AUTOMATIC IMAGE TAGGING BY USING IMAGE CONTENT ANALYSIS

Lailatul Qadri binti Zakaria[a], Paul Lewis[b], Wendy Hall[c]

Pusat Pengajian Teknologi Maklumat,
Fakulti Teknologi dan Sains Maklumat,
Universiti Kebangsaan Malaysia
laila@ftsm.ukm.my[a]

Intelligence, Agents, Multimedia (IAM) Group
School of Electronic and Computer Science
University of Southampton
phl@ecs.soton.ac.uk[b] wh@ecs.soton.ac.uk[c]

ABSTRACT

Image analysis and in particular automatic image classification or annotation, often begins with feature extraction for content representation. The content representation can be based on low level and high level feature extraction. In this paper, we present a simple image classifier which attempts to classify building images based on their low level features which are colour and edges. The classifier is developed by using Bayesian Inference method provided by Infer.net tool. Equivalently, the classifier can be regarded as an annotator which aims to annotate images with a specific tag when appropriate. In this example, the image classifier is developed to identify and distinguish building and non images. The method could be extended into other classes such as mountains, beaches and forest. This image classification is important because it can enable users to retrieve images that may not be well tagged and also to annotate images with information that they may want to use for retrieval purposes but not necessarily for explicit annotation. The performances are assessed using confusion matrices and ROC curves.

Keywords: Image Processing, Automatic Image Tagging, Colour Analysis, Edge Analysis.

1 INTRODUCTION

With the increase of storage capacities for digital cameras, they can generate hundreds of photos on an ongoing basis, and hence contribute to the huge number of photos uploaded in Web 2.0 websites. The ranges of the photos are massive, from vacation, sports, weddings, parties, travelling, friends, hobbies, pets, everyday life and many more. Typically, based on the owner's interest and time to spend, these photos are stored with some description such as title, captions and tags. Based on the study done by Ames and Naaman [1], the motivation for tagging can be classified into two main reasons: personal and social. The owner might tag the photos for personal organisation and communication which will add future recall and to facilitate remembering details about the photos. By adding tags to the photos, the owner will ensure that the photos can be easily found by specific people with whom the user might want to share or they may be discovered by anyone who may be interested in the photos.

Nevertheless, these images are usually too often left without any or with very crude descriptions. The only way for the images to be found is by browsing the directories, their name providing usually the date and the description with one or two words of the original event captured by the specific photos. Although most Web 2.0 websites do provide text based searching to find images by mapping query concepts with words in image's title, description or tags, the access has become more difficult as the number of photos increased, the photos are ill annotated and the query is done by two to three words only [2].

In an attempt to overcome these problems, the image analysis is used to analyse image content directly. Capturing the semantic description of images is usually motivated by the need for improved Content Based Image Retrieval (CBIR). CBIR is interested in organizing images based on their visual features such as colour, texture and shape. Nevertheless, these visual features do not allow users to query images by semantic meaning and most users are familiar with high level concepts which are normally presented in the form of text. Therefore, image annotation is required to label image with visual terms which provide semantic understanding about the image content. In the first instance, a classifier to tag images containing buildings is presented.

2 RELATED WORK

Szummer and Picard has used colour and texture features with the k-Nearest Neighbour (KNN) technique to classify indoor and outdoor images [3]. The colour features are based on Ohta colour space7 and the image is divided into 32 bins per-channel (32 x 3 channel) while the texture features are computed base on multi-resolution, simultaneous autoregressive model (MSAR). The classification performances are reported at 75.6% for colour features and 83.0% for texture feature separately .Integrating both features, colour and texture, has produced a stronger result which is 90.3%.

