ISSN2394-3777 (Print)
ISSN2394-3785 (Online)
Available online
atwww.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Special Issue 9, March 2018

AUTOMATED IMAGE LABELING USING SALIENT FEATURES OF UNKNOWN IMAGE

A.Sathiya[1]
Department of Computer Science and Engineering
Sathyabama University

N.S.Sudharsan[2]
Assistant Professor
Department of Computer Science and Engineering
T.S.M.Jain College of Technology

Abstract

Vigorous learning is useful in situations where labeled data is scarce, unlabeled data is available and labeling has some cost associated with it. In such situations energetic learning helps by identifying a minimal set of items to label that will allow the training of an effective classifier. Most existing image processing applications are designed for small-scale and local computation which does not scale well to web-sized problems with their large requirements for computational resources and storage, since a large number of images are not labeled properly which is considered as weakly labeled and some of the images were not labeled at all. In this Map Reducing concept is used to overcome this massive labeled and un-labeled analysis problem, since traditional data processing application software is inadequate to deal with them. In this a novel method is developed for achieving multi-label, multi-feature image annotation using a Map Reduce framework, where an image-level labels and region-level labels for both labeled and un-labeled images are obtained. The associations between semantic concepts and visual features are mined both at the image level and at the region level through which Multi-Label image correlations are obtained by a co-occurrence matrix of concept pairs using Convolution Neural Network.

Keywords: Chain code, Map Reducing, Social media Support Vector Machine, Un-labeled image, Weakly Labeled image.

I. INTRODUCTION

Nowadays, technology has become more and more advanced. This has led to cheap and superior multimedia strategies which have setascend to enormous data volumes. This gigantic data desires to be stored in database for a variety of applications. These applications have shaped anobligation for efficient and well-organized methods of storage, explore and reclamation of images via parallel processing techniques.Earlier, the images were stored with associated labelsand searching was done on the basis of these labels.

But, such a method is prone to many mistakes. If the multimedia image data's are wrongly annotated, irrelevant images might be retrieved. Furthermore, it is a arduousjob tosearch for an image which has been assigned a wrong label [1].

Multi-Label image annotation aims to gain knowledge of the relationship between the visual features and the pre known features (i.e.) labels [2]. The objective is to build up methods that can interpret a new image with relevant keywords from a knownterminology set. These results can be used as a tag suggestion for labeling images, or for image retrieval tasks, etc.

In existing multi-label image annotation learning technique, the comparisons among numerous labeled and un-labeled images and the relations between their illustration features and the relative ideas are not exploited satisfactorily to perk up the annotation performance [3].

Most of the preceding work didn't reflect on the segmentation step more scrupulously. Our proposed method introduces a system that applies multi-label annotation using the Map Reduce framework for obtaining features in both weakly labeled and un-labeled images and it also gives more concentrationinthe segmentation processof these images to provide better precision.

II. STATE OF ART

Modern days have witnessed the rapid expansion of multimedia data's over the internet. Social networks play a major role in sharing such data's in enormous amount. Websites like Facebook and Flickr allow users to upload multimedia images and illustrate the image content with tags. on the other hand, as mentionedearlier than, theimages on these websites are not fully labeled or even labeled. Elisseeff et al. [4] proposed a technique to deal noise in tag information using kernel values of images. Liu et al. [6] used a novel approachwith semi-parametric regularization for exploiting labeled and unlabeleddata and also for its optimization. Mostof the approaches use employed handcrafted features like[5] for solving imageannotation problem.

.

Even though the handcrafted features have made a huge improvement in multi-label annotation, the extracted featuresare not always up to the mark. Many recent researches use Convolution Neural Network (CNN) for addressing multi-label problems. It has been addressed that CNN have outperformed offered handcrafted featuresin most of the applications. Krizhevsky et al. [7] uses CNN for addressing multi-labeling problem which consists of images from 1,000 categories and achieved better result.

Jiebo Luoet al [8] proposes semi-supervised learning algorithm which is based on the kernel density estimation approach for automatic video annotation.Fei Wu et al. [9] proposed a system WeSed which used deep CNN for solving multi-label annotation with weakly labeled and un-labeled images. Fergus et al. [10] uses visual similarity between images to model the unlabeled data.In most of the existing system,for solving multi-label image issue, the comparison among multiple image tags and theassociations between their visual features and other semanticconcepts are used, but these approaches are not sufficientfor improving the interpretationperformance.

