UPMC/LIP6 at ImageCLEFphoto 2008: on the exploitation of visual concepts (VCDT)

Sabrina Tollari, Marcin Detyniecki, Ali Fakeri-Tabrizi, Massih-Reza Amini, Patrick Gallinari

Université Pierre et Marie Curie - Paris 6, UMR CNRS LIP6, France, firstname.lastname@lip6.fr

I Visual Concept Detection Task (VCDT)

1 Forests of Fuzzy Decision Trees (FFDT)

Learning step: a FFDT of 50 trees is constructed for each concept.
Classification step: each test image I obtains a degree $0 \leq d(I, X) \leq 50$ from each FFDT. A threshold value t is used to decide if an image presents a concept X or not: $d(I, X) \geq t$.

2 Cooccurrences Analysis

DTs learn each concept independently, but concepts can be related.

Exclusion discovery and rule
- let COOC be the concept cooccurrence matrix
- for each test image I:

- for each couple of concepts (A,B) where $COOC(A, B) \leq \epsilon$

  if $d(I, A) > d(I, B)$ then $d(I, A) = 0$ else $d(I, B) = 0$

Implication discovery and rule
- let COOCNEG be the concept cooccurrence asymmetric matrix between a concept and the negation of an other concept
- for each test image I:

- for each couple of concepts (A,B) where $COOCNEG(A, B) \leq \epsilon$

  $d(I, B) = \max(d(I, A), d(I, B))$

**Automatically discovered exclusion and implication**
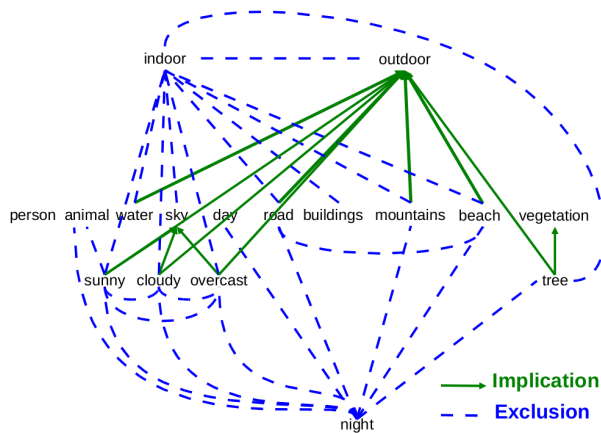
3 Experiments and Results

For each region, we compute a HSV histogram. Regions are symetrically located, because DTs are not able to automatically discover such a relation.

	Without class decision		With decision (t=25)	
	EER(AUC)	gains%	EER(AUC)	gains%
FFDT	24.6 (82.7)	-	26.2 (57.1)	-
FFDT+E.+I.	27.3 (72.0)	-11 (-13)	28.9 (53.8)	-10 (-6)

EER scores may not be adapted to boolean decisions, we use the Normalized Score (NS). The combined use of inclusion and exclusion gives the best results for any threshold.
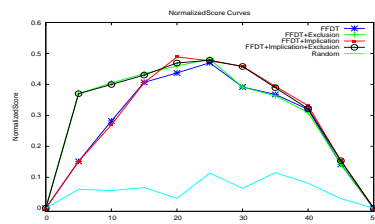
II Photo Retrieval Task (ImageCLEFphoto)

1 Using VCDT Concepts in ImageCLEFphoto

In order to answer if visual concept filtering can improve text only retrieval (based on TF-IDF or language model), we propose to use VCDT visual concepts.

```
<topic><title>animal swimming</title></topic>
```

We match VCDT concepts in the title of topics in two ways:
**VCDT filtering:** image text based ranks are filtered by the FFDT of **animal** concept
VCDTWN filtering: using WordNet, a synonym of **swimming** is *water sport*, and so image text ranks are filtered by the FFDT of **water** concept

	All 39 topics		11 topics modified by VCDT	
	P20	CR20	P20	CR20
TFIDF	0.250	0.300	0.155 (-)	0.161 (-)
TFIDF+VCDT	0.269	0.313	0.223(+44)	0.209(+30)
TFIDF +VCDTWN	filtering gives better results than TFIDF, but lower than TFIDF+VCDT			

2 Promote Diversity by Fast Clustering Visual Space

DIVVISU: We use the visual space clusters to rerank the first 50 retrieve images. If an image has the same visual space cluster as an image of highest rank, then this image is reranked at the end of the current 50 images list.
DIVALEA: As comparison point, we randomly permute the first 40 retrieve images.
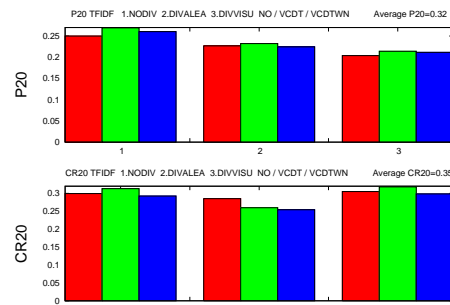
Conclusion

This study shows that there is a clear improvement, in terms of precision or cluster recall at 20, when using the visual concepts explicitly appearing in the query.

In addition...

We are part of the AVEIR consortium, which proposed the fusion of four french labs' runs and provided superior results in the ImageCLEFphoto 2008. We also submit to WikipediaMM 2008 task.