Vailaya et al. have used colour and edge features with a Bayesian framework [4]. The classification is done based on hierarchical classification of vacation images. Firstly, the image is classified into indoor or outdoor. A subset of outdoor is furthered classified into city or landscape, followed by a subset of landscape is then classified into sunset, forest and mountain classes. Indoor/outdoor classification is done by inferring LUV colour space, while city/non city is classified by observing edge directions histogram and finally sunset/forest/mountain classification is identified by using colour features in HSV space. The performance rate for indoor/outdoor classification is 90.5% which is comparable to [3]. Edge features has provided the best individual performance rate of 95.3% for city/landscape images. For landscape classifications, colour feature has provided the best accuracy of 96.6 for sunset/forest classification and 96% for forest/mountain classification.

van de Sande et al. has used different colour descriptors (RGB and HSV) to classify scene and object [5]. Such colour descriptors with different invariance such as invariance to light intensity, light intensity shifts and light colour change are tested. Furthermore, the SIFT descriptor which describes the local shape of a region using edge orientation histograms is also used. The analysis is done by employing SVMs, a similar method as presented by [6]. The test images are provided by PASCAL Visual Class Challenge which contains nearly 100000 images of different objects such as aeroplane, bicycle, horse and person. The result shows that integrating colour and SIFT variants have the best performance compared to others in classifying objects in the test set.

In general, most of the reviewed works have focused on natural and artificial scene classification. Features such as colour, texture, shape and edges have been used in the classification. Research done by Oliva et al. provides general clues to classify natural and artificial scenes based on edges/shape features [16]. Artificial scenes such as can be characterized by vertical and horizontal structure, for example a city building composed of tall building exhibiting vertical structure, while more balance vertical and horizontal directions (cross shape) for indoor scenes such as kitchen or living room. Natural panoramic scenes such as beach and field can be characterized by a vertical line. Oblique line (mainly orientations at 45 degree plus or minus 15 degree) can be found in natural images such as mountain, canyons and valleys. Circular orientation is commonly identified in highly texture environments such as forests and fields.

[12] & [4] both agree edges are a good identifier for classifying building/city images and integrating edge with other features such as colour does improve the performance. In most research, natural scene classification is by using colour and texture features. In comparison, colour features have performed better than texture and integrating both of these features has yielded a better performance rate ([7],[8]& [12]).

In term of classification strategies, Bayesian, SVM and KNN are imperfect. KNN is criticised for having a poor run-time performance thus, generally slow and difficult to determine the correct value of k from the validation set [7]. The SVM model run time is also time consuming and it is too large to be used in a practical system with limited space [9]. The Bayesian approach disadvantage is the normality assumption in general pattern recognition literature and neural networks are usually hard to optimise for generalisation [10]. It is also hard to compare the classifier performances between these researches due to two main factors. It is recorded that the Bayes classifier has the ability to classify images more effectively compared to SVM and KNN classification methods in image segmentation [11]. Hence we use a Bayes based machine learning tool from Microsoft called Infer.NET.

3 SYSTEM ARCHITECTURE

The image classifier construction is illustrated in Figure 1. Images are divided into two sets which are a training set and a test set. The training set consists of 210 images (105 building and 105 non building images). These images were obtained from Flickr and carefully allocated to the appropriate set. Figure 2 shows some examples for building and non building images used in the training set. In general, the building image set has a specific focus mainly on structure such as buildings and houses, while the non building image set has a wider focus covering such topics as people, flower, boat and etc. The test set consists of 1040 images (534 building images and 506 non building images).

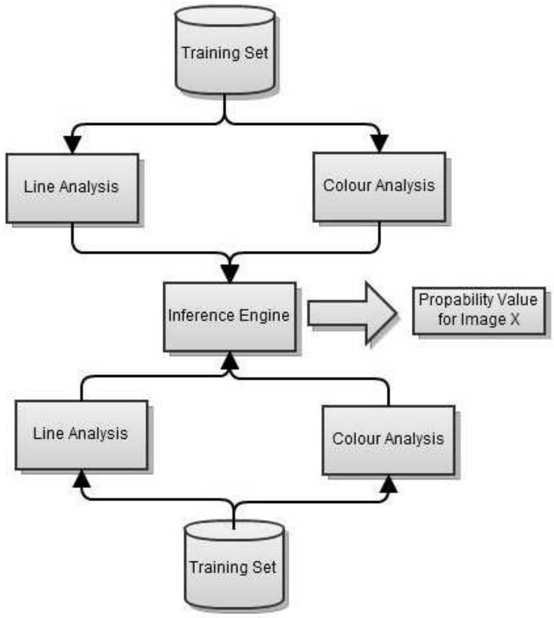


Figure 1: The construction of Building and Non Building Image Classifier

Figure 2: Example of Building and Non Building Images Used in Training Set

To initiate the analysis, known building and non building images (in the training set) are submitted to low level feature extractors for colour and line features. The City Landscape Identifier (CLI) from Photocopain [13] is used in the line analysis to generate line histograms. Both of these analyses produce results in the form of histograms. These histograms are normalized and used as the input to the Inference Engine.