In this research, the proposed system tries to resolve the image taggingannotation problem using an enhanced pattern matching technique along with Support Vector Machine (SVM) classification technique. The system considers only NUS-WIDE dataset images for training and testing, itcontains a large set ofimages with incomplete label information.

The proposed approach is done by using following steps i) collecting patterns of real world object ii) class generation using Support Vector Machine iii) Segmentation of giving un-labeled image iv) identification of pattern using SVM v) labeling objects in given un-labeled image. This approach shows promising result than the existing approaches.

III. SYSTEM OVERVIEW

A)COLLECTING PATTERNS OF REAL WORLD OBJECT

The first and foremost work is to generate the pattern for recognition. Much work related to pattern generation has been done for multi-labeling problem. G.Carnerio et al [11] proposed a method for multi-label annotation using pattern. Most of the pattern recognition is done either by color histogram or with a certain predefined pattern.

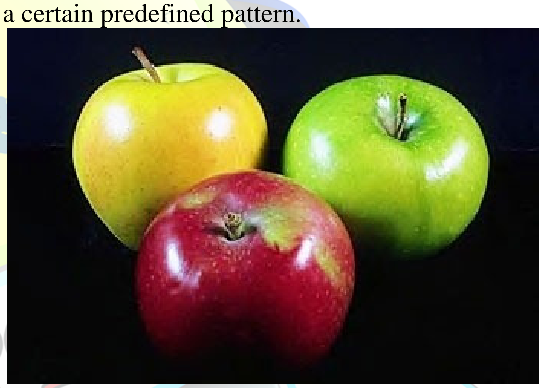

Fig. 1 Objects with color values

The proposed system takes both the color histogram values along with edge values of an object and these features together form the pattern for object identification. Fig1 shows the original color image of the objects.

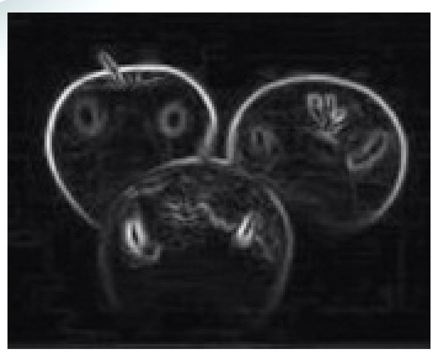

Fig. 2 Edge detection of the objects

ISSN2394-3777 (Print)
ISSN2394-3785 (Online)
Available online atwww.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Special Issue 9, March 2018

Fig 2 shows the edge vectors of the object and it is obtained using Sobel Edge Detection methods. The edge pattern is covered using Chain code representation. Boundary representation is given by the chain code representation [12]. In Chain Coding method,vectors between successive edge pixels are determined.

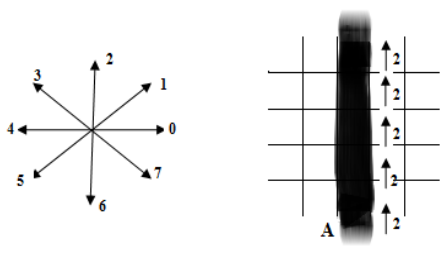

Fig. 3 Chain code for number '1'(010 010 010 010 010)

Typically, the Chain Code contains the start pixel vector address followed by a cord of code words. Such codes can be widespread by increasing the number of allowed path vectors between successive edge pixels of an object. Chain code representation is shown in Fig. 3.

The code will be more precise for the high definition image; Chain code for all objects will be identified in a similar manner as in Fig 4 and Fig 5 in many samples. It is obvious that Chain Code for trees in Fig 4 will be more or less similar, thus the pattern code for the trees will be generalized as 17464241 with the help of Chain code formed from trees A, B, C in Fig.4. In contrast, Fig 5 shows dissimilar Chain code representation since the orientation of object differs.

