. Xue, "A Hybrid Probabilistic Model for Unified Collaborative and Content-Based Image Tagging", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 33, NO. 7, Jul. 2011.

[23] R.Hong, M.Wang, Y.Gao, D.Tao, X. Li, and X.Wu, "Image Annotation By Multiple-Instance Learning With Discriminative Feature Mapping and Selection", IEEE Transactions On Cybernetics, Vol. 44, No. 5, May 2014.

[24] J.Kim, B.S.Kim and S.Savarese, "Comparing Image Classification Methods: K-Nearest-Neighbor and Support-Vector-Machines", Proceedings of the 6th WSEAS international conference on Computer Engineering and Applications, USA, Jan. 25 – 27, 2012.