The Inference Engine is implemented using Infer.net to provide a Bayesian Inference tool. The Inference Engine uses the training data in order to generate prediction values in the range 0 to 1.0 for each image in the set. A high value indicates a building image and a low value a non building image. A threshold value on the prediction must be chosen, above which images are tagged as building images. The query image / unknown image is submitted and follows a similar root to images in the training set. Line and colour histograms are extracted and passed to the Inference Engine and a probability value for the unknown image is generated and the threshold applied.

Even though the image classifier is only focusing on identifying building and non building images, it could be expanded to include other classifiers to detect the presence or absence of beaches/ocean, mountains, sunset and forest etc. To create classifiers for generating a wide range of tags it may be necessary to use more powerful low level features such as the visual terms used by some researchers [14] and [15].

3.1 Training Image Analysis

3.1.1 Line Analysis

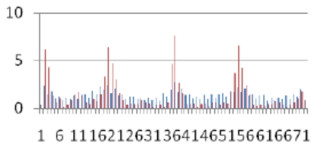
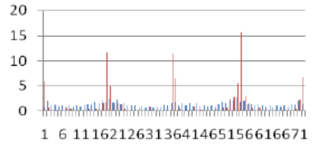
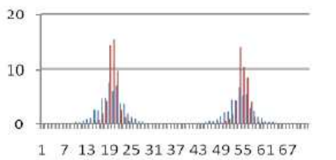
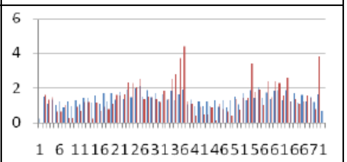
The extraction of line information is done by using an algorithm from the Photocopain system. Photocopain is a content based annotation tool which is integrated with the AKTiveMedia image annotation system to allow users to annotate images semi-automatically [13]. City Landscape Identifier (CLI) is one of the Photocopain components. The CLI is used to represent image content by calculating the edge direction coherence vector. The assumptions are that straight edges tend to appear more in artificial structures rather than natural structures.

The Line Analysis works as follows. An image is submitted to the CLI analyzer and the analyzer identifies edges at 72 directions. The edges are converted into lines and the number of pixels in a straight line section represents its length. Lines shorter than a particular threshold are regarded as incoherent and longer lines are regarded as coherent.

Histograms representing the incoherent lines and the coherent lines are created. The data in the histograms are normalized by summing up the total number of pixels in each direction and dividing each value by the total value. The normalize data gives the proportion of pixels as a percentage for each direction.

Table 1 shows a sample output generated by the CLI analyzer for building and non building images. Building images (Table 1 (a &b)) have generated high calculation of pixel counts in the region of 0, 90, 185 and 360 degree, which are presented in direction value of 0, 19, 36, 56 and 71 respectively. In general, images of buildings will produce a high pixel counts in such a range of angles compared to non building images. Based on the observation, we have decided to analyse images based on certain directions and compared them with the results for all directions.

Table 1: Examples of city and non city images with their normalized line histograms. The red bar in the line histograms represents long line data while the blue bar represents short line data.

Image	Edge Histogram
 (a)Building Image	
 (b)Building Image	
 (c) Non Building Image	
 (d) Non Building Image	

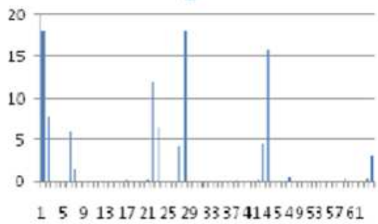
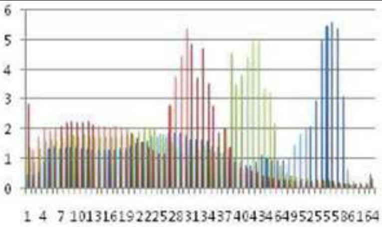
3.1.2 Colour Analysis

Colour analysis here involves extracting a colour histogram for each image to be used as the colour feature in the machine learning process. In the first instance we use the RGB colour model although it would be interesting to experiment with alternative colour models. For each image, the colour histogram is generated to provide information about the distribution of colours in the image. There are two different methods for generating the colour histogram. One is to divide the 3-D RGB colour space into cells and count the number of pixels falling in each cell. The other is to calculate a histogram for each of the three (R,G and B) colour channels separately.