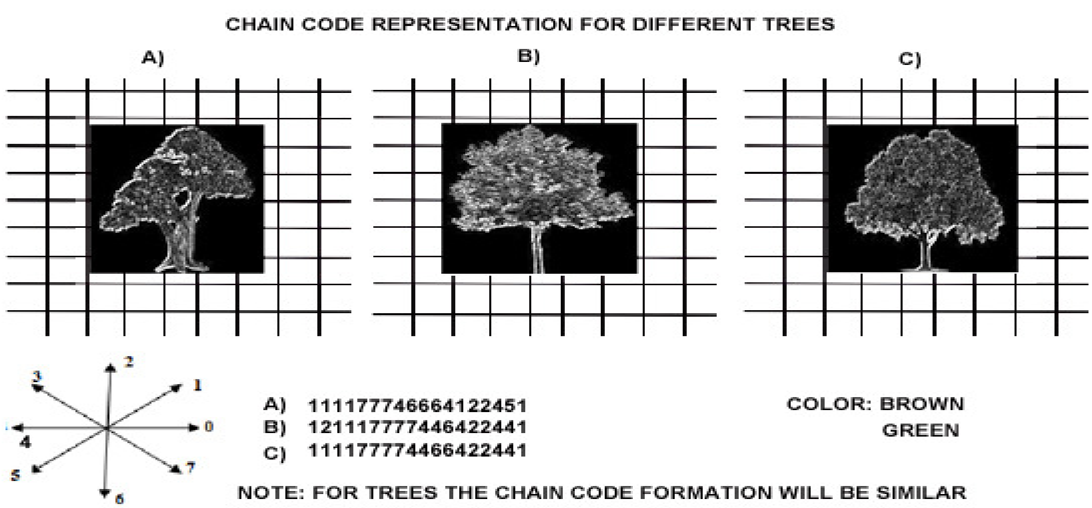

Fig 4 Chain code patterns for trees

ISSN2394-3777 (Print)
ISSN2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Special Issue 9, March 2018

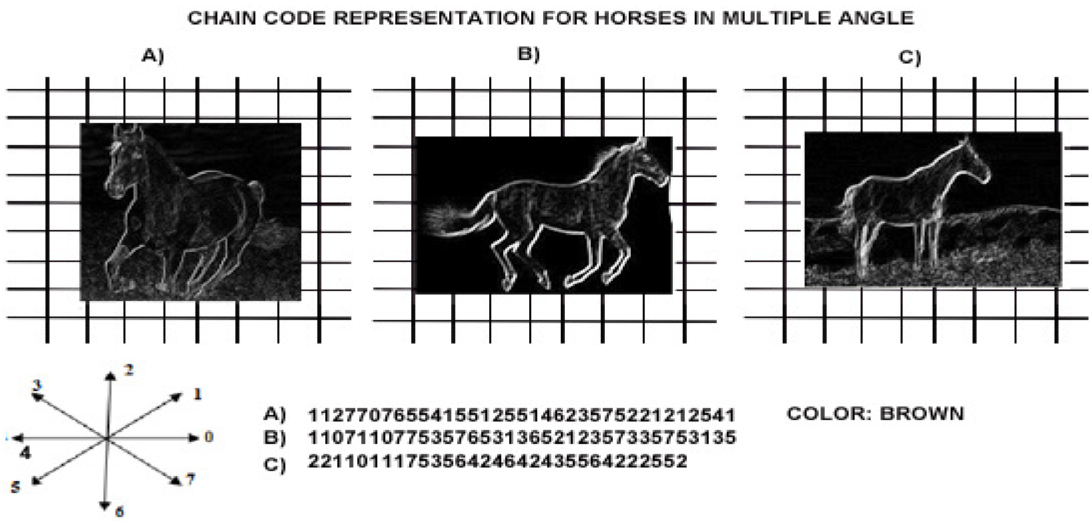

Fig 4 Chain code patterns for horses

This process has to be done for multiple horses; this will lead for generalization of horse pattern so easily. Likewise the chain code pattern will be obtained for all real world objects. The Chain Code along with the color histogram value forms the feature for SVM Classification. The direct pattern recognition method will yield slow recognition since it involves with 3*3 or 4*4 mapping. The images in the social network are in large number and with different features.The mapreduction features in proposed method will yield more precise and efficient result in the Big Data analysis.

B) Class generation using Support Vector Machine

Support vector machine [13] is wellorganizedcategorization technique used to categorize unknown data. It is one of the best machine-learning techniques ever proposed. The main advantage of SVM, is that it takes into descriptionof both experimental data and structural behavior for better generalization. The aim of SVM is to obtain the hyper-plane that separate two or more classes efficiently. There are many hyper-planes between two classes for separation, but SVM tends to find the best hyper-plane in all of that hyper-plane.

