A Text-Based Approach to the ImageCLEF 2010 Photo Annotation Task

Wei Li, Jinming Min, Gareth J. F. Jones

Center for Next Generation Localisation
School of Computing, Dublin City University
Dublin 9, Ireland
{wli, jmin, gjones}@computing.dcu.ie

Abstract. The challenges of searching the increasingly large collections of digital images which are appearing in many places mean that automated annotation of images is becoming an important task. We describe our participation in the ImageCLEF 2010 Visual Concept Detection and Annotation Task. Our approach used only the textual features (Flickr user tags and EXIF information) to perform the automatic annotation. Our approach was to explore the use of different techniques to improve the results of textual annotation. We identify the drawbacks of our approach and how these might be addressed and optimized in further work.

Keywords: Photo annotation, Document expansion, Feature extraction

1 Introduction

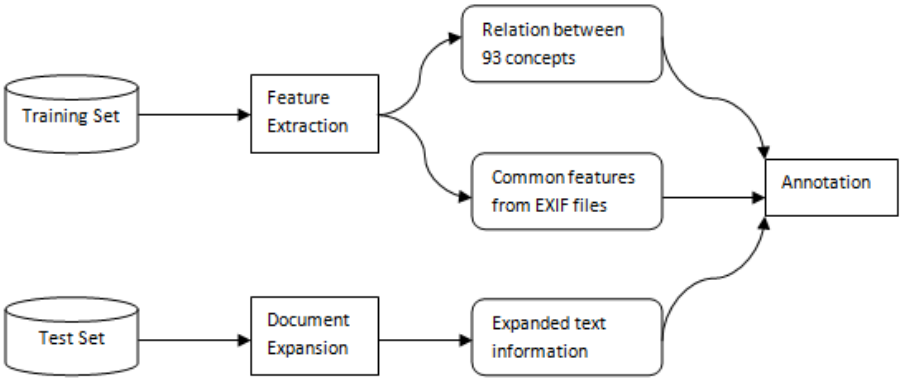
The exponential increase of images available on the World Wide Web has led to a great interest in the topics of Automatic Image Annotation (AIA) to support applications such as effective search for image collection. We describe our participation in the ImageCLEF 2010 Photo Annotation task which aims to explore methods for automatic annotation of large photo collections. The task involves assigning 93 concepts to images from the MIR Flickr 25.000 image dataset. The training and test sets consist of 8000 and 10,000 images respectively. The Flickr images in each collection include user assigned tags and EXIF data for the photos where they are present. Automatic image annotation can broadly be classified into three different approaches: visual; textual and hybrid models. In our work for ImageCLEF 2010 we concentrated only on use of text metadata for this task.

We submitted one run for the annotation task. The focus of our work was to attempt to exploit different methods to derive more text information from available resources to do the automatic annotation. We extracted features from the training set; used document expansion to enrich the existing text information resources; and use the ontology of concept. This paper is organized as follows: Section 2 describes our indexing and retrieval methods, Section 3 gives our experimental results and finally Section 4 concludes the paper.

2 Metadata Processing and Retrieval Strategies

Attempting to annotate images based on the available text information poses a significant challenge. Images are provided only with standard EXIF information and user tags of varying quality and scope. In fact investigation revealed that some images do not have any user tags at all. In our experiments, we investigated some different approaches to making use of the limited information which was available to capture more features from both training set and test set to assist with the annotation. These methods included document expansion and feature extraction which are introduced in the following subsections. The stages of processing and annotation are shown in Figure 1.

Figure 1. Flow diagram of image annotation approach



2.1 Document Expansion

The limitations of the text information associated with images can lead to significant problems for reliable processing of the images in applications such as search tools and classifiers. Particular problems can arise due to mismatch between the manually assigned tags when comparing individual images and between images and user queries, and due to the inadequacy of the tags assigned by users. In this experiment we attempted to enrich text information about images by using a process of document expansion [6]. In document expansion the existing text metadata for an image is used as a query to an information resource. Documents retrieved in response to the query are then processed to identify terms strongly associated with the images metadata which can then be added to the metadata, in the same manner as queries are expanded in traditional query expansion methods. For our work we use DBpedia as an external information resource for expansion of the image metadata “documents”.

We used document expansion to expand the image metadata and also the concepts which are to be used to annotate the images. Each concept usually consists of only 1 or 2 words. Thus it is hard to reliably match concepts to image metadata.