1. *3D Colour Histogram*: In the first method, each pixel in the images is projected into a 3D RGB colour space. In this work the 3D colour space is divided into 4x4x4 and 6x6x6 cells which generate colour histogram with 64 bins and 216 bins respectively. The number of pixels in each cell is counted and stored in the colour histogram. The histogram is normalized by summing up the total number of pixels in each bin and dividing each value by the total value. The normalized data gives the proportion of pixels as a percentage for each bin.
2. *Separate Colour Histogram*: In the second method, the intensity histogram of each colour, red, green and blue, is observed separately. In an RGB image each pixel is represented by three colours, which are red green and blue, and the intensity level for each colour is observed and extracted to generate three colour histograms. Each colour histogram is divided into 64 bins producing 3 colour histograms for each image. The colour histogram is also normalized by using the same normalisation as described earlier.

Table 2 shows examples of building and non building images and their colour histograms generated by extracting colour features from the images. The second column shows the 3-D colour histograms and the third column shows the separate R,G and B histograms superimposed in different colours. For the 3D colour histogram, the x-axis represents the bin number while the fraction of pixels for each bin is indicated in the y-axis. For the 3 separate colour histograms, the x-axis bin numbers indicate the colour intensity while the y-axis indicates the fraction of pixels with that intensity for that each colour.

Table 2: Example of 3D and Separate Colour Histograms for Building Images

Image Example	
3D Colour Histogram	
Separate Colour Histogram	

3.2 Finding the Optimal Threshold Value to Identify Building Images

Three experiments were involved in finding the optimal threshold value to identify building and non building images. These experiments were done using just the training set of images as this is all part of the classifier training process. All of the images in the training set are already labelled with either building or non building tags, thus making it possible to evaluate classification performance with different thresholds. The classification performances are evaluated by using confusion matrices and ROC curves to identify optimum threshold values to classify building and non building images. Detailed descriptions for the processes involved in the experiments are presented for Experiment 1, and the processes are repeated for Experiment 2 and Experiment 3.

1. Experiment 1: Finding threshold values using line histograms
2. Experiment 2: Finding threshold values using colour histograms
3. Experiment 3: Finding threshold values using line and colour histograms

3.2.1 Experiment 1: Optimal Threshold Value Selection Using Line Histograms

The line histograms generated in the line analysis consist of two types of data which are the short line and the long line histograms. In this experiment, we divided the analysis into three sub experiments:

1. Experiment Line 1 : observing long lines only with all directions. In this analysis, all 72 data values in the long line histogram are observed.
2. Experiment Line 2 : observing long lines with merged directions. In this analysis, only 24 directions are observed. The 72 bins in the long line histogram are merged into 24 bins by summing 3 bins for each direction.
3. Experiment Line 3 : observing short and long lines with merged directions. In this analysis, 24 bins are used in the short line histogram and 24 bins in the long line histogram.

Using data from all directions means all 72 data/value are used. Each data item in the line histogram indicates the percentage of pixels with edge directions in each 5 degree range. Merged direction refers to merging the data to generate a wider orientation direction.

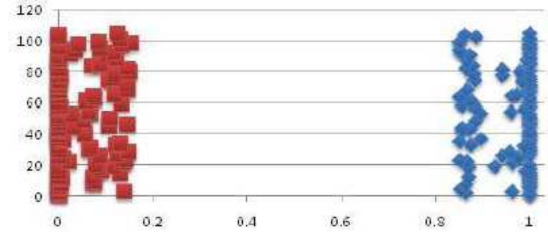
Table 3 shows examples of the building and non building image prediction generated by the Inference Engine. The prediction values range from 0 to 1. If the image prediction value is close to 1, it shows high probability that the image is a building image. If the image prediction value is close to 0, it shows high probability that the image is a non building image.