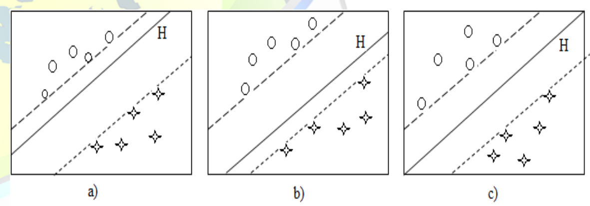

Fig. 5. Hyper-plane (H) differentiation for same class with different factor

In such cases the hyper-planes are in the form $f(x)=w.x+b$. SVM finds an optimal hyper-plane. There are two methods for finding optimal hyper-plane [8], in that the first method is by finding a plane that bisects two closest points of two convex hulls defined by a set of points of each class. The other methods are by maximizing the margin between two sustaining planes as shown in Fig. 5, where H is the hyper-plane and the dotted lines are marginal for the hyper-plane. In order to find the best hyper-plane, we need to maximize the hyper-plane H1($w.x+b=+1$) from hyper-plane H above and hyper-plane H2 ($w.x+b=-1$) below, till it touches the lowest value of both the classes. The width between the plane H to any hyper-plane H1 or H2 is width w. Let x_n be any point over the plane and x be the point over the plane $w^t x+b=0$, where $\left|w^t x_n+b\right|=1$,

ISSN2394-3777 (Print)
ISSN2394-3785 (Online)
Available online atwww.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Special Issue 9, March 2018

where vector w is perpendicular to the plane in χ space. Take xi and xii on the plane, such that $w^t x^i + b = 0$ and $w^t x^{ii} + b = 0$ which gives $w^t (x^i - x^{ii}) = 0$. Now the projection of x_n - x on w gives the distance between H1 to H $\hat{w}$ = $\frac{1}{\|w\|}$, which in turn gives the distance between H1 and H2 as $\frac{2}{\|w\|}$. Therefore to maximize the margin, we need to minimize $|w| = w^t w$ with the condition that no data point's lies between H1 and H2, this happens when $w.x + b \geq +1$ for $y_i = +1$, and $w.x + b \leq -1$ for $y_i = -1$. Combining the two conditions we have $y_i(w.x + b) \geq 1$. So the problem can be defined as minimizing $\frac{1}{2} w^t.w$, this can be easily solved by introducing Lagrange multipliers. This can be given as follows,

$$\min L(w,b,\alpha) = \frac{1}{2} w^t.w - \sum_{i=1}^{l} \alpha_i (y_i (w^t x_i + b) - 1)$$

w.r.t w and b and maximize w.r.t each $\alpha_i \geq 0$, subject to the constraint that the gradient of $L(w,b,\alpha)$ w.r.t variables w and b leads to

$$L(\alpha) = \sum_{i=1}^{l} \alpha_n - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} y_i y_j \alpha_i \alpha_j x_i{}^t x_j$$.

Maxmize $L(\alpha)$ w.r.t α subject to $\alpha_i \geq 0$ for i= 0,1,……l and $\sum_{i=1}^{l} \alpha_i y_i = 0$, this leads to quadratic solution

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} y_i y_j \alpha_i \alpha_j x_i{}^t x_j - \sum_{i=1}^{l} \alpha_i$$

, the quadratic program will hands us alpha value from which w can be obtained as $w = \sum_{i=1}^{l} \alpha_i y_i x_i$, by Khum Tucker condition for i = 0,1,2,…,l $\alpha_i (y_i (w^t x_n + b) - 1) = 0$, when α_i>0 it gives x_n and it is the support vectors. With theseconditions along with the features of Chain code and Color histogram values the system will classify the given pattern more precisely.

C) SEGMENTATION OF GIVEN UN-LABELEDIMAGE

Segmentation is done for giving un-labeled image in order to segment the images into multiple objects from noise.Image segmentation involves in partitioning of an image into discrete regions containing each pixel with similar attributes. To be significant and helpful for image analysis and understanding, the regions should strongly recount to depict objects or features of interest. First step in image processing is starts with segmentation of image into meaningful patterns.It is done by converting a greyscale or color image into high-level image representation in terms of features, objects, and scenes. Segmentation techniques fall into two category, which is contextual or non-contextual type. Non-Contextual type takes no portrayal of spatial associations between features in an image and collection of pixels together on the basis of some global attribute like grey value, color etc. Contextual techniques make use of these associations, e.g. group of similar color or grey level intensity and close spatial locations.