Thus it is interesting to try to expand concepts to include words related to the concept or which describe the concept. We thus hoped that after this expansion concepts could be more reliably matched to image metadata. To perform concept expansion each of the concepts was treated as a query and again applied to external DBpedia information resource. Selected expansion terms were then added to the concept.

Our document expansion method uses the Okapi feedback method. For expansion of the concepts, we assumed that the top 100 retrieved ranked DBpedia documents were relevant to the concept, we then added 10 top scoring words from the retrieved documents to the concept. For user tags a slightly more complex procedure was used. We still added 10 words to the metadata data of each image. However, since some user tags are sentences, they may contain stop words or other words which are not central to the focus of the tag. If we use the simple document expansion method which treats every word with the same weight, some stop words or other words not related to the topic of the tag may be added to user tags. To help avoid this problem, we used the document expansion method introduced in [1]. In this procedure DBpedia documents are first reduced by removing stop words and other words not likely to be significant to the document. The document expansion stage is then performed to add additional words to the image metadata. To perform the concept assignment, words in the expanded concepts and metadata documents were first stemmed, then the similarity between each expanded image tag and concept was computed to perform the annotation.

While this approach has the potential to assign good concept annotations for images which have manual tags to seed the expansion process, it does not work well for images which have do not have manual tags as a starting point for expansion. In order to be able to annotate these images another method is required.

2.2 Feature Extraction

The annotation scheme has been set up in such a way to make it easy to extend it with new keywords without having to go through all images again [2]. In this part, we present a way we used to refine it. The ImageCLEF 2010 task provides 93 annotation concepts. The relation between these concepts is another useful way for us to perform the annotation.

2.2.1 Find affiliation between concepts

From the training set, some general concepts were found. They cover some proper subtopics , see Table 1. According to this relationship, if any subtopic is annotated in one photo, then its corresponding general topic will be annotated in the same photo.

Table 1. Some examples of the affiliation in 93 concepts (not all)

General Concept	Sub Concept
Sky	clouds, shadow
Water	lake, rive, sea
City_life	car, vehicle, bicycle, ship, train, airplane
Animals	dog, cat, bird, horse, fish
winter	snow
architecture	building-_sight, church, bridge

2.2.2 Find opposite relation between concepts

Besides the affiliation, another relationship was also found. Some concepts are opposite, see Table 2. This means that if one concept occurred in a photo, its opposite concept is unlikely to have occurred in the same photo. In this experiment, only two of these opposite pairs were found (the pair with '*' mark in Table 2). How to find more of these opposites is a challenge for future work in this kind of task.

Table 2. Some examples of the opponent relation in 93 concepts (not all)

Concept	Opponent Concept
Indoor	Outdoor
*day	*night
No_visual_Time	day, night
*no_person	*single_person, female, male, baby, child, teenager, adult, old_person

2.2.3 Extract features from EXIF file

For concept classification, each concept was treated as an individual classification task. For each concept, there is an annotated image collection. Find out the common features of all images in this collection from their EXIF information file. Then this common feature can be used to annotate this concept on test set.

3 Task Submission and Evaluation

We made only one submission for this task. This used all the methods introduced above to collect information which can be used to automatic annotate test dataset. The official result of this run is reported in Table 3.

For this task, 64 runs were submitted in total, only two runs chose to use the text-based approach (our submission and another from the MLKD group). Based on the

reported MAP measure, these two runs got very close results, were ranked at approximately 42 (MLKD group) and 45 (our run) out of 64 submitted runs, respectively. The best run used the hybrid approach.

Table 3. Result of Runs evaluated by MAP, EER and AUC

Submission run	MAP	Avg. EER	Avg. AUC
Text-Based Run (DCU__1277149866992_ _test_annotation.txt)	0.228428	0.450835	0.194407

For each concept, the EER (Equal Error Rate) and AUC (Area Under Curve) were calculated. The results of every concept are shown in Figure 2 (the x axis indicates the 93 concepts; the y axis indicates the Accuracy Rate). From the figure we can see that the results of our experiment are not good, indeed some concepts are not detected at all. One of the main reasons that contributes to this poor result is that text resource is not sufficient for this task. Some images do not have tags and EXIF files at all. This is a big problem when using a text-based approach to do the annotation task. Another is that both the EER and AUC evaluation methods require confidence scores of every annotated concept. However, our experiment cannot provide this score information.