Table 3: Prediction Values for building and Non building Images. (Prediction value is Probability of images being building images). Value generated by the Inference Engine for Experiment Line 1.

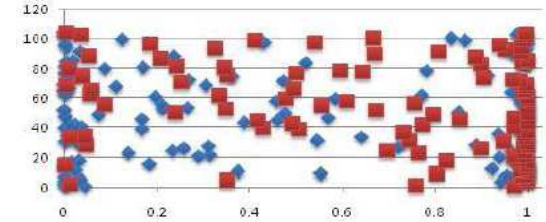
Building Image ID	Prediction Value	Non Building Image ID	Prediction Value
1	1.0000	1	0.0006247
2	1.0000	2	1.04E-13
3	0.8629 3	3	1.96E-07
4	0.9607 4	4	0.1375
5	5 0.8539 5	5	1.30E-16
….	….	….	….
105	1.0000 105	105	4.81E-13

Full results for line histogram analyses are illustrated in the form of image distributions in Figure 3. The x-axis shows a cut-off point / threshold value candidate, and the Y-axis show number of building and non-building images in the training set. Building images are represented in blue boxes while non building images are represented in red boxes.

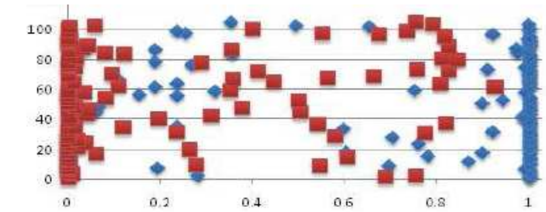
Figure 3(a) shows an excellent classification. There is a clear separation between building and non building images. Building images are well classified with cut-off point 0.8 and above, while non building images are classified using cut-off point 0.2 and under. Figure 3(c) shows a good classification. Most of the building images fall close to 1, while there are a few images distributed in the range of 0 to 0.8, and vice versa. Most of non building images are distributed close to 0, while a few images are distributed in range of 0.2 to 1. Figure 3(b) shows a bad classification. Most of building images are distributed evenly at each end. For non building images, most of them are distributed close to 1, indicating these images have low probability to be identified as non building images, but high probability to be identified as building images.



(a)Experiment Line 1 by observing long Lines with 72 Directions



(b) Experiment Line 2 by observing long lines with 24 directions each.



(c) Experiment Line 3 by observing short and long lines only with 48 directions.

Figure 3: Building Image and non building image distributions using line histograms.

The classification performances for each experiment are analysed to identify the optimum threshold value to identify building and non building images by using Confusion Matrix and ROC Curve data analysis. The Confusion Matrix is used to display actual and predicted classification by the image classifier. The entries in the Confusion Matrix have the following meaning in the context of our study:

Table 4: Data Matrix

Actual\Predicted	Negative (Non Building)	Positive (Building)
Negative (Non Building)	A	c
Positive (Building)	B	d

The entries in the confusion matrix have the following meaning:

- *a* is the number of correct non building images identified as non building.

- *b* is the number of incorrect building identified as non building
- *c* is the number of incorrect non building images identified as building
- *d* is the number of correct building images identified as building

There are six standards indicators that can be generated from the Confusion Matrix which are True Positive (or Sensitivity), Positive Fraction, True Negative (or Specificity), False Negative Fraction, Accuracy and Precision. Two indicators are used to evaluate the image classification performance, the Sensitivity and the Specificity.

1. The True Positive Fraction or sensitivity (TP) is the proportion of positive cases that are correctly identified, as calculated using the equation:
   TP (Sensitivity) = d/(b+d)
2. The True Negative Fraction or specificity (TN) is defined as the proportion of negatives cases that are classified correctly, as calculated using the equation:
   TN (specificity) = a/(a+c)

Figure 4 shows ROC curves produced from Sensitivity and 1-Specificity for Line 1,2 and 3 . The ROC curves are generated by plotting 1- Specificity on X-axis and Sensitivity on Y-axis. The ROC curve allows visual representation analysis of the trade off between Sensitivity and Specificity.