Here Contextual region growing segmentation is used, since segmentation is more important as it plays a major role in identifying patterns in the given unknown image for labeling. The given image is segmented using color and boundary region. Fig. 7shows the segmented image using RGB threshold value.

ISSN2394-3777 (Print)
ISSN2394-3785 (Online)
Available online
atwww.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Special Issue 9, March 2018

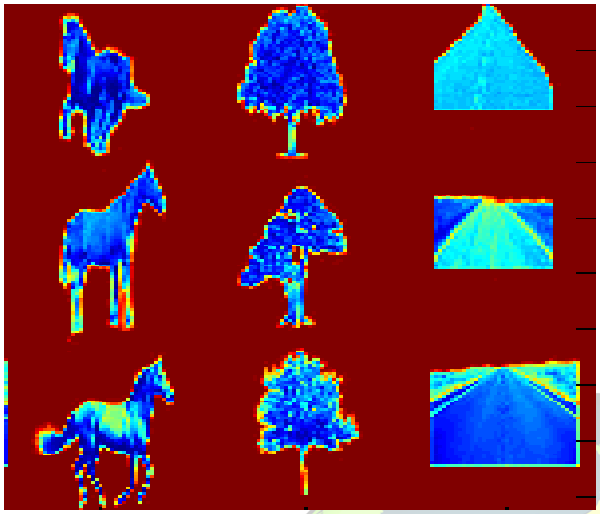

Fig.7 Segmented image

D) IDENTIFICATION OF PATTERN USING SVM

The segmented objects in the given un-labeled image as in Fig.4 are processed under edge detection method and then using Chain Code the features will be obtained for individual object. The features obtained from section D will be given as input to classification testing in SVM.

E) LABELING OBJECTS IN GIVEN UN-LABELED IMAGE.

When testing unknown image, the system will calculate the bag of features like Chain code and color histogram and it will be feed as input to SVM, with that the SVM will return the category of each object in the given testing image using already trained features and thus each object will be labeled so precisely than the other existing method.

IV. EXPERIMENTAL RESULTS

NUS-WIDE unlabeled image dataset is used foranalyzing process [14]. In training, the dataset samples are classified correctly up to slack variable ε_i (margin error), if the training sample predicted wrongly and lies on the wrong side of the hyper-plane, then ε_i must be greater than or equal to 1, the factor C is a constant used to maximize the margin and minimizing the amount of slack. Thus, by both the margin error and the factor C the Lagrange multipliers can be expressed in the term as,

$$\underset{\alpha}{\max imize} \sum_{i=1}^{l} \alpha_n - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} y_i y_j \alpha_i \alpha_j x_i{}^t x_j$$

Subject to: $\sum_{i=1}^{l} \alpha_i y_i = 0$, $0 \le \alpha_i \le C$.

The error rate is calculated using Cross Validation (CV), in this 10% percent of the training data is used for testing and the remaining data is used for training the system. In this proposed system the CV error is computed using different values of ε_i and C in SVM classifiers. Finally the value for ε is 0.0156 and for C is 10 gives the lowest CV error, thus it is used for training the system.

<table>
<tr><th colspan="3">Performance Evaluation of proposed system</th></tr>
<tr><td>No of un-labeled samples</td><td colspan="2">80,000</td></tr>
<tr><td>No of Training samples</td><td colspan="2">72,000</td></tr>
<tr><td>No of tested samples</td><td colspan="2">8,000</td></tr>
<tr><td>Correctly labeled images</td><td>Not determined</td><td>Wrong prediction</td></tr>
<tr><td>89%</td><td>4.1%</td><td>6.9%</td></tr>
</table>

Table.1 Performance evaluation table

The Cohen's kappa coefficient measure is used for analyzing the performance of proposed system. The Cohen's kappa coefficient is a statistical measure of the inter-rater agreement for qualitative items. The coefficient measure can be given by the following formula.

$$\kappa = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)} .$$

Where Pr(a) is the relative observed agreement among raters, and Pr(e) is the hypothetical probability of chance agreement, using the observed data to calculate the probabilities of each observer randomly saying each category. κ indicates the agreed rate. The proposed system shows accuracy rate of 89%. Fig 8 shows an accuracy rate of the proposed system.