Figure 2(a). EER and AUC of concept 0 to 46

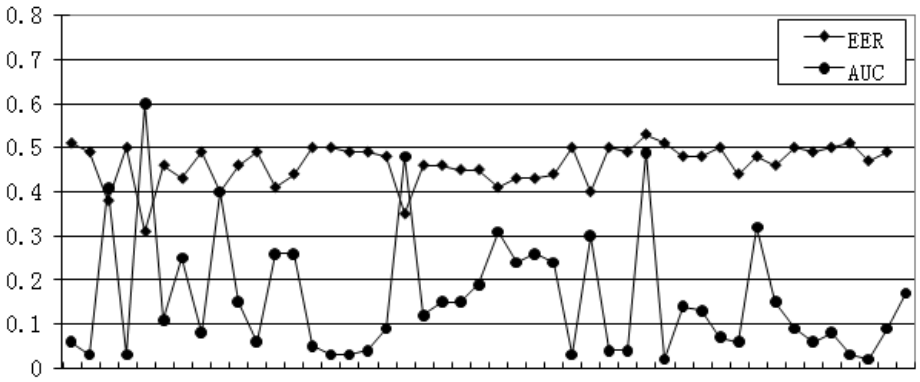
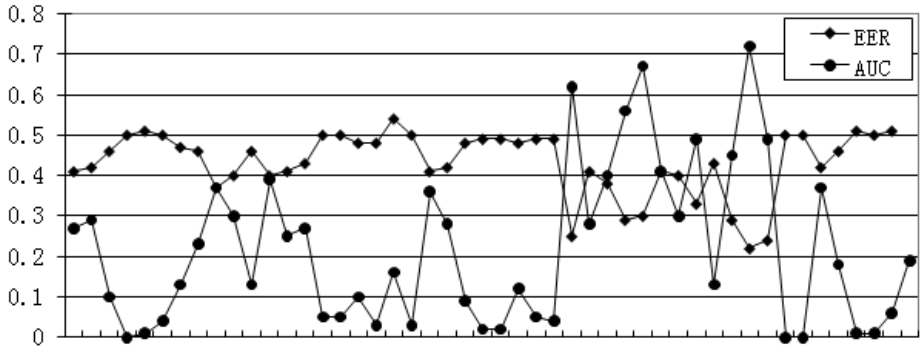


Figure 2(b). EER and AUC of concept 47 to 93



4 Conclusion

We have presented and analysed our submission to the ImageCLEF 2020 Photo Annotation Task and compared our results to those of other participants. Although the text-based approach does not get good results, it still has potential to be improved. In this experiment we use the document expansion to improve this task, we found that the external resource affects the final results a lot. In future work we will try some other external resources. The limitation of the existing text resource is another big problem for text-based approach. Some images which do not have tags and EXIF information cannot be annotated at all. So how to find more features and information from this limited resource is a big challenge for text-based approach. All of these problems define our future works.

5 Acknowledgements

This research is supported by the Science Foundation Ireland (Grant 07/CE/I1142) as part of the Centre for Next Generation Localisation (CNGL) at Dublin City University.

References

1. J. Min, J. Leveling, G.J.F.Jones: Document Expansion for Image Retrieval. RIAO conference, 2009
2. MIRFLICKR Image Collection Website. http://press.liacs.nl/mirflickr/
3. Tsikrika, T., Kludas, J.: Overview of the WikipediaMM Task at ImageCLEF 2009. In: Working Notes for the CLEF 2009 Workshop, Corfu, Greece (2009)
4. Jiquan Ngiam, Hanlin Goh: I2R ImageCLEF Photo Annotation 2009 Working Notes. ImageCLEF task 2009

5. Supheakmungkol SARIN, Wataru KAMEYAMA: Joint Equal Contribution of Global and Local Features for Image Annotation. ImageCLEF Photo Annotation task 2009 working note.
6. J. Min, P. Wilkins, J. Leveling, and G.J.F.Jones: DCU at WikipediaMM 2009: Document expansion from wikipedia abstracts. In Working Notes for the CLEF 2009 Workshop, Corfu, Greece, 30 September to 2 October, 2009.
7. A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain: Content-based image retrieval at the end of the early years. IEEE Transactions on Pattern Analysis and Machine Intelligence.
8. String Metrics Introduction. Wikipedia