Figure 4 shows the classifier has achieve perfect classification for *Experiment Line 1* and a good classification in *Experiment Line 3*. Nevertheless, for *Experiment Line 2*, the negative curve shows most of the building images and non building images are distributed at the wrong end of the classification groups. Based on the calculations and visualization given by the Confusion Matrix and the ROC Curves, the optimum threshold values for classifying building images and non building images are identified. The optimum threshold value to identify building images is at cut-off point 0.8, while 0.2 for non building images. Therefore, images with prediction value of 0.8 or higher would have a greater possibility to be identified as building images and images with prediction value 0.2 or lower would have a greater possibility to be identified as non building images.

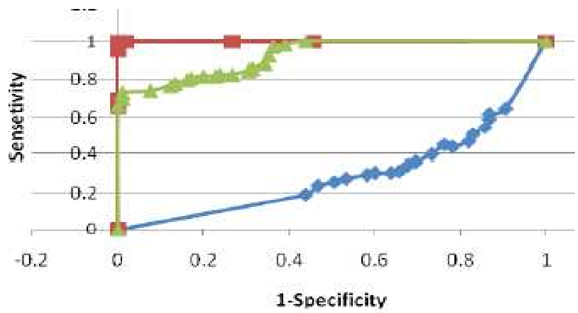


Figure 4: The ROC for Line Analyses. Experiment Line 1, 2 and 3 are presented in line colour red, blue and green respectively.

Selecting the optimum threshold value is a trade off task between Sensitivity and Specificity. The threshold values are selected based on results from Experiment Line 1. The table shows that all the building images are identified at cut-off point 0.8, and all non building images are correctly identified at cut-off point 0.2. Similar threshold values can be agreed by observing results in Experiment Line 3.

3.2.2 Experiment 2: Optimal Threshold Selection by Observing Colour Histograms

In this, colour histograms are divided into three sub experiments:

1. Experiment Colour 1 : observing 3D colour histograms by using 64 bins (RGB Colour Space is divided into 4x4x4 cells)
2. Experiment Colour 2 : observing 3D colour histograms by using 216 bins

(RGB Colour Space is divided into 6x6x6 cells)
3. Experiment Colour 3 : observing 3 colour histograms (Red, Green and Blue) by using 64 bins for each colour separately (3x64 = 192 bins)

The experimental processes are similar to Experiment 1. Figure 5 shows ROC curves produced from Sensitivity and 1-Specificity for Colour 1, 2 and 3 . Perfect classification was achieved by Experiment Colour 2 and 3. Colour 1 has produced a negative curve as most of the building images have prediction values lower than 0.5 and non building images are higher than 0.5. The optimum threshold values selected for this experiment were at 0.8 for building and 0.2 for non building images. The values are selected based on the Colour 3 results.

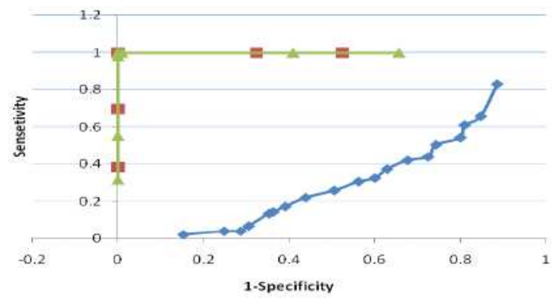


Figure 5: The ROC for Colour Analysis. Experiment Colour 1, 2, and 3 is presented in line colour blue, red and green respectively.

3.2.3 Experiment 3: Optimal Threshold Selection by Integrating Line and Colour Histograms

In Experiment 3, both line histograms and colour histograms are used in the classification.

The experiment is divided into six sub experiments:

1. Training 1: Experiment Colour 1 with Line 1
2. Training 2: Experiment Colour 1 with Line 3
3. Training 3: Experiment Colour 2 with Line 1
4. Training 4: Experiment Colour 2 with Line 3
5. Training 5: Experiment Colour 3 with Line 1
6. Training 6: Experiment Colour 3 with Line 3

Figure 6 shows ROC curves produced from Sensitivity and 1-Specificity for Training 1, 2, 3, 4, 5 and 6. All of the trainings runs have produced perfect classifications. Earlier in the colour analysis experiment, the colour histogram with a small number of bins has produced a negative result. Training 1 and 2 shows integrating the colour histogram (small number of bins) with the line histograms could improve the results. The optimum threshold values selected for each test are presented in Table 6.