ISSN2394-3777 (Print)
ISSN2394-3785 (Online)
Available online
atwww.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Special Issue 9, March 2018

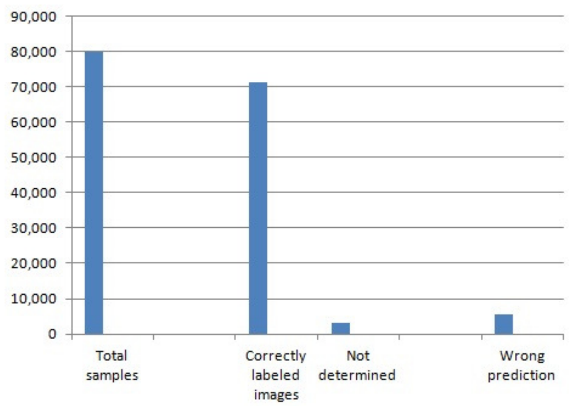

Fig. 8 Performance graph

V. Conclusion And Future Work

Most of the research uses Neural Network and other common Pattern Matching technique for multi-label annotationproblem. However,this processcost more and it involves a lot of possibly uncertaindecisions. The proposed system uses Chain code and color histogram as features for training system using SVM, which shows better performance in labeling un-labeled or weakly labeled images than other systems. The undetermined results and wrong prediction could be avoided by training the system with more number of object patterns.

VI. References

[1] Sitalakhmi and S. kulkarni "MapReduce neural network framework for efficient content based image retrieval from large datasets in the cloud" IEEE Transactions on Hybrid Intelligent Systems, pp. 63 – 68,2012.

[2] Y. Gong, Y. Jia, T. Leung, A. Toshev, and S. Ioffe, "Deep convolutional ranking for multilabel image annotation," CoRR, vol. abs/1312.4894, 2013. Available: http//arxiv.org/abs/1312.4894.

[3] R. Jin, S. Wang, and Z.-H. Zhou. Learning a distance metric from multi-instance multi-label data. In CVPR, 2009.

[4] A. Elisseeff and J.Weston. A kernel method for multilabelled classification. In NIPS, 2002.

[5] A. Makadia, V. Pavlovic, and S. Kumar. A new baseline for image annotation. In ECCV, 2008.

[6] D. Liu, S. Yan, Y. Rui, and H.-J. Zhang. Unified tag analysis with multi-edge graph. ACM MM 2010.

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classificationwith deep convolutional neural networks," in Proc. Adv.Neural Inf. Process. Syst., 2012, pp. 1097–1105.

[8] Fei Wu, Zhuhao Wang, Zhongfei Zhang, Yi Yang, Jiebo Luo, Fellow, IEEE, Wenwu Zhu, Fellow, IEEE, and Yueting Zhuang, "Weakly Semi-Supervised Deep Learning for Multi-Label Image Annotation". IEEE TRANSACTIONS ON BIG DATA, VOL. 1, NO. 3, JULY-SEPTEMBER 2015.

[9] M. Wang, X. Liu, and X. Wu, "Visual classification by '1-hypergraph modeling," IEEE Trans. Knowl. Data Eng., vol. 27, no. 9, pp. 2564–2574, Sep. 2015.

[10] R. Fergus, Y. Weiss, and A. Torralba, "Semi-supervised learning in gigantic image collections," in Proc. Adv. Neural Inf. Process. Syst., 2009, pp. 522–530.

[11] G. Carneiro, A. B. Chan, P. J. Moreno, and N. Vasconcelos, "Supervised learning of semantic classes for image annotation and retrieval," IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 3, pp. 394–410, Mar. 2007.

[12] F. Arrebola, A. Bandera, P. Camacho and F. Sandoval, "Corner detection by local histograms of contour chain code", IEEE Electronics Letters, vol. 33, 9th October 1997.

[13] Kathy F Chen, "Offline and Online SVM performance analysis", Massachusetts Institute of Technology, February 2007.

[14] Available: http://lms.comp.nus.edu.sg/research/ NUS-WIDE.html