Table 6: Thresholds values identified for building and non building selected based on trainings results.

Training	Threshold value for Building	Threshold Value for Non Building
Training 1	≥ 0.85	≤ 0.15
Training 2	≥ 0.85	≤ 0.20
Training 3	≥ 0.85	≤ 0.15
Training 4	≥ 0.85	≤ 0.15
Training 5	≥ 0.80	≤ 0.20
Training 6	≥ 0.85	≤ 0.20

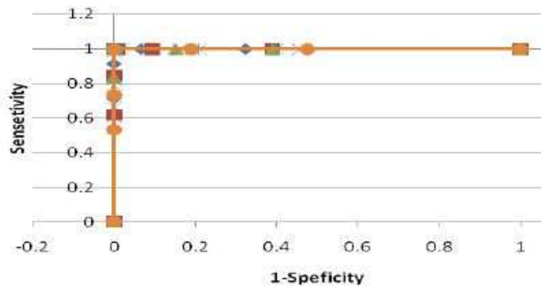


Figure 6: ROC Curves for integrating line and colour histograms integrations. All the training sets have produced a perfect classification.

4 CLASSIFYING BUILDING IMAGES AND NON BUILDING IMAGES: RESULTS AND DISCUSSIONS

This section presents image classification results for integrating line and colour analysis. The test set consists of Test Set 1 and Test Set 2. Each of the test sets consists of 534 building images and 506 non building images respectively. The classifications were done based on the thresholds value identified in training set analysis.

The image classifications results for integrating line and colour analyses are illustrated in ROC curves Figure 7. The experiments are divided into six test sets:

1. Test 1: Colour 1 with Line 1
2. Test 2: Colour 1 with Line 3
3. Test 3: Colour 2 with Line 1
4. Test 4: Colour 2 with Line 3
5. Test 5: Colour 3 with Line 1
6. Test 6: Colour 3 with Line 3

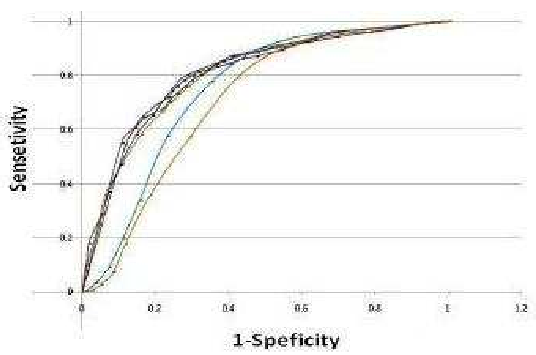


Figure 7: ROC Curves illustrating image classification performance results by integrating Colour and Line Analysis.

The ROC curves show all tests have produced positive results. Test 1, Test 2, Test 3 and Test 4 have produced similar curves, indicating performances for these tests are similar. Tests 5 and Test 6 have produced lower curves indicating poorer performances. The quantitative analysis shows that Test 2 has identified more building and non building images than the rest of the tests which is 729 images, with 431 building images and 298 non building images. Based on accuracy rate, Test 4 has the highest accuracy rate for finding building images which is 96.74%, while Test 6 has the highest accuracy rate for identifying non building images which is 91.67%. Nevertheless, Test 6 only found less than 30 images as building and non building images, showing use of strict threshold values has increased quality in image classification, but also decreased the quantity of images that could be identified. Therefore, the best threshold value for annotating "building" on unknown images is based on Test 2, which is $X \geq 0.80$. Table 5 shows quantitative analysis result for Test 2.

Table 5: Quantitative Analysis Results for Test 2

Threshold	Building	Non Building	Unknown	Total
	$X \geq 0.8$	$X \leq 0.2$	$0.8<X<0.2$	
Test Set 1 (Building)	344	57	133	534
Test Set 2 (Non Building)	87	241	178	506
Total Identified	431	298	311	1040
Identified (%)	41.44	28.65	29.90	-
Correct (%)	79.81	80.87	-	-
Incorrect (%)	20.19	19.13	-	-

5 CONCLUSIONS AND FUTURE WORK

In this paper we have developed a classifier which can be used for automatically annotating images with a "buildings" or "non-buildings" tag. Although just a proof of concept it can be seen that other image classifiers could be developed to add additional annotations to images based on their content and hence provide enhanced retrieval facilities for those images.

In the future work, we will use the image classifier as a template to other set of images such as sunsets, beach and mountain. It will increase the capability of image classifiers to annotate more unknown images which are unsearchable by the text based search.

6 REFERENCES

1. Ames, M., Naaman, M.: Why We Tag: Motivations For Annotation In Mobile And Online Media. In CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 971--980, New York (2007)
2. Jansen, B. J., Spink, A. How Are We Searching The World Wide Web?: A Comparison Of Nine Search Engine Transaction Logs. Inf. Process. Manage., 42(1):248--263 (2006).
3. Szummer, M., Picard, R. W.: Indoor-Outdoor Image Classification. In CAIVD '98: Proceedings of the 1998 International Workshop on Content-Based Access of Image and Video Databases (CAIVD '98), pp: 42--51, Washington (1998)
4. Vailaya, A., Figueiredo, M. A. T., Jain A. K., Zhang, H.-J.: Image Classification For Content-Based Indexing. IEEE Transactions on Image Processing, 10:117--130 (2001)
5. van de Sande K, E.A., Gevers, T., Snoek, C. Gm M.: Evaluating Color Descriptors For Object And Scene Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32:1582--1596 (2010)
6. Zhang, J., Marsza, M., Lazebnik, S., Schmid, C.: Local Features And Kernels For Classification Of Texture And Object Categories: A comprehensive Study. Int. J. Computer Vision, 73(2):213-238 (2007)
7. Serrano, N., Savakis, A. E., J. Luo, J. Improved Scene Classification Using Efficient Low-Level Features And Semantic Cues. Pattern Recognition, 37(9):1773--1784 (2004).
8. Luo , J., Savakis, A.: Indoor Vs Outdoor Classification Of Consumer Photographs Using Low-Level And Semantic Features. Image Processing, Proceedings. 2001 International Conference, 2:745--748 (2001)
9. Zhang, L. Li M., Zhang, H. J.: Boosting Image Orientation Detection With Indoor Vs. Outdoor Classification. In WACV '02: Proceedings of the Sixth IEEE Workshop on Applications of Computer Vision, pp: 95--98, Washington, DC, USA (2002)
10. Payne A., Singh, S.: Indoor Vs. Outdoor Scene Classification In Digital Photographs. Pattern Recognition, 38(10):1533-1545 (2005).
11. Rahimizadeh, H., Marhaban, M. H., Kamil, R. M., Ismail, N. B.: Color Image Segmentation Based On Bayesian Theorem And Kernel Density Estimation. European Journal of Scientific Research, 26:430-436 (2009)
12. Herve, N., Boujemaa, N.: Image Annotation: Which Approach For Realistic Databases? In CIVR '07: Proceedings of the 6th ACM international conference on Image and video retrieval, pp:170--177, New York (2007).
13. Tufield, M., Harris, S., Dupplaw, D. P., Cravarthy, A., Brewster, C., Gibbins, N., O'Hara, K., Ciravegna, F., Sleeman, D., Wilks, Y., Shadbolt, N. R.: Image Annotation With Photocopain. In First International Workshop On Semantic Web Annotations for Multimedia (SWAMM 2006) at WWW2006, pp:22--26, Edinburgh (2006).
14. Hare, J., Lewis, P.: Automatically Annotating The MIR Flickr Dataset: Experimental Protocols, Openly Available Data And Semantic Spaces. In MIR '10: Proceedings of the international conference on Multimedia information retrieval, pp: 547-556, (2010)
15. Sivic J., Zisserman, A.: Video Google: A Text Retrieval Approach to Object Matching in Videos. Computer Vision, IEEE International Conference on, 2:1470--1477 (2003).
16. Oliva, A., Torralba, A., Dugue, A. G., Herault J.,: Global Semantic Classification Of Scenes Using Power Spectrum Templates. Proceedings of the Challenge of Image Retrieval, Newcastle. Electronic Workshop Computer Series, Springer-Verlag (1999).